



HAL
open science

The European approach to regulating AI through technical standards

Mélanie Gornet

► **To cite this version:**

Mélanie Gornet. The European approach to regulating AI through technical standards. 2023. hal-04254949

HAL Id: hal-04254949

<https://hal.science/hal-04254949>

Preprint submitted on 23 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The European approach to regulating AI through technical standards

Mélanie Gornet

Abstract

European institutions are currently finalizing the AI Act, a new regulation on artificial intelligence. The AI Act will require manufacturers of high-risk AI systems to affix a European Conformity (CE) mark on their products. The CE mark will show compliance with the essential requirements of the AI Act and permit AI products to circulate freely in Europe. For the mark to be issued, manufacturers will have to apply harmonized technical standards (HSs). The CE mark and HSs are long-established European regulatory tools to deal with product safety, and already apply to a wide range of products, like toys, machinery or medical devices. To date, however, they have never been used to attest to compliance with fundamental rights, yet this is part of the goal of the AI Act.

This article examines the role of HSs and CE marking in the AI Act, and in particular how these product safety regulatory techniques have been expanded to cover protection of fundamental rights. This article further examines the June 22, 2023 opinion of the CJEU Advocate General in the *Public.Resource.Org* case appeal, which characterizes HSs as regulatory instruments under EU law, and how this qualification raises questions on democratic processes and transparency in standardization organizations. Finally, this article examines issues raised by the EU's choice, in the AI Act, to cover fundamental rights through HSs and the CE mark. This article concludes that unlike compliance with product safety, compliance with fundamental rights cannot be attested to by technical standards and CE marking because violations of rights are too context-specific and require a judicial determination. However, technical standards have an important role to play in encouraging best practices in AI governance, which take protection of fundamental rights into account.

1 Introduction

In April 2021, the European Commission revealed its first draft for the future regulation laying down harmonized rules on artificial intelligence (AI)¹, also known as the AI Act [29]. The text proposes a legal framework to regulate AI systems and lays down requirements that they should meet. At the time of writing, the three European institutions - Commission, Council and Parliament - are debating the content of the Act in a *trialogue* phase.

The AI Act is not the first digital law in Europe, it follows, notably, the adoption of the General Data Protection Regulation (GDPR) in 2016. However, the AI Act takes a different route from the GDPR, choosing to draw inspiration from European product safety rules. In particular, AI systems will require a conformity assessment (CA) that will be based on harmonized standards (HSs), i.e. technical specifications drawn up by European Standardization Organizations (ESOs) and possessing various legal properties, such as generating a presumption of conformity with the legislation. This conformity assessment procedure will then lead to the European Conformity (CE)

¹This work uses the term "AI" to refer both to machine learning algorithms and logic- and knowledge-based systems, in a similar way to the Council version of the AI Act [59].

marking of the AI product, a seal affixed to show compliance to regulations. However, unlike other product safety regulations, the AI Act is not only intended to protect against risks to safety, but also against adverse effects on fundamental rights. Consequently, HSs and CE marking could also apply to the protection of fundamental rights. This extension of product safety approach to fundamental rights is new, and raises difficult questions that this article attempts to address.

This article starts by laying down, in Part 2, the structure of the AI Act and how it makes use of the product safety regulatory approach, in part to protect fundamental rights. Part 3 looks in more detail at the status of HSs in EU law, and shows that although they are considered legal acts, their scope is intended to remain technical, i.e. outside the realm of political judgment. Finally, the article highlights, in Part 4, the shortcomings of the application of HSs and CE marking to the protection of fundamental rights, as well as the legitimacy problem faced by ESOs.

2 The AI Act uses product safety tools to protect fundamental rights

2.1 The AI Act has a risk-based approach

The AI Act pursues a dual objective of protecting individuals' fundamental rights² and enabling the free movement of data and AI systems within the Union. The text is structured around risk. It classifies systems based on their level of risk: *unacceptable risk*, *high-risk*, *limited risk*, and *minimal risk*. "Risk" is understood as the "combination of the probability of an occurrence of harm and the severity of that harm"³, as stated in the Parliament version (Parl. ver.) of the AI Act [61] (art. 3.1.a Parl. ver.). For limited risk systems, only transparency requirements apply; for minimal risk systems no regulatory burden applies; and systems presenting an unacceptable risk are prohibited entirely. The core focus of the Act is on high-risk AI systems, for which Annex III provides a non-comprehensive list (art. 6.2). This list can be amended by the Commission, if a new use case is found to create high risks (art. 7.1). Systems that are considered high-risk must comply with the requirements set forth in Title III, Chapter 2, in relation to risk management, data and data governance, technical documentation, record keeping, transparency and provision of information to users, human oversight, accuracy, robustness and cybersecurity. Within those requirements, risk management is a key element, particularly when AI is used in high-stakes situations [64]. Providers of high-risk AI systems must establish, implement, document and maintain a risk management system, comprising notably of the identification of known and foreseeable risks and adoption of appropriate measures to eliminate or mitigate those risks (art. 9). Residual risks must be reduced to a "reasonable" level, dictated by the state-of-the-art [41].

2.2 CE marking will show that AI systems comply with the regulation

The AI Act establishes an "ex ante" accountability framework for AI [22], in which proof of compliance with general requirements is a prerequisite for the "placing on the market or putting into service" of AI systems (art. 2).

²The Commission's explanatory memorandum (Expl. Memo.), presented just before the text of the AI Act and constituting an important aid to the interpretation of the legislation, contains a list of rights whose protection should be enhanced by the AI Act (section 3.5). It includes, for example, the right to human dignity, respect for private life and protection of personal data, non-discrimination, equality between women and men, freedom of expression, freedom of assembly, right to an effective remedy and to a fair trial, rights of defence and the presumption of innocence, the general principle of good administration, etc.

³Note that a similar definition is given in the General Product Safety Regulation [12]

The AI Act is inspired by European product safety rules, based on the so-called New Legislative Framework (NLF). The rules applicable to products under the NLF are explained in an official Commission publication, the Blue Guide [30]. Under the NLF, European legislation (directives and regulations) does not directly define technical specifications, but rather sets out the "essential requirements" that products must meet, leaving manufacturers some flexibility as to the means of achieving compliance [23]. For a product covered by a NLF legislation to enter the European market, it must be CE marked⁴. CE marking has a dual use: it allows consumers to benefit from the same level of (presumably) high protection throughout Europe [25], and allows the free movement of products within Europe by harmonizing legislation. Products bearing the CE mark can be traded in Europe without restrictions [25]. Before development of the CE mark, trade was limited by differences in national product requirements between member states [47].

Manufacturers are responsible for CE marking. They must check the applicable European legislation and ensure their products meet the essential requirements. They must then carry out the conformity assessment, set up the technical file, issue the EU declaration of conformity, and affix the CE mark to the product [27]. The AI Act stipulates that high-risk AI systems must undergo a conformity assessment procedure and, when they are found to be compliant, providers must draw up a EU declaration of conformity and affix the CE mark on the product (art. 19).

2.3 Harmonized standards will provide a technical means of assessing compliance

In the field of product safety, HSs [26] define the technical requirements that would enable a product to comply with the essential requirements set out in a specific product directive or regulation. EU legislation sets what goals to reach, and HSs define how to reach them [48]. A harmonized standard is only one possible way to comply with a legal requirement (Blue Guide, p.50), and is thus indented to be voluntary like any other standard (art. 2(1) [10]), but it is in practice the most important pathway for compliance.

HSs are developed by one of the three European Standards Organisations (ESOs): the European Committee for Standardization (CEN), the European Committee for Electrotechnical Standardization (CENELEC), or the European Telecommunications Standards Institute (ETSI). If a directive or regulation needs to be supported by HSs, the European Commission issues a standardization request to one or more ESOs, describing the main topics the standards should cover⁵. Once the standards have been drafted by the ESOs and approved, they are generally published in the Official Journal of the European Union (OJEU)⁶

HSs are, in this context, seen as a way to operationalize mandatory requirements (Expl. Memo., p.6) while reducing costs (Expl. Memo., p.7). Recital 61 of the AI Act further states that "standardisation should play a key role to provide technical solutions to providers to ensure compliance". The AI Act also explicitly requires providers to consult HSs, for instance to determine risk management measures (art. 9.3), for record keeping (art. 12.1), or for quality management (art. 17.1(e)). Some experts therefore believe that it is in standardization that the real rule-making will occur [70].

In December 2022, the Commission issued a draft standardization request to ESOs regarding standards for the AI Act [32], asking them to cover ten subjects related to the requirements for high-risk systems: risk management system for AI systems, governance and quality of datasets

⁴CE marking is applicable throughout the European Economic Area (EEA).

⁵Not all standards developed by ESOs, are HSs, only those following a request from the Commission (Reg. 1025/2012 [10], art. 2(1)(b)&(c)).

⁶Not all harmonized standards are cited in the OJEU. Some might be requested by the European Commission to address standardization gaps, without supporting a specific legislation [48]

used to build AI systems, record keeping through logging capacities by AI systems, transparency and information provisions, human oversight of AI systems, accuracy specifications for AI systems, robustness specifications for AI systems, cybersecurity specifications for AI systems, quality management system for providers of AI systems, including post-market monitoring process, and conformity assessment for AI systems. These topics correspond to the requirements for high-risk AI systems set out in Title III, Chapter 2 of the Act. ESOs are now working on HSs for these topics, as well as other topics, at their own discretion.

2.4 Private organizations will draft harmonized standards

European and international standardization organizations are private associations that are tasked to develop technical standards. They are composed of experts which have signed a service contract with a national standardization bodies. Experts can come from private companies, research institutes, public establishments, or work on their own behalf. Anyone can apply to join an national standardization body to take part in standards development and committee voting, in exchange for membership fees, paid by the expert's home company. Once experts are part of their national standardization body, they can ask to join the working groups at European or international level. This includes the three ESOs and the three international standardization bodies: the International Organization for Standardization (ISO), the International Electrotechnical Commission (IEC) and the International Telecommunication Union (ITU).

The Vienna and Frankfurt agreements between CEN and ISO, and CENELEC and IEC respectively, facilitate the exchange of information between the organizations and avoid duplication of work [51, 40]. This collaboration extends to the adoption of standards, since ISO and IEC standards can be incorporated into the catalogue of European standards by ratification by CEN-CENELEC. At present, almost 33% of CEN publications come from ISO, and 73% of CENELEC publications come from IEC. As far as HSs are concerned, ISO and IEC standards take precedence where they exist, unless it can be proved that the Commission's request cannot be met by standards issued by these international bodies [33]. This collaboration makes the composition of international standards organizations even more relevant to European issues, since their standards are likely to become HSs.

Additionally, the largest group of ISO stakeholders is the industry [55]. This composition gives standardization organizations access to beneficial industrial expertise [54], an essential competence for the development of technical requirements related to product safety.

2.5 The AI Act takes standards into the realm of fundamental rights protection

The Commission insists on its desire to integrate ethical considerations into the supervision of AI systems. The explanatory memorandum to the proposed AI Act states that the proposed essential requirements are inspired by the Ethics Guidelines of the High-Level Expert Group on AI (HLEG) [60]. The amendments proposed by the European Parliament go even further, stating that the text should ensure a high level of protection "in order to achieve trustworthy AI" (recital 2 Parl. ver.). The Parliament continues by stating some "general principles applicable to all AI systems" (art. 4 Parl. ver.), directly taken from the seven key requirements⁷ set out by the HLEG.

Some of the "general principles" set out by the Parliament touch directly upon fundamental rights, such as "transparency" or "diversity, non-discrimination and fairness", which relate to the fundamental rights to information and non-discrimination. The explanatory memorandum also

⁷With the exception of "accountability", as it is assumed that the regulation will enable this key requirement to be enforced.

states that it is in the Union’s interest to ”ensure that Europeans can benefit from new technologies developed and functioning according to Union values, fundamental rights and principles” (Expl. Memo., p1). Another example of how fundamental rights are taken into account can be found directly in the text of the Act: a system shall be considered high-risk if it ”poses a risk of harm to the health and safety, or a risk of adverse impact on fundamental rights” (art. 7). This list is extended by the Parliament, which also considers harms to ”the environment, democracy and the rule of law” (art. 7.2(c) Parl. ver.).

The Parliament also introduces a new mechanism to assess trustworthiness: the fundamental rights impact assessment (FRIA) (art. 29.a Parl. ver.), inspired by the data protection and privacy impact assessments of the GDPR, and whose absence in the Commission’s version of the Act was criticized [38]. A FRIA would be mandatory for high-risk systems listed in Annex III. It would generate a list of foreseeable impacts on fundamental rights and risks of harm on marginalised persons or vulnerable groups, as well as a detailed plan describing the measures and tools to help mitigate these risks.

Although fundamental rights have already been addressed and protected by European law - the GDPR for example - the AI Act is the first attempt to integrate fundamental rights into a product safety approach, using HSs and CE marking. The Commission has recognized that standards no longer only deal with technical components, but also ”incorporate core EU democratic values and interests, as well as green and social principles” [31]. The Parliament version of the Act further invites all actors involved in the standardization process to take into account the principles for trustworthy AI (art. 40 Parl. ver.).

Despite this apparent desire to extend the scope of technical standards, the standardization request by the European Commission [32] does not expressly refer to a standard on ”trustworthiness”. Trustworthiness is rather seen as a cross-cutting theme, not being tackled in a specific standard but being a constitutive part of every standard. CEN-CENELEC, however, continues to address this topic through its working group on foundational and societal aspects of AI systems (CEN-CLC JTC 21/WG 4⁸), a European equivalent to the ISO/IEC working group on AI trustworthiness (ISO/IEC JTC 1/SC 42/WG 3⁹). Its work includes standards on ”AI trustworthiness characterisation”, ”AI-enhanced nudging” and ”competence requirements for AI ethicists professionals”, among others¹⁰, despite the absence of these topics in the Commission’s request. This shows that ESOs are free to venture beyond the strict limits defined in the Commission’s request.

Other organisations, such as the NIST or IEEE, are rushing to adopt recommendations, guidelines, or draft standards¹¹ on different aspects of trustworthy AI, including fairness, explainability, and privacy. Some of the technical documents relating to trustworthy AI focus on particular measurements, others focus on processes¹² that AI developers are supposed to implement to manage risks, including for fundamental rights. These recommendations, guidelines and draft standards on AI are not HSs, but they may influence the development of HSs for AI, either by becoming HSs like ISO standards, or by establishing themselves on the market and influencing the state-of-the-art.

⁸For the structure of JTC 21, see <https://standards.iteh.ai/catalog/tc/cen/5af9e506-b1dc-4fcd-a3af-84d65edbf2bb/cen-clc-jtc-21>

⁹For the structure of SC 42, see <https://standards.iteh.ai/catalog/tc/iso/a8b53a70-2bb4-40a8-abf1-f42dde4432c5/iso-iec-jtc-1-sc-42>

¹⁰For a complete list of published standards and standards under development, see <https://standards.cencenelec.eu/dyn/www/f?p=205:105:0::::>

¹¹See for instance the IEEE 7000 standards series: <https://standards.ieee.org/initiatives/autonomous-intelligence-systems/standards/#p7000> or the NIST risk management framework [58]. For an overview of standards related to ethics, see [46].

¹²It is worth noting that even outside of AI trustworthiness, standards are often classified as ”product” or ”process” standards [65, 53].

3 The status of harmonized standards in EU law

3.1 Harmonized standards were not originally designed to cover fundamental rights

HSs owe their legal existence to Regulation no. 1025/2012 [10] on European standardization. Regulation 1025/2012 lists the elements that can be considered as technical specifications (art. 2.4.a). The Regulation mentions environmental protection, health and safety, but does not mention ethical criteria or fundamental rights.

The NLF was intended first as a legislative instrument to bring together all the elements of product safety legislation (Blue Guide, p.12). This emphasis on safety has gradually shifted to include other criteria. The 2022 version of the Blue Guide [30] specifies, in brackets, that "environmental and health policies also have recourse to a number of these elements" (Blue Guide, p.12), but this is clearly a secondary objective of the NLF, which is above all safety-oriented. After "safety" risks, the most commonly addressed risks are health risks, and then, more rarely, environmental risks. Recently, other criteria have begun to appear in the texts on product safety. For instance, Regulation 765/2008 on market surveillance and the marketing of products [7] creates a framework to provide "a high level of protection of public interests, such as health and safety [...], the protection of consumers, protection of the environment and security" (art. 1.2). Regulation 2019/1020 [11] further states that a product should be suspended from free circulation on the market when it presents a "serious risk to health, safety, the environment or *any other public interest*" (art. 26.1(e), emphasis added). The term "any other public interest" could encompass risks to fundamental rights. However, this is never explicitly stated in the texts.

3.2 Harmonized standards have legal effects and can be considered part of EU law

In Europe, HSs create legal effects. Products manufactured in accordance with HSs benefit from a "presumption of conformity". This means that the essential requirements covered by HSs are presumed to be automatically met if the products comply with that standard. Manufacturers may then benefit from simplified conformity assessment procedures [48]. If they choose not to apply HSs, manufacturers must demonstrate by other means how the specifications they use permit products to comply the essential requirements (Blue Guide, p.55), a more challenging task than if the manufacturer simply applied a HS. The presumption of conformity afforded by HSs encourages their adoption and avoids legal claims concerning HSs when a manufacturer's position on the market is affected by these standards [63].

The legal significance of technical standards in the EU has grown, because regulations cannot be understood without their relevant standards, making them *de facto* binding [43, 39]. Some consider that the development of technical standards has entered a stage of "juridification" [63], a term taken up by the recent Opinion of the Advocate General in the Public.Resource.Org case ([5] §29). HSs are now regarded as a form of implementing acts [67].

A number of cases have involved the analysis of the scope of HSs. The Fra.bo case [2] showed that HSs can have *de facto* mandatory effects, due to the presumption of conformity granted to them that renders any other means of achieving compliance more costly and time consuming. Additionally, the CJEU held in the James Elliott case [3] that HSs form part of EU law due to these legal effects. The last case to date, Public.Resource.Org, Inc., Right to Know CLG v European Commission [4] examined whether HSs could be subject to copyright protection. After an initial ruling by the General Court, the relevance of the claim to copyright protection is currently being re-examined in

an appeal. To this end, the Advocate General, in his 22 June 2023 Opinion, conducted a detailed analysis of HSs [5].

Even if Regulation 1025/2012 [10] considers HSs to be, in theory, voluntary, as there are other ways to demonstrate compliance, in practice, it is difficult if not impossible for manufacturers to choose a different avenue. Recourse to HSs is thus quasi obligatory for economic players if they want to stay competitive [68]. Another advantage is that the presumption of conformity reverses the burden of proof, since the company does not have to prove that it complies with the legislation, as this is automatically presumed. If a manufacturer chooses not to comply with HSs, the onus is on him to prove that his product complies with the legislation, which represents a huge commercial risk that no manufacturer would take ([5] §42). As noted by the Advocate General in the Public.Resource.Org case appeal, the whole architecture of the EU standardization system presupposes that all actors use HSs (§47). Furthermore, there are no realistic alternatives, because ESOs are too focused on HSs development to propose other standards, and there is no financial incentive for other private actors to compete with them (§48).

3.3 The commercial operating mode of ESOs is at odds with the legal scope of harmonized standards

In Public.Resource.Org [4], the plaintiffs were two non-profit organizations who requested access to several HSs, referenced in the OJEU but whose full text was not public and behind a paywall. The Commission refused to grant them this access on the basis of the first indent of Article 4(2) of Regulation 1049/2001 [6]. This article lists the exceptions to the free access of the EU institutions documents, and states that access can be refused "where disclosure would undermine the protection of commercial interests [...] including intellectual properties". A first judgement was made on 24 July 2021 by the General Court, in favor of the Commission [4]. In their appeal, the organizations assert that the General Court erred in incorrectly assessing the copyright protection of HSs, since HSs are part of the law and cannot be copyrighted, and if they were allowed copyright protection, free access to the law would take precedence over copyright protection.

According to the Vademecum [28] of the European Commission, HSs are only a means to support the implementation of legislation. In the Public.Resource.Org case appeal, the Advocate General questions this claim, affirming that they are more than a simple aid and are actually an "essential tool" for the correct implementation of EU legislation (§33 and 36). One of the Advocate General's conclusions is therefore that, due to the heavy reliance of EU legislation on HSs, the effectiveness of the legislation is compromised in the absence of a publicly accessible version of these standards. HSs are indeed considered by the Advocate General to be "indispensable" for enforcing the corresponding EU legislation, thus, the public cannot exercise their rights if they do not have access to HSs (§46-47). To ensure that everyone can have the possibility to know the law and respect it, every act, including HSs, should respect the principle of transparency and right of access to documents, recognised by the Treaty on EU [9] (art. 1§2, 10.3, 11.2&3) as well as the Charter of Fundamental Rights [8] (art. 42). This is at odds with the operating mode of ESOs that usually charge for access to technical standards and keep the intellectual property of all their standards.

3.4 The Commission is responsible for political choices while the ESOs are responsible for technical choices

Today HSs are published in the OJEU under the letter L, for legislation, where previously it was published at C, for information and notice (§9). HSs therefore seem to be the equivalent of a legally binding regulation, even though they are developed by institutions - the ESOs - without any

democratic accountability. In reality, HSs are developed under the direction of the Commission.

The James Elliott case [3] found that the Commission has significant control over the procedure of drafting and considered HSs as constituting acts of the institutions of the EU. Not only does the Commission request HSs, it also supervises the drafting and adopts them. After the draft harmonized standard has been proposed by the ESOs and before publication in the OJEU, the Commission is empowered to send back the document to the ESOs for modification if the draft does not comply with the request. Ultimately, publication in the OJEU depends on acceptance by the Commission. The cycle of an HSs thus starts and ends with the Commission. This led the Advocate General in his Opinion on the Public.Resource.Org case appeal [5] to conclude that the Commission has the power to transform a preparatory document into an act that forms part of EU law (§28). The Advocate General further advises that the Commission should be seen as the institution adopting HSs and that ESOs are only preparatory bodies (§17).

The Commission itself has declared that more power needs to be transferred from the ESOs to the Commission [31]. One way of achieving this would be to allow the Commission to draw up technical solutions directly, as an alternative to the HSs drawn up by the ESOs. The AI Act acknowledge this possibility: the Commission is tasked to draft "common specifications", where HSs do not exist or are considered insufficient or when there is "a need to address specific safety or fundamental right concerns" (art. 41.1).

However, despite the Commission involvement, democratic oversight of HSs is still lacking, as neither the European Parliament nor the Member States have a right to veto standards. Additionally, the Commission's right to refuse publication of a HSs is burdened by technical limitations and human resources costs that prevent it from carrying out a comprehensive examination [37].

4 Fundamental rights and technical standards

4.1 It is hard to separate a technical question from a fundamental rights question

ANEC, the organization that defends the interests of European consumers in standardization matters, has already recognized the many difficulties involved in transposing EU fundamental rights and values into technical standards [45]. In an ideal world, technical standards should be separated from value judgments. In reality, however, it is hard to separate the two.

For instance, the concept of fairness in AI systems has several meanings, both morally, legally and technically [56]. In a general sense, fairness means "the quality of treating people equally or in a way that is right or reasonable" [1]. This relates in law to the principle of non-discrimination protected by Article 21 on the EU Charter of Fundamental Rights [8] and Article 10 of the Treaty on the Functioning of the European Union [9]. There are many technical definitions of fairness, and a system that is fair according to one definition is not necessarily fair according to another. Many definitions cannot even be satisfied at the same time [24]. For instance, the COMPAS software, used in the United States to predict the recidivism rate of criminals, has been accused of penalising African-Americans according to a certain fairness criterion [15], whereas it respected fairness according to another measurement method [57]. By defining technical formulas to measure fairness in a standard, we run the risk of choosing an approach to non-discrimination that will lead to injustice in certain situations. This example shows that a seemingly technical definition of fairness can hide a normative choice affecting fundamental rights, the kind of normative choice that generally is made by lawmakers and judges.

Another example is the NIST study on demographic differential for facial recognition [36] which displays a few "equity measures" for facial recognition systems. For all of them, error rates are

calculated for different groups of people, based on sensitive personal information like gender or ethnicity. Some measures are based on a comparison between the error rates of the two groups on which the system performs best and worst, and other measures are based on the average of all error rates, for example. The first case is, unfortunately, not very robust, and even a slight change in parameters can produce a totally different result. On the contrary, an average-based measure will be more robust but will erase the difference between groups: a system whose performance is very poor in one group but excellent in the others could end up with the same score as a system whose performance is correct in all groups. Thus, the poor performance on this one group could go unnoticed. Yet, if a system does not work well for a certain category of population, it can lead to discrimination, such as black people being wrongly accused of committing crimes because an algorithm has matched their face to that of a criminal [49].

The NIST also proposes a benchmark that evaluates the fairness of systems against their performance¹³. A manufacturer can choose to focus on optimizing their score in the given performance or fairness criteria. They can also choose which fairness metric they should improve: the benchmark includes demographic variations by false match rate (FMR) or false non-match rate (FNMR). A low FMR aims to avoid mistakes where a person is wrongly judged to be the same as in a certain image, which usually involves higher security and social stakes to avoid intrusions into a building or station, and false accusations in case of police use. A low FNMR avoids systemic rejection of certain people.

A choice of standard signals a preference for a specific logic and set of priorities [66]. Standards organize social life, and it is crucial to question what choices have been made and how they could have been made differently [66]. However, in the context of AI standards, these choices are often presented as purely technical, and therefore non value-laden choices. Moreover, by trying to define good ethical behavior in technical standards, we risk reducing ethics to a set of tools, which trivializes moral reasoning [16].

4.2 Compliance with standards can lead to ethics washing and CE marking may give citizens an unjustified sense of protection

The diversity of approaches to AI ethical development, such as the multitude of fairness measures, is likely to lead to strategic simplification choices [13]. Manufacturers will display the measure that shows that their system is free of bias and therefore fair according to them, and not the other measures showing the system is discriminatory. The introduction of these mathematical measures in a standard is likely to accentuate this trend, by giving greater legitimacy to any chosen measure included in the standard.

Additionally, the protection granted by standards is limited and having in place a risk management system will not guarantee that all possible harms have been taken into account, or that the protective measures are sufficient. For instance, respecting a mathematical notion of fairness does not guarantee that the system will not discriminate [50]. Certification to technical standards is often perceived by consumers as a guarantee of safety [35]. This is particularly true of CE marking, often regarded as the cornerstone of the European trustworthiness model, a system that European citizens have come to internalize and respect [19]. But the mark is also often wrongly understood by consumers as a guarantee of quality, when in fact it only signifies compliance with regulations. Indeed, studies have shown that it is difficult for citizens to understand what the CE mark represents [19]. Products covered by the NLF do not require pre-market approval to be sold in the EU. The CE mark therefore does not indicate that a product has been approved by a government agency or by the EU [25]. As recalled by the Blue Guide [30], CE marking is a key indicator of a product's

¹³Available here: <https://pages.nist.gov/frvt/html/frvt11.html>

compliance with EU legislation, but it is not a proof of that compliance (Blue Guide, p.64). As such, a CE marked product may also have safety flaws [71]. Several high-profile cases have involved medical devices (breast implants [69, 62] and glucose monitors [71]) that had the CE marking but which were seriously defective. In the same way as for CE marking and safety standards, it is likely that a CE marking relating to fundamental rights may be incorrectly interpreted by citizens as meaning that a given AI system respects fundamental rights.

4.3 ESOs have a legitimacy problem as regard to fundamental rights

As previously seen, standardization organizations are private law bodies, mostly led by the industry. There is also a lack of representation of certain stakeholders. Those impacted by the use of AI have no role to play in standardization or certification processes [38]. Associations representing the interests of consumers, such as ANEC, as well as those representing workers or small businesses, do not officially have the right to participate in the work of ISO and IEC. They therefore have no say in the development of these standards, even if they are to be adopted by Europe [33]. This industry-led composition also raises risks of regulatory capture¹⁴ and conflicts of interest, since industrial stakeholders are drafting the very same laws by which they will be governed.

Additionally, while a large proportion of ISO's members come from Western Europe, almost half come from elsewhere in the world, particularly Asia and North America [55]. This could create tensions, as Europe would want both to rely on the work of international standards and to adopt standards that represent European values. For instance, ANEC has called for ESOs to address EU values and "not just adopt international standards which might not reflect our values and principles" [44].

Finally, even if responsibility for issuing the HSs is shouldered in large part by the Commission, ESOs that develop the standards are governed by private law, lacking the democratic legitimacy of the Commission and the other EU institutions. However, these legitimacy concerns about private standard-setting for public regulation are often outweighed by the positive externalities associated with the existence of relevant technical requirements [33]. The legitimacy of ESOs is further challenged by the AI Act, as standards will encompass fundamental rights issues and ESOs lack the expertise to assess them [70]. The standardization request thus states that CEN-CENELEC should ensure to "gather relevant expertise in the area of fundamental rights" ([32] art. 2.1). This is necessary to ensure the relevance of technical standards with judicial norms, yet it might not be sufficient to guarantee the legitimacy of the ESOs in the establishment of EU legal acts dealing with the protection of fundamental rights.

4.4 Standards can cover fundamental rights topics if they do not try to set thresholds or evaluate trade-offs

As seen previously, standards have difficulty in addressing fundamental rights issues, and when they attempt to do so, they can lead to ethics washing and consumer deception. ANEC has already advised that HSs should not be used to define or apply fundamental rights, legal or ethical principles [45]. If standards cannot attest to respect for fundamental rights, what purpose do they serve, and what should they contain?

Let us take the example of a standard on fairness. Such a standard can be used by a company to benchmark itself against the competition and assess its own progress. If the results are good enough, the company will use the standard as a marketing tool, like the NIST benchmark for facial recognition for which companies compete to achieve the best results based on different fairness tests.

¹⁴According to [34], regulatory capture is "the process through which special interests affect state intervention".

This fosters competition between companies and encourages them to innovate [18]. A standard can also enhance transparency and redress information asymmetries [43] by presenting to users and citizens a standardized score of different performance parameters, including for fairness, thereby permitting better comparison between products. Finally, standards, such as HSs, that are linked to legal compliance obligations, provide public authorities with a uniform method for assessing compliance.

These different uses of standards hint to what they can and cannot contain. For compliance, HSs will help clarify the AI Act’s approach to risk, for instance by defining how to conduct a risk management system, or detail what elements a conformity assessment should contain¹⁵. Additionally, standards can help harmonize how to conduct an algorithmic impact assessment [21], or a FRIA as recommended by the Parliament. As regard to governance, standards can provide guidance on the structure to be put in place within the company - perhaps with a digital ethics officer or an ethics board, the competences required for this position, or the type of decisions they can and cannot make¹⁶. Product-based standards can define tools to help make better design decisions. For example, they can define all the evaluation measures known in the literature¹⁷ (paying attention to selection biases), or the technical means to avoid a system malfunction that could lead to fundamental rights violations in the long term. In short, standards can help define tools and provide a common vocabulary for comparison between products or companies. These tools can help market actors transparently compete on fundamental rights issues, showing they have responsible processes in place, and that on certain metrics, they have achieved a certain score on an issue such as fairness.

However, there are some things that AI standards should not try to do. Even when following a standard on risk management, the evaluation of risks will remain under the responsibility of the provider. A standard can therefore never say what risks are acceptable or unacceptable [42]. Fairness standards, should not say what definition of fairness should be used for a given use case¹⁸ or what the acceptable threshold of unfairness is. In case there is a trade-off to be made between fairness and performance, a standard should not say what that trade-off should be. A standard can only provide different ways of defining and measuring fairness, making sure everyone is using the same taxonomy and methodology to measure the different aspects of fairness, but will not say which aspect of fairness should be given priority, or whether a residual level of unfairness can be tolerated in a given situation.

Performance standards¹⁹ are quite common in product safety. They specify how the product is to be built, what materials are to be used, how they are to be assembled, and so on. They also specify the tests the product must meet, such as the exact temperature or pressure it must withstand. In product safety, it is not unusual for a standard to define a threshold, for example a level of resistance to fire, or the error rate of a safety component for machinery. These standards are, however, nearly impossible to establish today for AI systems due to their probabilistic nature, which makes their reaction to certain tests highly dependent on the situation, the data on which the system has been trained, etc. This is even truer for standards that have a direct impact on people’s fundamental rights, such as fairness standards. Setting a threshold for these measures would be like setting a threshold for the level of discrimination that may be accepted: it is neither a universal decision, nor something acceptable from a legal standpoint. Setting a fairness threshold

¹⁵These topics are notably present in the standardization request [32].

¹⁶Like for instance the standard on ”competence requirements for AI ethicists professionals” that is being prepared by CEN-CENELEC.

¹⁷Like the ISO standard on bias mitigation [52], which lists all the means known in the literature for assessing and dealing with bias.

¹⁸This includes both the metric used and the population groups on which the system is evaluated.

¹⁹Following [14] terminology, also referred to as quality standards [17].

could also be abused by claiming that a system is "fair enough", without any concern for improving fairness further [20]. Whether a fairness score is acceptable, or is the right metric to be using in this situation, should remain outside of standards and determined by the regulator and judge.

As thresholds cannot be set for standards relating to the protection of fundamental rights, the development of HSs on these subjects for the purpose of assessing compliance with the AI Act seems like a difficult - and not necessary desirable - task. Because of their legal effects, HSs will always aim to set thresholds. But outside of the safety realm, HSs are less suitable, as they cannot define what is an "acceptable" level of protection to fundamental rights. This article therefore invites standardization actors to develop standards, whether HSs or other standards, which contribute to the protection of fundamental rights through the dissemination of good practices, but which avoid making value-laden societal judgments.

5 Conclusion

This article shows the AI Act's attempt to operate at two levels : ex ante compliance, inspired by product safety rules with the use of HSs and CE marking, and the protection of fundamental rights. It examines recent case laws that has determined the role of HSs in European law, as well as the June 22, 2023 Advocate General's Opinion in the Public.Resource.Org case appeal [5], which found that HSs are to be regarded as EU legal acts and that, while the Commission is to be held responsible for the political dimension of HS, the ESOs are responsible for the technical content.

However, product safety tools such as HSs and CE marking are not meant to cover fundamental rights. Standards on fundamental rights would be both difficult to establish and could lead to ethical washing and consumer deception. The field of expertise of ESOs, made up mainly of industrial experts, is not that of fundamental rights, and they could face a legitimacy problem if they tried to take on this role reserved for legislators and judges. This does not mean, however, that standards cannot address fundamental rights, as they still have an important role to play in encouraging best practices in processes and measurement techniques, but they can never attempt to decide on a trade-off or on a level of acceptability of a given fundamental right risk.

The AI Act approach calls into question the very nature of standards and their limits. It might also pose problems for the interpretation of standards by the courts, as in the past the boundaries between the technical and legal worlds were well defined, whereas today there is a certain overlap. In this context, even more than in the case of safety standards, ESOs will have to account for the power they hold. The HSs to be developed in support of the AI Act will set the tone for future regulations in the field of digital law. Europe should, however, be cautious about the power it grants to HSs, particularly if they continue their foray into fundamental rights.

References

- [1] fairness. *Cambridge Dictionary*. URL: <https://dictionary.cambridge.org/dictionary/english/fairness>.
- [2] Judgment of the Court (Fourth Chamber), 12 July 2012. *Fra.bo SpA v Deutsche Vereinigung des Gas- und Wasserfaches eV (DVGW) — Technisch-Wissenschaftlicher Verein*. Reference for a preliminary ruling from the *Oberlandesgericht Düsseldorf*. Case C-171/11.
- [3] Judgment of the Court (Third Chamber) of 27 October 2016. *James Elliott Construction Limited v Irish Asphalt Limited*. Request for a preliminary ruling from the *Supreme Court (Ireland)*. Case C-613/14.
- [4] Judgment of the General Court (Fifth Chamber, Extended Composition) of 14 July 2021. *Public.Resource.Org, Inc. and Right to Know CLG v European Commission*. Case T-185/19.

- [5] Opinion of Advocate General Medina delivered on 22 June 2023. Case C-588/21 P. Public.Resource.Org, Inc., Right to Know CLG v European Commission.
- [6] Regulation (EC) No 1049/2001 of the European Parliament and of the Council of 30 May 2001 regarding public access to European Parliament, Council and Commission documents. OJ L145, p.43–48, May 2001.
- [7] Regulation (EC) No 765/2008 of the European Parliament and of the Council of 9 July 2008 setting out the requirements for accreditation and market surveillance relating to the marketing of products and repealing Regulation (EEC) No 339/93. OJ L218, p.30–47, July 2008.
- [8] Charter of Fundamental Rights of the European Union. OJ C326, p.391–407, October 2012.
- [9] Consolidated version of the Treaty on European Union. OJ C326, p.13–390, October 2012.
- [10] Regulation (EU) No 1025/2012 of the European Parliament and of the Council of 25 October 2012 on European standardisation. OJ L316, p.12–33, October 2012.
- [11] Regulation (EU) 2019/1020 of the European Parliament and of the Council of 20 June 2019 on market surveillance and compliance of products. OJ L169, p.1–44, June 2019.
- [12] Regulation (EU) 2023/988 of the European Parliament and of the Council of 10 May 2023 on general product safety. PE/79/2022/REV/1. OJ L135, p.1–51, May 2023.
- [13] Ulrich Aivodji, Hiromi Arai, Olivier Fortineau, Sébastien Gambis, Satoshi Hara, and Alain Tapp. Fairwashing: the risk of rationalization. In *Proceedings of the 36th International Conference on Machine Learning*, pages 161–170. PMLR, May 2019.
- [14] Robert H Allen and Ram D Sriram. The Role of Standards in Innovation. *Technological Forecasting and Social Change*, 64(2):171–181, June 2000.
- [15] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner. Machine Bias. *ProPublica*, May 2016. URL: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- [16] Elettra Bietti. From ethics washing to ethics bashing: a view on tech ethics from within moral philosophy. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, FAT* '20, pages 210–219, New York, NY, USA, January 2020. Association for Computing Machinery.
- [17] Knut Blind. *The Economics of Standards*. Edward Elgar Publishing, 2004.
- [18] Knut Blind. The impact of standardisation and standards on innovation. In *Handbook of Innovation Policy Impact*, pages 423–449. Edward Elgar Publishing, July 2016.
- [19] Hakan Burden and Susanne Stenberg. Regulating Trust – An Ongoing Analysis of the AI Act. Technical report, RISE Research Institutes of Sweden, 2022.
- [20] Maarten Buyl and Tijn De Bie. Inherent Limitations of AI Fairness. *arXiv:2212.06495*, December 2022.
- [21] Alessandra Calvi and Dimitris Kotzinos. Enhancing AI fairness through impact assessment in the European Union: a legal and computer science perspective. In *2023 ACM Conference on Fairness, Accountability, and Transparency*, pages 1229–1245, Chicago IL USA, June 2023. ACM.
- [22] Céline Castets-Renard and Philippe Besse. Ex ante Accountability of the AI Act: Between Certification and Standardization, in Pursuit of Fundamental Rights in the Country of Compliance. In Céline Castets-Renard and Jessica Eynard, editors, *Artificial Intelligence Law: Between Sectoral Rules and Comprehensive Regime. Comparative Law Perspectives*. Bruylant, September 2022.
- [23] CEN. The 'New Approach'. URL: <https://boss.cen.eu/reference-material/guidancedoc/pages/newapproach/>.
- [24] Alexandra Chouldechova. Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments. *Big Data. Mary Ann Liebert, Inc., publishers*, 5(2):153–163, June 2017.

- [25] European Commission. CE marking. URL: https://europa.eu/youreurope/business/product-requirements/labels-markings/ce-marking/index_en.htm.
- [26] European Commission. Harmonised Standards. URL: https://single-market-economy.ec.europa.eu/single-market/european-standards/harmonised-standards_en.
- [27] European Commission. Manufacturers. URL: https://single-market-economy.ec.europa.eu/single-market/ce-marking/manufacturers_en.
- [28] European Commission. SWD(2015) 205 final of 27.10.2015, Vademecum on European standardisation in support of Union legislation and policies - Part 1 Role of the Commission's Standardisation requests to the European standardisation organisations.
- [29] European Commission. Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and Amending certain Union legislative acts. COM/2021/206 final, April 2021.
- [30] European Commission. Commission Notice - The 'Blue Guide' on the implementation of EU product rules 2022, June 2022.
- [31] European Commission. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - An EU Strategy on Standardisation - Setting global standards in support of a resilient, green and digital EU single market. COM(2022) 31 final, February 2022.
- [32] European Commission. Draft standardisation request to the European Standardisation Organisations in support of safe and trustworthy artificial intelligence, December 2022.
- [33] Pierluigi Cuccuru. Interest Representation in European Standardisation: The Case of CEN and CEN-ELEC. Amsterdam Law School Research Paper No. 2019-52, Amsterdam Centre for European Law and Governance Research Paper No. 2019-06. 2019.
- [34] E. Dal Bo. Regulatory Capture: A Review. *Oxford Review of Economic Policy*, 22(2):203–225, June 2006.
- [35] Av de Tervueren. ANEC Position Paper on CE marking "Caveat Emptor - Buyer Beware". Technical report, The European Association for the Co-ordination of Consumer Representation in Standardisation (ANEC), 2012.
- [36] David L Duewer. Face Recognition Vendor Test (FRVT) Part 8: Summarizing Demographic Differentials. Technical Report NIST IR 8429, National Institute of Standards and Technology (NIST), July 2022.
- [37] Martin Ebers. Standardizing AI - The Case of the European Commission's Proposal for an Artificial Intelligence Act. *The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics*, August 2021.
- [38] Lilian Edwards. Expert opinion. Regulating AI in Europe: four problems and four solutions. Technical report, Ada Lovelace Institute, March 2022.
- [39] Michelle Everson, Giandomenico Majone, Les Metcalfe, and Adriaan Schout. The Role of Specialised Agencies in Decentralising EU Governance. Technical report, Report Presented to the Commission, 1999.
- [40] European Committee for Electrotechnical Standardization (CENELEC). CENELEC Guide 13 FAQ. Frequently Asked Questions on the Frankfurt Agreement. Edition 1, 2017.
- [41] Henry Fraser and José-Miguel Bello y Villarino. Acceptable Risks in Europe's Proposed AI Act: Reasonableness and Other Principles for Deciding How Much Risk Management Is Enough. *European Journal of Risk Regulation*. Cambridge University Press, August 2023.
- [42] Henry L. Fraser and Jose-Miguel Bello y Villarino. Where Residual Risks Reside: A Comparative Approach to Art 9(4) of the European Union's Proposed AI Regulation, September 2021.

- [43] Marta Cantero Gamito. Europeanization through Standardization: ICT and Telecommunications. *Yearbook of European Law*, 37:395–423, January 2018.
- [44] Chiara Giovannini. ANEC comments on the European Commission proposal for an Artificial Intelligence Act. Technical report, The European consumer voice in standardisation (ANEC), July 2021.
- [45] Chiara Giovannini. The role of standards in meeting consumer needs and expectations of AI in the European Commission proposal for an Artificial Intelligence Act. Technical report, The European consumer voice in standardisation (ANEC), December 2021.
- [46] Mélanie Gornet and Winston Maxwell. Normes techniques et éthique de l’IA. In *Conférence Nationale en Intelligence Artificielle (CNIA)*, Strasbourg, France, July 2023.
- [47] David Hanson. *CE Marking, Product Standards and World Trade*. Edward Elgar Publishing, January 2005.
- [48] Laurens Hernalsteen and Constant Kohler. Drafting Harmonized Standards in support of the Artificial Intelligence Act (AIA) - CEN-CENELEC, 2022. URL: https://www.cencenelec.eu/media/CEN-CENELEC/AreasOfWork/CEN-CENELEC_Topics/Artificial%20Intelligence/jtc-21-harmonized-standards-webinar_for-website.pdf.
- [49] Kashmir Hill. Wrongfully Accused by an Algorithm. *The New York Times*, June 2020. URL: <https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html>.
- [50] Anna Lauren Hoffmann. Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. *Information, Communication & Society*, 22(7):900–915, June 2019.
- [51] ISO and CEN. Foire aux questions relatives à l’Accord de Vienne, 2016.
- [52] ISO/IEC. ISO/IEC TR 24027:2021, Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision making, November 2021.
- [53] Raphael Kaplinsky. The Role of Standards in Global Value Chains. World Bank Policy Research Working Paper No. 5396. August 2010.
- [54] Mark McFadden, Kate Jones, Emily Taylor, and Georgia Osborn. Harmonising Artificial Intelligence: The role of standards in the EU AI Regulation. Technical report, Oxford Information Labs, December 2021.
- [55] Mari Morikawa and Jason Morrison. A Survey of Participation in ISO’s International Standards Development Processes. October 2004.
- [56] Deirdre K. Mulligan, Joshua A. Kroll, Nitin Kohli, and Richmond Y. Wong. This Thing Called Fairness: Disciplinary Confusion Realizing a Value in Technology. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW, article 119), November 2019.
- [57] Northpointe. Practitioner’s Guide to COMPAS Core. Technical report, April 2019.
- [58] National Institute of Standards and Technology (NIST). AI Risk Management Framework: AI RMF (1.0). Technical Report NIST AI 100-1, January 2023.
- [59] Council of the European Union. Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and Amending certain Union legislative acts - General approach, November 2022.
- [60] Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission (HLEG). Ethics guidelines for trustworthy AI. Technical report, European Commission, April 2019.
- [61] European Parliament. Artificial Intelligence Act. Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)). P9_TA(2023)0236, June 2023.

- [62] Peter Rott. Certification of Medical Devices: Lessons from the PIP Scandal. In Peter Rott, editor, *Certification – Trust, Accountability, Liability*, volume 16, pages 189–211. Springer International Publishing, Studies in European Economic Law and Regulation, 2019.
- [63] Harm Schapel. The New Approach to the New Approach: The Juridification of Harmonized Standards in EU Law. *Maastricht Journal of European and Comparative Law*, 20(4):521–533, December 2013.
- [64] Jonas Schuett. Risk Management in the Artificial Intelligence Act. *European Journal of Risk Regulation*. Cambridge University Press, February 2023.
- [65] Gregory Tassej. Standardization in technology-based markets. *Research Policy*, 29(4):587–602, April 2000.
- [66] Stefan Timmermans and Steven Epstein. A World of Standards but not a Standard World: Toward a Sociology of Standards and Standardization. *Annual Review of Sociology*, 36(1):69–89, June 2010.
- [67] Carlo Tovo. Judicial review of harmonized standards: Changing the paradigms of legality and legitimacy of private rulemaking under EU law. *Common Market Law Review*, 55(4), August 2018.
- [68] Koos Van Elk and Rob Van der Horst. Access to Standardisation. Study for the European Commission, Enterprise and Industry Directorate-General. Final Report. Technical report, EIM Business & Policy Research, March 2009.
- [69] Barend Van Leeuwen. PIP Breast Implants, the EU’s New Approach for Goods and Market Surveillance by Notified Bodies. *European Journal of Risk Regulation*, 5(3):338–350, September 2014.
- [70] Michael Veale and Frederik Zuiderveen Borgesius. Demystifying the Draft EU Artificial Intelligence Act. *Computer Law Review International*, November 2021.
- [71] I. M. E. Wentholt, J. B. L. Hoekstra, A. Zwart, and J. H. DeVries. Pendra goes Dutch: lessons for the CE mark in Europe. *Diabetologia*, 48(6):1055–1058, June 2005.