

# Algorithmic discrimination in Spain: limits and potential of the legal framework

---

A programme of



red.es



# About Digital Future Society

Digital Future Society is a non-profit transnational initiative that engages policymakers, civic society organisations, academic experts and entrepreneurs from around the world to explore, experiment and explain how technologies can be designed, used and governed in ways that create the conditions for a more inclusive and equitable society.

Our aim is to help policymakers identify, understand and prioritise key challenges and opportunities now and in the next ten years in the areas of public innovation, digital trust and equitable growth.

Visit [digitalfuturesociety.com](https://digitalfuturesociety.com) to learn more

A programme of



red.es



## Permission to share

This publication is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/) (CC BY-SA 4.0).

## Published

September 2022

## Disclaimer

The information and views set out in this report do not necessarily reflect the official opinion of Mobile World Capital Foundation. The Foundation does not guarantee the accuracy of the data included in this report. Neither the Foundation nor any person acting on the Foundation's behalf may be held responsible for the use which may be made of the information contained herein.

# Table of contents

<b>1. Introduction</b>	<b>4</b>
<b>2. Algorithmic discrimination</b>	<b>6</b>
Algorithmic discrimination and the ethical challenges of using artificial intelligence	6
Aspects of algorithmic discrimination	9
Data protection, inadequate to tackle algorithmic discrimination	11
<b>3. The challenge algorithmic discrimination poses to European and Spanish anti-discrimination law</b>	<b>13</b>
Challenges of anti-discrimination law categories: direct and indirect discrimination	14
Problems of intersectionality and data granularity	15
Obstacles in the practical handling of algorithmic discrimination cases	16
Difficulties attributing liability for algorithmic discrimination	17
<b>4. Options for intervention</b>	<b>19</b>
Implementing transparency mechanisms	19
Regulating automated or semi-automated decision-making processes	21
Developing algorithmic discrimination policies	23
Applying technological solutions to algorithmic biases	24
<b>5. Conclusions and recommendations</b>	<b>26</b>
<b>References</b>	<b>29</b>
<b>Acknowledgements</b>	<b>33</b>

# 1. Introduction

The use of applications based on artificial intelligence (AI) is becoming increasingly widespread. Its reach into all areas of our personal lives and societies is causing mounting concerns and posing both ethical and technological challenges.

One such challenge is algorithmic discrimination and the discriminatory outcomes that can result from the use of automated or semi-automated decision-making systems and other AI-based applications.

While the regulatory framework for the use of AI is still in development, it is important to define algorithmic discrimination and evaluate how it can be addressed.

Most cases currently in the press and public discourse involve discriminatory outcomes resulting from the use of sensitive personal data such as sex or membership of an ethnic or racial group. But the existing data protection regulations do not cover all the criteria or mechanisms of discrimination. Nor do they provide measures to redress these cases.

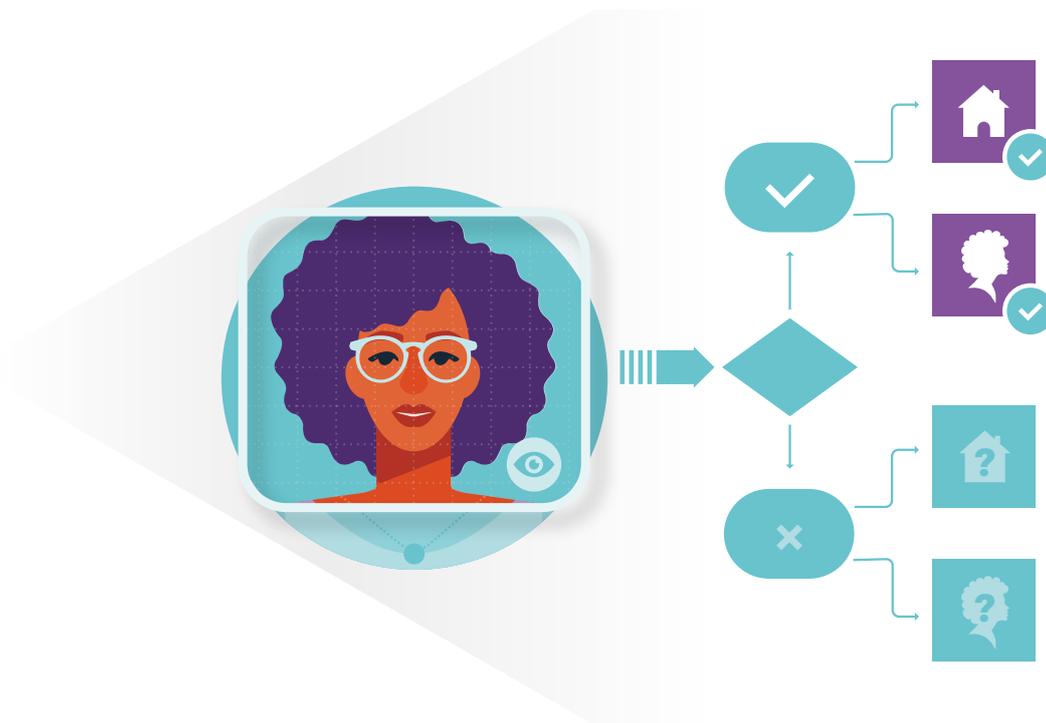
Discrimination should be addressed directly when developing regulations for the use of AI. It needs to be addressed in connection with existing anti-discrimination mechanisms under the Spanish Constitution and law, so that no aspect of it can be exempt from the principle of equality, and to prevent the proliferation of AI-based systems that ignore or infringe on fundamental rights and equality.

This report is part of a wider debate and the work of the Digital Future Society Think Tank examining the challenges to the public sector of implementing automated decision-making systems. It continues the work of understanding how these systems can discriminate against certain groups and perpetuate and exacerbate biases in societies (Digital Future Society 2020 a,b; Digital Future Society 2021).

The work arises from the particular need to understand the legal implications of these systems and whether Spain's anti-discrimination and gender equality laws can tackle algorithmic discrimination adequately. The focus of the analysis is thus on algorithmic discrimination, though it also addresses disagreements among technologists, lawyers and public policy analysts in their understanding of the problem, and the difficulties that may result.

It also examines the challenges algorithmic discrimination poses to anti-discrimination law and doctrine to reveal some of the current framework's weaknesses in providing protection against the various aspects of algorithmic discrimination. Lastly, the report proposes recommendations for both the AI regulatory framework and for a public anti-discrimination policy targeting algorithmic discrimination.

Algorithmic discrimination does not create a new problem. And it is not the result of errors occurring in systems whose development is still in progress and have a purely technological fix. Algorithmic discrimination is the result of the complex interaction between the combined use of information and communication technologies (ICTs) and AI, and the existing social dynamics of inequality, stereotypes and social hierarchies. To address algorithmic discrimination is to tackle this complexity.



## 2. Algorithmic discrimination

### Algorithmic discrimination and the ethical challenges of using artificial intelligence

Automated and semi-automated decisions are increasingly common. Both governments and private companies use AI and algorithms to support or generate decisions in various areas. AI-based solutions are applied to issues as diverse as the music we listen to, the ads we see on our computers and mobile phones, and the information we receive through them such as job offers, assessment of creditworthiness, the credibility given to complaints, and the calculation of our risk of becoming a victim of crime.

AI promises more neutral, efficient, data-driven decisions that are faster and lower-cost than traditional human-driven processes. The increasing use of AI, especially the role of AI and algorithms in decision-making, draws enthusiasm and fervour to the extent of “technological fetishism”, as well as growing distrust and scepticism due to the opacity of many of the systems and the technical difficulties of determining how they work.

An increasing number of studies shows that automated or semi-automated decision-making systems based on algorithms, both in the public and private sphere, may infringe on fundamental rights.

In February 2020, the Hague district court declared an algorithmic system called *Systeem Risico Indicatie (SyRI)* used by the Dutch government to be illegal.<sup>1</sup> The algorithm's purpose was to prevent and combat fraud against social security and the treasury. It used big data to profile individual citizens, without their knowledge, to determine their level of risk and likelihood of committing benefit fraud.

The court ruled that SyRI did not meet the requirements of proportionality and transparency. It also infringed on the right to privacy under Article 8(2) of the European Convention on Human Rights of mostly low-income people of immigrant origin, living in "problem districts" (Eguiluz Castañeira 2020).

In another case, in December 2020, a court in Italy (Bologna) ruled that Deliveroo's "Frank" algorithm, used by the company to rank its riders, was discriminatory. The algorithm judged the riders' reliability according to their absences and failure to cancel shifts on time in the system, without distinguishing the reasons. It treated all absences the same whether they were due to misconduct or were for valid reasons such as illness or exercising the right to strike (Oliva León 2021).

Ignoring the reasons behind a cancellation amounted to discrimination, the court ruled, as the algorithm unfairly penalised workers with legally legitimate reasons not to work (Ibid).

Political concern about the challenges to society of using AI is evident in the publication of standards and guidelines that address the phenomenon — from the European Commission's White Paper<sup>2</sup> to the recent European Commission Proposal on AI (AI Act).<sup>3</sup> The creation of monitoring bodies such as the European AI Committee<sup>4</sup> and the *Agencia Española de Supervisión de Inteligencia Artificial* (Spanish AI Supervisory Agency) (AESIA)<sup>5</sup> are further evidence of this.

There are studies on the challenges to the protection of fundamental rights posed by the use of new ICT and AI systems, especially in relation to personal data protection and privacy, but the challenges extend much further (the European Union Agency for Fundamental Rights and the Council of Europe 2020).

---

<sup>1</sup> Judgment of the District Court of The Hague of 5 February 2020, <https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878>

<sup>2</sup> European Commission. White Paper on Artificial Intelligence - A European approach to excellence and trust, 19 February 2020, [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_en.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf)

<sup>3</sup> 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, 21 April 2021, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>

<sup>4</sup> Articles 56—58 of the Proposal for a Regulation on AI.

<sup>5</sup> Pérez, E. Los algoritmos de las redes sociales serán controlados por una agencia pública, el Gobierno vigilará cómo se aplica la IA, 19 November 2021, <https://www.xataka.com/legislacion-y-derechos/algoritmos-redes-sociales-seran-controlados-agencia-publica-gobierno-quiere-vigilar-como-se-aplica-ia>

The uncontrolled use of AI can have harmful effects on democracies and the rule of law. So-called echo chambers that social media algorithms produce are a case in point. These have an impact on freedom of information and the right to vote (Battaglini 2019) and amplify disinformation and hate speech, undermining human dignity and the right to honour.

New generations of algorithmic tools are offering unprecedented opportunities for indiscriminate or mass surveillance, so-called surveillance capitalism (Zuboff 2020), and citizen profiling and classification.

AI systems can also have an impact on effective judicial protection and due process. Their lack of transparency shrouds the internal processes the algorithms use to generate their results. This creates uncertainty in determining the applicability and enforcement of legal provisions that protect fundamental rights, attribute liability and facilitate compensation claims.

Alongside political action, there has been an increase in scientific contributions to the field with the creation of interdisciplinary forums (ICT, engineering, law and ethics) and cross-sectoral forums (tech companies, research groups and civil society organisations).

While most academic attention has focused on the infringement of rights (such as the right to privacy or general principles of criminal law) that can result from the use, misuse or fraudulent use of technologies, the explicit problem of algorithmic discrimination is beginning to receive attention and visibility.

To date, research on discrimination and algorithms has mainly focused on cases in the US (Barocas and Selbst 2016; Gillis and Spiess 2019). Studies of the phenomenon in Europe however, are now gathering pace and being published (Hacker 2018; Zuiderveen Borgesius 2020; Xenidis and Senden 2020).

The relationship between AI and discrimination, and the impact that AI can have on inequality, the principle of equality and the principle of non-discrimination, have been examined in studies commissioned by both the Council of Europe (Zuiderveen Borgesius 2018) and the European Commission (Gerards and Xenidis 2020), as well as by international and national agencies (Orwat 2020; Défenseur des droits and CNIL 2020; Balayn and Gürses 2021).

## Aspects of algorithmic discrimination

According to the Independent High-Level Expert Group on AI, algorithmic discrimination, AI bias or algorithmic bias consists of “systematic and repeatable errors in a computer system that create unfair outcomes” from a legal point of view, such as “favouring one arbitrary group of users over others” (Independent High-Level Expert Group on AI 2020).

There are several types of bias relevant to the context of AI (NIST 2021):

- **Automation bias:** results when humans over-rely on the results or solutions generated by algorithms and as a result undervalue their own skills and judgement.
- **Legacy bias:** an AI system using biased training data will reproduce that bias, which will affect the future outcome (e.g. a CV screening algorithm that draws on gender-biased data and learns that women are less suited to a job).
- **Correlation bias:** an algorithm correlating different data sets can lead to discriminatory practices (e.g. associating gender with lower work productivity or job performance, not due to a causal relationship, but because women have historically been rated lower than men for the same job performance).
- **Amplification bias:** the ability of algorithms to process massive amounts of data at high speed means that high numbers of people can be affected in too short a time for problems to be noticed, found or mitigated. This can reproduce the stereotypes and discriminatory biases reflected in the data given to the AI system in an exacerbated and accelerated form.

Algorithmic discrimination can be associated with a number of factors.

The first is **data quality**. The use of incomplete, biased, incorrect or outdated data is a primary source of algorithmic discrimination. Some authors contend that algorithms do not actually discriminate, that they are unable to do so being no more than mathematical configurations that simply collect and process data. If the output is discriminatory, it is argued, look at the data. This is vividly expressed as “garbage in, garbage out” (Xenidis and Senden 2020, p. 157).

The differentiated impact of the prohibited grounds for discrimination — gender, racial or ethnic origin, disability and age — shown in various studies (Barocas and Selbst 2016; Eubanks 2018) is attributed to structural inequalities embedded in the data used by the algorithms.

Data quality can lead to bias at various junctures. Training data used in machine learning systems may be biased for instance. Equally, the data the system accesses may reflect entrenched social hierarchies, incorrect or inadequate representations of certain social groups, or unequal asset distribution, opportunities or burdens, etc.

The designers and developers of AI models might also introduce their own biases and prejudices. They may do so intentionally or unintentionally when preparing training samples or defining the weight given to each variable. The algorithms used in facial recognition or gender classification systems which use biased training datasets are a model example of this. They under-represent women and darker-skinned people and produce results with a higher margin of error (more than 34.4%) for non-white women compared with white men (Hao 2019).

Humans can make biased decisions based on the prejudices and hierarchies contained in the data the algorithms use. The difference is a **problem of scale, or amplification**. Because the algorithms apply decisions to more people more quickly, there is less likelihood of identifying and addressing any issues in time. Machine learning algorithms can also discriminate on mass through solutions that reproduce patterns of inequality and stereotypes embedded in the data they learn from.

Another factor is the **unpredictability of the results**. If an AI system is too vast and complex, it may not be understood by humans, not even experts, which leads to a lack of transparency and accountability in the decision-making process. This was the case, for instance, of modern deep neural networks for large-scale text data (Buomso and Park 2020).

The algorithms' behaviour may also be partially autonomous. For instance, they may operate by interpreting certain input data and using a set of pre-determined instructions, but not be confined to these, even if the system's behaviour is restricted and targeted at achieving its assigned objective.<sup>6</sup> In such cases, the possibility of understanding, controlling or predicting the operation of the algorithm is compromised, even for the system's programmers. This further complicates determining who is liable for the impact or outcome of using the system.

Finally, another related factor is **opacity, or the black box effect**. Automated or semi-automated decision-making systems may contain their programmers' unintentional prejudices and biases. If they are intentional, they can be hidden or disguised behind very complex coding. This makes it very difficult to find the root cause of the discrimination and address it.

Algorithms may produce discriminatory results for some groups not because they use a prohibited category in their coding, but due to their proxies and correlations established in the big data context. This makes it very challenging — if not impossible — to track or revise the results. Opacity, or the black box effect, involves both how the algorithm works and, very often, the quality of the data sets it processes making it harder to determine where the discrimination problem lies.

Algorithmic opacity raises both technical and legal issues. The difficulty of understanding the coding and processing means it lacks transparency, and accessibility is difficult or impossible. The models and algorithms are also protected under intellectual property and business secrecy laws. Developers and their customers are generally unwilling to make them openly available to the public.

---

<sup>6</sup> European Parliament Resolution on artificial intelligence: questions of interpretation and application of international law in so far as the EU is affected in the areas of civil and military uses and of state authority outside the scope of criminal justice. Retrieved from: [https://www.europarl.europa.eu/doceo/document/TA-9-2021-0009\\_EN.html](https://www.europarl.europa.eu/doceo/document/TA-9-2021-0009_EN.html)

This problem is illustrated by the dispute between the civil platform CIVIO and a Spanish government ministry, *Ministerio para la Transición Ecológica* (Ministry for Ecological Transition). In response to the ministry's denial of a rate subsidy to people who met the requirements to receive it, CIVIO filed a petition for information on the BOSCO system. (Power companies use BOSCO to process applicants' data and inform them whether or not they qualify for the subsidy.) The court ruled the code was protected under intellectual property rights so the program could not be checked for errors or complaints processed.<sup>7, 8</sup>

Characterising algorithmic discrimination as bias or system errors leading to unacceptable or unfair outcomes, also raises problems with the legal definition as well as the social and political meaning of discrimination. This is examined in more detail in the section on the challenges algorithmic discrimination poses to anti-discrimination law.

## Data protection, inadequate to tackle algorithmic discrimination

Most discussion of ethical challenges in the context of AI have focused on data protection and privacy. Some cases have had discriminatory impacts, but to date, these have been addressed through data protection instruments, rather than under anti-discrimination regulations. This was the case for SyRi: while the court did mention the system's discriminatory impact, it was treated as a privacy case.

Because protected personal data categories and grounds for discrimination overlap in some cases, data protection regulations can sometimes be used in cases of discrimination in the application of AI systems.

In January 2011, for instance, the Dutch Data Protection Agency ordered Charlois District Council to stop processing race-related data in making decisions on public order measures and to delete the information from its databases. In January 2020, the same agency fined TikTok €750,000 because its privacy information was only available in English, not Dutch. This was considered to infringe on the privacy of children, who are legally considered a vulnerable group.<sup>9</sup>

In these cases, there was discriminatory harm to groups protected under anti-discrimination clauses on the grounds of belonging to a racial or ethnic group, and on the grounds of age, respectively. The decisions did not examine the discriminatory impacts, however. The protection from discrimination received by the groups concerned was merely an indirect advantage of the protection of their personal data.

---

<sup>7</sup> For more information see, <https://civio.es/novedades/2022/02/17/spanish-court-blocks-the-way-for-making-public-code-open-and-transparent/>

<sup>8</sup> CIVIO has appealed the court's decision and the Consejo de Transparencia y Buen Gobierno (Council for Transparency and Good Governance) which initially sided with the ministry, has withdrawn its opposition to disclosure of the source code.

<sup>9</sup> Data Protection Enforcement in the Netherlands, <https://www.globalcompliance.com/data-privacy/data-protection-enforcement-in-the-netherlands/>

**Data protection regulations have certain limitations** and pose a number of problems rendering them inadequate for discrimination cases.

Firstly, algorithm-based decisions **can have discriminatory effects without even using personal data**. Algorithms **establish probabilistic patterns** through inferences and proxies in all kinds of mass data processing.

The assessment of a bank customer's creditworthiness may result from a correlation established between their postcode and a series of ratings of consumer habits. The algorithm may associate payment defaults or particular rental or ownership patterns with that postcode for instance. Given the ethnic, racial and class-based spatial segregation in European cities, a postcode can act as a proxy for prohibited discrimination criteria (such as race or ethnicity) or social and structural discrimination criteria (such as class or purchasing power, which are not prohibited in all anti-discrimination systems).

Secondly, data protection regulations **do not apply** to decision-making processes with discriminatory outcomes **if the personal data is processed in third-party countries not subject to the European General Data Protection Regulation (GDPR)** and the data are imported into Europe anonymised or pseudonymised (sanitised).

Finally, **there have been problems with the practical application of data protection regulations due to their breadth and degree of abstraction** (Zuiderveen Borgesius 2020). The European Parliament has pointed to the lack of resources and powers of the data protection authorities.<sup>10</sup> So it is unlikely that these regulations will have more than a sporadic impact on algorithmic discrimination.

---

<sup>10</sup> European Parliament resolution on the Commission evaluation report on the implementation of the General Data Protection Regulation two years after its application, [https://www.europarl.europa.eu/doceo/document/B-9-2021-0211\\_EN.html](https://www.europarl.europa.eu/doceo/document/B-9-2021-0211_EN.html)



### **3. The challenges algorithmic discrimination poses to European and Spanish anti-discrimination law**

Algorithmic discrimination, or the discriminatory effects that can result from the use of algorithm-based decision-making processes, poses a number of challenges to anti-discrimination law in Europe and particularly Spain.

Note that the legal definition of discrimination does not match the technical definition of algorithmic discrimination or algorithmic bias given above. The legal definition of discrimination also fails to cover the full spectrum of inequalities considered socially and politically discriminatory.

The definition of algorithmic discrimination used in technical contexts is based on the idea of bias or error (in the design of the model or as a result of poor data quality), while legal and social ideas of discrimination focus on the concept of unfair disadvantage. Consequently, the problem as raised by technologists and lawyers when referring to algorithmic discrimination may deviate. And the solutions they seek may be irrelevant to each other or divergent.

## The challenges of anti-discrimination law categories: direct and indirect discrimination

European anti-discrimination law prohibits both direct and indirect discrimination. Spanish anti-discrimination law also covers both categories — article 6 of the Organic Law on Equality,<sup>11</sup> regarding sex, and article 28(1)(b) of Act 63/2003,<sup>12</sup> which transposes the 2000 anti-discrimination directives into Spanish law,<sup>13</sup> regarding racial or ethnic origin, religion or belief, disability, age and sexual orientation.

**Direct discrimination** is when a person is, has been or would be treated less favourably than another in similar circumstances based on any of the prohibited grounds.<sup>14</sup> In EU law, prohibited grounds are sex, race or ethnicity, religion or belief, sexual orientation, age and disability. In Spanish law, the prohibition of discrimination in Article 14 of the Constitution does not give an exhaustive list of grounds. Instead, it contains what is called an open clause.<sup>15</sup> This means there may be more grounds for discrimination under Spanish law than those listed in the European directives.

In general, the category of direct discrimination is considered unlikely to be applicable or useful in cases of algorithmic discrimination (Hacker 2018; Xenidis and Senden, 2020; Gerards and Xenidis 2020). However, this is not, as is sometimes argued, because algorithms are mere mathematical configurations with no bias or intention. Unlike in US anti-discrimination law, in Europe, intention plays no role in determining discrimination, not even direct discrimination.

It is rather that automated algorithm-based decision-making processes handle large amounts of data to generate their results and the decision is unlikely to be based on one of the prohibited grounds alone. It may not even be possible to determine that (or the extent to which) a prohibited ground has been used in a decision. This is due to the complexity of the algorithmic code and the processes giving rise to such decisions — the problem of opacity, or the black box effect discussed previously.

---

<sup>11</sup> Organic Law 3/2007 of 22 March, for the effective equality of women and men: [BOE-A-2007-6115](#)

<sup>12</sup> Act 63/2003 of 30 December 2003 on tax, administrative and social measures: [BOE-A-2003-23936#A27](#)

<sup>13</sup> Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=LEGISSUM:l33114> and Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32000L0078&from=en>

<sup>14</sup> Act 63/2003 defectively transposed the European directives. Unlike the Organic Law on Equality, it does not envisage cases of past (has been) or hypothetical (could be) discrimination in the prohibition of direct discrimination on the grounds of racial or ethnic origin, religion or belief, disability, age and sexual orientation.

<sup>15</sup> Article 14 of the Spanish Constitution, "Spaniards are equal before the law and may not in any way be discriminated against on account of birth, race, sex, religion, opinion or any other personal or social condition or circumstance."

Discriminatory effects in the use of algorithms are often due to biases in the data used in both the algorithm's training phase and at later stages. The data used to train the algorithm may not be sufficiently representative of certain groups, explaining why the facial recognition systems trained largely on images of white men failed on women and non-white people — especially non-white women (Buolamwini and Gebru 2018).

The data accessed by the algorithm may also reflect stereotypes and structural discrimination against certain groups in society (e.g., data reflecting the segregation of women in different employment sectors or the pay gap), and the algorithm simply acts on, reproduces and exacerbates these biases.

Therefore, **indirect discrimination** may be more feasible for tackling algorithmic discrimination because looks at effects rather than actions.

Indirect discrimination occurs when an apparently neutral provision, criterion or practice puts a person at a disadvantage compared with others on the grounds of sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation, unless that provision, criterion or practice can be objectively justified by a legitimate aim and the means of achieving that aim are necessary and proportionate.<sup>16</sup>

To determine a case of indirect discrimination we do not need to know whether or not the algorithm used prohibited criteria in the process of reaching a solution or decision. The difficulties arising from the black box effect, such as the use of proxies and correlations, or biases in the data accessed by the algorithm, are not relevant. All we need to know is whether the solution or decision has different effects on any of the groups protected under European legislation or constitutional doctrine.

## Problems of intersectionality and data granularity

Algorithmic discrimination shares the problem of intersectionality with other areas of discrimination, but in an exacerbated form. Intersectionality refers to disadvantage from the convergence of multiple discriminatory factors or inequalities — the multiple ways black women experience discrimination compared with white women or black men, for instance (Crenshaw 1989).

In the context of algorithmic decision-making, especially machine learning, outcomes are unlikely to be based on a single protected ground. Instead, solutions tend to take into account a multitude of statistically-correlated factors and variables in large amounts of data (Hacker 2018, p. 1151).

---

<sup>16</sup> In addition to the prohibitions on indirect discrimination in Act 63/2003 and Organic Law 3/2007, in 1991 the Constitutional Court ruled that the prohibition of discrimination in Article 14 of the Constitution refers to both direct and indirect discrimination. Constitutional Court Judgment of 1 July 1991: [ECLI:EN:TC:1991:145](https://eur-lex.europa.eu/eli/tyl/1991/145)

The consequent risk is, in fact, is the opposite. The discrimination becomes more refined, more granular and highly intersectional, and extends beyond the risk from the limited number of protected categories (Xenidis and Senden 2020, p. 163). The work of Virginia Eubanks (Eubanks 2018), for instance, shows how the assessment made by algorithms creates differentiated opportunities for homeless people to access social resources, depending on their specific relationship history with social services and the particular urban area they are located in.

Even if the algorithms differentiate between men and women, by age group, or race or ethnicity, the disaggregation of the protected grounds by means of interaction with the other axes, makes it very difficult to approach a case by comparing groups who do not share the same protected grounds.

Neither European nor Spanish anti-discrimination law contains specific rules on intersectional discrimination. And there are no guidelines on how to deal with such cases. Nevertheless, debate on how to address the difficulties posed by intersectionality in anti-discrimination law has been growing over the last two decades, especially since the adoption of the anti-discrimination Directives in 2000, which multiplied the grounds for discrimination prohibited under European law.

Yet no solution has been found and there are many doubts in the specialised literature that the concept can become effective within the framework of current anti-discrimination law (Schiek 2016). The existing anti-discrimination legal framework is thus ill-equipped to deal with one of the salient features of discrimination cases in the context of AI: its high degree of intersectionality or granularity.

## Obstacles in the practical handling of algorithmic discrimination cases

Conceptually, indirect discrimination appears to be a suitable approach to the phenomenon of algorithmic discrimination. Yet the practical application of anti-discrimination protection raises a number of issues.

First, **it can be difficult to establish the disadvantage to a protected group**. While indirect discrimination does not require comparison with a group in similar circumstances, it is necessary to identify a group nonetheless.<sup>17</sup> As mentioned above, the high level of intersectionality or granularity in the categories used by algorithms to produce their results can make it difficult to identify a disadvantaged group and that group's representativeness of the protected categories.

---

<sup>17</sup> Judgments of the CJEU of 10 March 2005, Nikoloudi: [ECLI:EU:C:2005:141](#) and 17 July 2014, Leone: [ECLI:EU:C:2014:2090](#)

In other words, it may be difficult to perceive a set of people as a group if the algorithm offers, for example, less advantageous solutions to a very particular sub-group of women in certain neighbourhoods or postcodes, or in certain occupational sectors or with very specific characteristics as consumers of goods or services. It may also be difficult for that specific sub-group to be considered representative of the “sex” category.

Some other practical difficulties **stem from a lack of transparency**. Victims of indirect discrimination are often unaware they are victims. This is especially so if their circumstances are due to the application of an algorithm to a potential universe of victims who have no link or relationship with one other. This would happen if, for instance, when offering jobs through a platform, an algorithm discriminated against women or members of ethnic or racial minorities who have no way of comparing the system’s outputs with one another.

In general, victims of indirect discrimination, especially when it comes to structural phenomena of discrimination such as solutions based on biased data, are not in a position to understand the processes that disadvantage them. They also do not have access to the data or resources (knowledge, time and money) to file discrimination cases in court.

Another problematic point is that, even if the solution’s disadvantage to a protected group is demonstrated, **it is not considered discrimination if an objective justification is provided**. Much of the scientific literature considers that the opportunities offered by objective justification undermine the ability to tackle algorithmic discrimination due to indirect discrimination.

Objective justification requires that the user of an algorithm-based decision-making system show that they are pursuing a legitimate aim and that the use is appropriate and proportionate. Some authors (Hacker 2018) argue that the predictive capability of the algorithms would, in itself, function as a legitimate purpose (e.g., a mechanism for measuring labour productivity or creditworthiness) and would appear to be appropriate to that purpose. It would be difficult for judges to establish that another algorithm or solutions offered by human decision-makers are less biased or apply less discriminatory correlations.

## Difficulties attributing liability for algorithmic discrimination

In anti-discrimination law, both direct and indirect discrimination must be attributable to a party (employers, the government, service providers, etc.). Both European and Spanish anti-discrimination law prohibit instructions to discriminate and consider them discriminatory.

But in the case of algorithmic discrimination, attributing liability or making an accusation of discriminatory action, whether direct or indirect, is more problematic because many actors may be involved. These include the machine learning system itself (which acts relatively autonomously in producing solutions), the system developers and the company offering the products or services (software) to implement the automated or semi-automated decision-making systems. They clearly also include the clients who purchase those products or services and implement the system.

The Court of Justice of the European Union (CJEU) has ruled that there must be a link between employers' actions and discriminatory effects for liability to be established in a case concerning different pay conditions between direct employees and workers recruited through a temporary employment agency.<sup>18</sup>

In contrast, in Spain, both case law and legislation have alternative mechanisms making it possible to establish presumed or strict liability. For instance, the Supreme Court has established a presumption of indirect discrimination when there are discriminatory effects and there is no minimum transparency threshold in the criteria on which the decision is based (as might be the case for customers who have purchased services or products to implement automated solutions and cannot explain the criteria for adopting the solutions).<sup>19</sup>

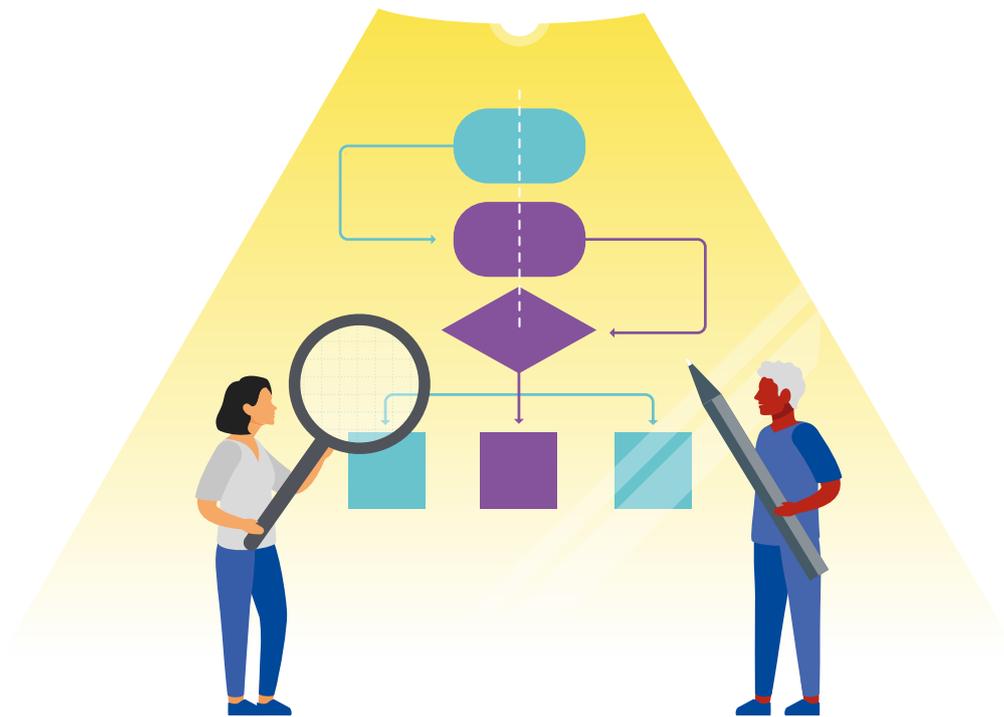
In addition to presumed liability, legislators may choose to establish forms of strict liability, i.e. attributing default liability to a particular actor, irrespective of whether the action may be attributed to that actor or not.

One form of strict liability is found, for instance, in the 2020 guidelines of the Spanish Data Protection Agency (AEPD). These state that any person making a decision on the use of AI cannot deny liability on the grounds of insufficient information or technical knowledge. Liability cannot be shifted to the developers or the company offering the AI products.

---

<sup>18</sup> Judgment of the CJEU of 13 January 2004, Allonby: [ECLI:EU:C:2004:18](#)

<sup>19</sup> Constitutional Court judgment of 18 July 2011: [ECLI:EN:TS:2011:5798](#)



## 4. Options for intervention

### Implementing transparency mechanisms

As shown, one of the fundamental problems in identifying and combating algorithmic discrimination is the opacity of AI systems and the logic followed by the companies that develop and own them. It is even difficult to find out who uses algorithm-based decision-making systems and when they do so.

One clear option for intervention is to increase transparency in the use of these technologies. Spain has adopted regulations designed to increase transparency which could serve this purpose.

Royal Decree-Law 9/2021 (the Rider's Law) introduced a new paragraph in article 64.4 of the Statute of Workers' Rights that gives the works council the right to "be informed by the employer of the parameters, rules and instructions that form the basis for AI algorithms or systems that affect decision-making that may influence working conditions and access to and maintenance of employment, including profiling".<sup>20</sup>

It entitles the workers' representatives to be informed of the use of algorithms by the employer in the case of both automated decision-making systems and when the algorithm affects decision-making (even if only humans make the decisions) (Monclús and Martínez 2021).

<sup>20</sup> Act 12/2021 of 28 September, amending the consolidated text of the Statute of Workers' Rights, approved by Legislative Royal Decree 2/2015 of 23 October, to safeguard the employment rights of people who work in delivery within the field of digital platforms: [BOE-A-2021-15767](https://www.boe.es/boe/BOE-A-2021-15767)

Article 82 of the recent amendment of the Consolidated Text of the General Consumer and User Defence Act (TRLGDCU),<sup>21</sup> includes general information on the main parameters determining the ranking of goods and services in online searches made by consumers and users. This opens up the possibility of legally demanding meaningful information from businesses about how the datasets interact and the parameters used by the algorithm.

And article 13 of the AI Act being prepared by the European Commission requires the design of high-risk AI systems to be sufficiently transparent so that users can correctly interpret and use their results.<sup>22</sup> AI systems must be accompanied by instructions for use and include concise, complete and clear information. This includes the characteristics, capabilities and limitations of the AI system's operation, such as specifications regarding input data, training data and validation data, etc., as well as the human oversight measures in place.

Another relevant legal provision, limited to the field of personal data, is article 13(2)(f) of the GDPR, which requires informing data subjects of the existence of automated decisions, including profiling. In the case of profiling, article 11.2 of Organic Law 3/2018 on the protection of personal data and safeguarding of digital rights entitles data subjects to be informed of and to oppose the adoption of automated individual decisions that have legal effects for them or significantly affect them in a similar way.<sup>23</sup> The scope of this rule is reduced by setting limits on the right to information in the form of intellectual property rights and trade secrets.

Finally, Act 19/2013 of 9 December on transparency, access to public information and good governance could easily be applied to the use of algorithm-based decision-making systems.<sup>24</sup> Article 12 of this act grants the right to access public information. This includes content or documents produced or acquired by governments and others subject to the duty of transparency in their work such as management bodies, state agencies and publicly-owned organisations or investee companies.

If automated decisions are used to distribute public funds or determine access to services, their code and logic should be subject to the same requirements set out in articles 6, 7 and 8 of the above act for policy planning and production. They should also be subject to acts with legally relevance, those with economic and budgetary implications, and any statistics required to assess the performance and quality of public services should be provided.

Nevertheless, when assessing the importance of establishing transparency requirements for the use and characteristics of algorithms in decision-making, it should not just be a matter of providing information. Instead, it should take the form of the right to explanation (Selbst and Powles 2017), meaning the information should be comprehensible by its recipients.

This right is contested (Wachter et al. 2017). It is often argued that the codes of many algorithms are not comprehensible or traceable even to AI experts. It would be advisable for rules imposing transparency requirements to also provide mechanisms to ensure the right to explanation or mitigate its absence whenever it is not technically possible to trace or comprehend the algorithm.

---

<sup>21</sup> Legislative Royal Decree 1/2007 of 16 November, approving the consolidated text of the General Consumer and User Defence Act and other complementary laws: [BOE-A-2007-20555](#)

<sup>22</sup> Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52021PC0206>

<sup>23</sup> Organic Law 3/2018 of 5 December, on the protection of personal data and the safeguarding of digital rights: [BOE-A-2018-16673](#)

<sup>24</sup> Act 19/2013 of 9 December, on transparency, access to public information and good governance: [BOE-1-2013-12887](#)

## Regulating automated or semi-automated decision-making processes

The explicit regulation of automated or semi-automated decision-making processes to prevent discrimination is a major legal challenge for the European Union.

The current proposal for an AI Act, for instance, categorises AI systems according to their risk to the fundamental rights and freedoms of citizens. The principle of non-discrimination features prominently among the motivations behind the document, which goes so far as to say: "... the proposal complements existing Union law on non-discrimination with specific requirements that aim to minimise the risk of algorithmic discrimination".<sup>25</sup> Discrimination in itself, however, is not mentioned once in the proposal.

The proposal establishes a classification scheme for AI systems based on risk levels. An AI system is considered high risk based on the function it performs and the specific purpose and modalities for which the system is used. In Annex III, the proposal lists AI systems for which risks have already materialised and authorises the Commission to expand this list.

Specific prevention and mitigation obligations are set out for the different risk levels. High-risk systems have obligations related to the design of risk assessment systems, data quality, technical documentation, system logging and traceability, human oversight safeguards, and cybersecurity and system robustness.

Many organisations have criticised this approach: 61 of them wrote an open letter to the Commission calling for red lines to be drawn for the use of AI systems and for a ban on uses that reproduce structural discrimination or infringe on fundamental rights (EDRi 2021).

The risk-based approach has also been criticised for its lack of rigour (Mahler 2021). And because it does not primarily aim to manage risk. Instead, it seeks a balance, or proportionality, between the risks posed by AI systems and fundamental rights or democratic values (Access Now 2021).

This level-of-risk approach is not intended to address the risk of discrimination specifically. This is despite the fact that the list of high-risk systems in Annex III contains areas and decision-making functions in which discrimination is common (access to education or employment, migration management, access to goods and services, etc.).

In other words, the obligations set out in the proposal for AI systems used in these domains with ranking, selection and even prediction functions (probability calculation), do not take into account the specific features of algorithmic discrimination, such as the use of proxies, the granularity of categories, the difficulty of identifying disadvantage and the problem of objective justification. The latter is particularly relevant in view of the importance the proposal attaches to the idea of proportionality.

---

<sup>25</sup> Proposal for an AI Regulation, p. 4.

In addition, the Proposal for a Regulation on Digital Services (Digital Services Act) recognises that certain groups or individuals may be vulnerable or disadvantaged in the use of online services due to their gender, racial or ethnic origin, religion or belief, disability, age or sexual orientation.<sup>26</sup> They can be affected by “(unconscious or conscious) biases potentially embedded in the notification systems by users and third parties, as well as replicated in automated content moderation tools”.

This proposal aims to mitigate this risk of discrimination, including the discriminatory display of advertising with an impact on citizens’ equal treatment and opportunities. The proposal also includes the regulation of online internal complaint-handling systems and online out-of-court dispute settlement for service recipients. It also provides for audits and a supranational monitoring mechanism, with a particular focus on large online platforms.

While the proposals for European regulations underline the risk of discrimination and the impact on fundamental rights, they clearly do not establish legal anti-discrimination mechanisms. They do not even prohibit indirect discrimination or discriminatory instructions. What they do establish are mechanisms that increase transparency, traceability and access to information in AI systems, as well as some obligations regarding risk assessment and the establishment of mitigation plans.

In Spanish legislation there are certain laws that could contribute to regulating the risk of discrimination in the use of AI.

The Unfair Competition Act 3/1991, for instance, states that “discriminatory treatment of the consumer in terms of prices and other sale conditions shall be deemed unfair, unless there is just cause”.<sup>27</sup> In other words, if an automated pricing system on a website offered higher prices based on the postcodes of areas inhabited by ethnic minorities, and the company could not provide an objective justification for the price difference, it would be a case of prohibited discrimination and unfair competition (Battaglini 2020).

The General Consumer and User Defence Act classifies an offence as any discriminatory action against vulnerable consumers regardless of the reason, or against any consumer or user exercising their rights under this act, unfair trade practices, and discriminatory conduct in access to goods and services – especially regarding equality between women and men.<sup>28</sup>

While this law does not specifically criminalise relationships mediated by AI systems, the prohibition of discrimination is sufficiently broad to include them. But there is little case law for guidance in the area of discrimination in access to goods and services. The lack of case law on discrimination in this context (with studies demonstrating the incidence of the use of automated or AI-based systems (Pandey and Caliskan 2021) is largely due to the difficulties discussed above concerning victims’ awareness of being discriminated against and access to information and support systems to file complaints.

---

<sup>26</sup> European Commission. Proposal for a Regulation of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en>

<sup>27</sup> Unfair Competition Act 3/1991 of 10 January 1991: <BOE-A-1991-628>

<sup>28</sup> Legislative Royal Decree 1/2007 of 16 November, approving the consolidated text of the General Consumer and User Defence Act and other complementary laws: <BOE-A-2007-20555>

## Developing algorithmic discrimination policies

A further measure that would strengthen anti-discrimination protections is the development of public policies to control, monitor and evaluate algorithmic discrimination. This could be managed by a specialised agency (such as the proposed AESIA) or operate through coordination mechanisms among specialised agencies (such as the Anti-discrimination Office of the Employment Inspectorate, the Ministry of Equality, etc.).

There is ample room for improvement in this area as anti-discrimination law in Spain is generally not accompanied by anti-discrimination policies or monitoring and control. There are no public bodies or public-private partnerships that offer support and protection to victims of discrimination, especially victims of structural or systemic forms of discrimination.

Regarding monitoring and control, there is no policy to generate knowledge and evidence of the effectiveness of implementing anti-discrimination standards in general or on the incidence of discrimination in its various forms (both the impact on different groups of people and on the different sectors or scopes in which discrimination occurs).

One option for intervention of particular interest is algorithmic auditing.<sup>29</sup> This is precisely one of the measures to be implemented by the AESIA in future.

There are already projects such as the Observatory of Algorithms with Social Impact (OASI), developed by the pioneering company Eticas Foundation, consisting of a search engine that provides information on the algorithms used by governments and companies to make crucial automated decisions about citizens' lives. The OASI reports who developed the algorithms, who uses them, their scope of application (both geographic and sectoral), whether they have been audited, their aims, their social impact and the threats they pose (Jiménez 2021).

The OASI project is also considering the possibility of auditing the algorithms of companies and institutions without accessing their code with external algorithmic audits. These analyse an algorithm's impact after it has been implemented and would enable an analysis under the category of indirect discrimination.

---

<sup>29</sup> Auditability refers to the ability of an AI system to undergo the assessment of its algorithms, data and design processes, according to the Assessment List for Trustworthy AI (Independent High-Level Expert Group on Artificial Intelligence 2020).

In addition to the overall effect of knowledge accumulation and controllability, algorithmic auditing can also be beneficial to companies and governments that use AI systems:

- In relation to the **infringement of fundamental rights**: systematic infringement of rights may lead to civil, administrative and even criminal liability, not to mention the moral damage to citizens who have their hopes crushed and their rights and freedoms curtailed.
- In terms of **economy and efficiency**: discriminatory practices may be a financial disaster to governments or companies that implement AI systems. The scale and numerical impact of automated decision-making processes means there is a vast potential number of victims who could claim compensation or damages.
- In terms of **reputational cost**: the possibility of a company or institution being mired in a case of algorithmic discrimination may be a major marketing blow. It could also leave citizens and consumers more distrusting of the use of technologies.

## Applying technological solutions to algorithmic bias

Most proposals for intervention in algorithmic discrimination in EU documentation are technological solutions. This technological approach is driven by the view of discrimination in terms of biases resulting from poor data quality. Discrimination is seen not so much as a reflection of a structural problem of social inequality, but as a technical problem in the model design, training data, data collection and analysis, etc. It is seen as a problem that can, and should, be solved through technical debiasing solutions.

Debiasing mechanisms are methods that address biases by introducing a form of statistical parity (a so-called fairness metric) in the dataset, algorithm or results. The same approach also applies to audits, which refer to evaluation processes concerning the statistical parity of the system being examined.

Some studies already point to the limitations of this technocentric approach to discrimination in the context of AI and to debiasing solutions (Balayn and Gürses 2021). Technological solutions are criticised because the definition of their biases and problems, in both the White Paper and the proposed AI Act, is much too vague to serve as guidance. Debiasing mechanisms are not a universal panacea for any type of AI application. To date there is very limited proven effectiveness. The European documentation appears to ignore this (Ibid.).

Similarly, the technocentric approach to discrimination as data bias cannot account for structural discrimination (which may correspond statistically yet be unfair). Debiasing mechanisms are not designed to confront the problem of discrimination in a social context.

Finally, it has also been noted that technological debiasing solutions concentrate even more power in technology companies which are, ultimately, commercial actors whose primary interest is their own profit and not the important political and social issues that underlie decisions on equality and discrimination.

Reducing discrimination to biases may lead to incorrect assessment of discrimination and debiasing mechanisms are no panacea against algorithmic discrimination. But it is worth noting that AI can also contribute to monitoring and tracking the use of AI systems, establish evidence of more general discrimination and inequality, and reveal our own biases (Kleinberg and Sunstein 2019).

The Claudette system, for instance, uses machine learning to identify abusive clauses in terms and conditions of service or privacy information on online platforms (Lagioia and Sartor 2020). This system, developed collaboratively between academia and consumer organisations, also demonstrates the advantages of expanding the use of technology to actors other than Big Tech or governments, and mitigates one of the risks already mentioned in the technocentric approach.

So, as the Fundamental Rights Agency (European Union Agency for Fundamental Rights and Council of Europe 2018) and the European Parliament predict, big data and AI could also be opportunities and instruments in the fight against discrimination and in the protection of fundamental rights.<sup>30</sup>

---

<sup>30</sup> European Parliament Resolution with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies, [https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275\\_EN.html](https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html) recitals 27 et seq.

## 5. Conclusions and recommendations

Among the ethical challenges posed by the use of AI, algorithmic discrimination is beginning to draw attention for its capacity to affect fundamental values and rights in our democracies, social cohesion and policy efforts against inequality.

Algorithmic discrimination comprises distinct profiles that require targeted solutions. Data protection regulations are clearly inadequate to address the problems it poses.

A key issue for the design of anti-discrimination solutions in the context of AI is the divergence between technologists', lawyers' and policymakers' various conceptions of algorithmic discrimination. The interdisciplinary collaboration needed to find solutions to a problem as complex as this requires this divergence to be recognised and addressed. It is the only way to prevent these different understandings from being used interchangeably, leading to confusion and misguided solutions to ill-considered problems.

This becomes especially urgent in the development of laws to regulate the use and application of AI, such as the proposed European AI Act. These raise the problem of discrimination in a legal and political sense in their explanatory memoranda, but then limit their solutions to technological fixes.

There is a need for a higher degree of cross-disciplinary work in the design of technological, legal and public policy solutions.

In Spain, the anti-discrimination framework has distinct weaknesses in dealing with algorithmic discrimination. This may be an opportunity to strengthen the regulatory framework in the face of structural discrimination and inequality, by developing a genuine anti-discrimination public policy.

The regulations governing AI need to explicitly prohibit direct, indirect and intersectional discrimination. The use of AI systems, automated or semi-automated decision-making systems or other applications should not become grounds for exemption from the discriminatory outcomes that come from the use.

Some aspects of anti-discrimination law render it too limited to address algorithmic discrimination. Both the mechanisms for determining disadvantage and the formulas for resolving intersectionality issues must be reviewed.

Anti-discrimination law could be expanded to apply to other forms of structural discrimination such as inequality or segregation. Beyond this, it is necessary to develop an anti-discrimination policy that can guarantee protection that is currently not accessible to individual victims, and which the courts cannot provide without greater training and support.

Coordination among various public actors starting with the nascent AESIA is essential in the development of anti-discrimination policies. It requires the involvement of those already responsible for combating discrimination, such as the Women's and Equal Opportunities Institute, the Employment Inspectorate's new Anti-discrimination Office, and the Ombudsman, etc. The expansion and strengthening of this policy will also depend on collaboration with relevant civil stakeholders, such as the tech companies themselves, interested organisations and associations, trade unions and research centres.

This development rests on the premise of transparency: while the algorithms are not fully comprehensible to the human mind, we still need to know when and for what AI systems are being used and what impact they are having. Responsibility should also be taken for their use — models and codes should be monitored, and there should be mechanisms for certification, monitoring and impact assessment.

Finally, AI should be considered a tool, not an inevitability. When no measures of prevention or mitigation can safeguard the right to non-discrimination or prevent an infringement of fundamental rights, the use of AI systems can also be limited, made conditional or proscribed.

We propose the following **set of recommendations** to identify and address algorithmic discrimination.

- **Promote a truly integrated and multidisciplinary approach to discrimination.**  
Technological solutions to algorithmic biases should set concrete objectives and criteria and not be presented as single, all-encompassing solutions or as solutions that limit discrimination to aspects they can address (certain types of biases in certain types of AI systems). Encourage collaboration that takes into account the three-fold technological, empirical and ethical-political dimension of algorithmic discrimination.
- **Analyse the continuity between anti-discrimination legislation and AI regulation.**  
This analysis is necessary to avoid contradictions, loopholes and 'windows of exemption' from the application of the principle of equality and non-discrimination. Include explicit prohibition of direct, indirect and intersectional discrimination in legislation regulating the use of AI.
- **Establish compulsory transparency mechanisms for all public and private sector users of AI systems.** Records could be created of the AI systems used and their functions and modalities of application, similar to wage, overtime and equality plan records.
- **Establish an anti-discrimination public policy to be applied within the context of AI.**  
Such an anti-discrimination policy should fulfil at least three functions: generate knowledge and evidence of discrimination resulting from the use of AI; monitor and assess the impact of AI systems; and support and empower individuals and organisations to file complaints in cases of discrimination and infringement of fundamental rights.

- **Provide AESIA with the capacity and resources to monitor, assess and penalise cases of algorithmic discrimination.** Establish AESIA's algorithm auditing and certification functions, and encourage collaboration with public and private research and consultancy bodies. Algorithmic audits could work according to the model for pay audits. Registration and certification of AI systems could be a condition for accessing public funds or contracts, as is now the case for equality plans.
- **Coordinate the efforts of various public actors in the implementation of anti-discrimination policy.** Promote collaboration mechanisms among different authorities, institutions and public agencies to comprehensively tackle algorithmic discrimination, covering different areas: data protection (AEPD), employment (Ministry of Employment, the Employment Inspectorate's Anti-discrimination Office), AI (AESIA), and gender (Women's Institute).
- **Promote research into algorithmic discrimination.** Such research should be able to generate evidence of the impact of AI systems that can be of use to potential victims, associations and governments in claiming and securing rights. Promote AI-based research to generate knowledge of discrimination in general and develop more ethical technological solutions to counteract commercial AI systems that are more focused on profit maximisation.

## References

- Access Now. (2021). The EU should regulate AI on the basis of rights, not risks. [online] Available at: <https://www.accessnow.org/eu-regulation-ai-risk-based-approach/> (Accessed: 25 March 2022).
- Balayn, A. and Gürses, S. (2021). Beyond Debiasing: Regulating AI and its inequalities. European Digital Rights. [PDF] Available at: [https://edri.org/wp-content/uploads/2021/09/EDRi\\_Beyond-Debiasing-Report\\_Online.pdf](https://edri.org/wp-content/uploads/2021/09/EDRi_Beyond-Debiasing-Report_Online.pdf) (Accessed: 9 March 2022).
- Barocas, S. and Selbst, D. (2016). Big Data's Disparate Impact. California Law Review vol. 104. [PDF] Available at: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2477899](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899) (Accessed: 9 March 2022).
- Battaglini, M. (2019). Las próximas elecciones votarás a quien mejor te haya manipulado. ¿Crees que eres inmune? Pues no lo eres tanto. LinkedIn. [online] Available at: <https://www.linkedin.com/pulse/las-pr%C3%B3ximas-elecciones-votar%C3%A1s-quien-mejor-te-haya-manuela/?originalSubdomain=es> (Accessed: 9 March 2022).
- Battaglini, M. (2020). Cómo el Código Postal sirve para la discriminación de precios. LinkedIn. [online] Available at: <https://www.linkedin.com/pulse/c%C3%B3mo-el-c%C3%B3digo-postal-sirve-para-la-discriminaci%C3%B3n-de-manuela/> (Accessed: 9 March 2022).
- Boulamwini, J. and Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of Machine Learning Research 81:1-15, 2. [PDF] Available at: <https://proceedings.mlr.press/v81/boulamwini18a/boulamwini18a.pdf> (Accessed: 9 March 2022).
- Buomso, K. and Park, J. (2020). Transparency and accountability in AI decision support: Explaining and visualizing convolutional neural networks for text information. ScienceDirect. [PDF] Available at: <https://www.sciencedirect.com/science/article/abs/pii/S0167923620300579> (Accessed: 9 March 2022).
- Crenshaw, K. (1989). Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. University of Chicago Legal Forum. [PDF] Available at: <https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1052&context=uclf> (Accessed: 9 March 2022).
- Défenseur des droits and Commission Nationale de l'Informatique et des Libertés, CNIL. (2020). Algorithms: preventing automated discrimination. [PDF] Available at: [https://www.defenseurdesdroits.fr/sites/default/files/atoms/files/836200280\\_ddd\\_synthalgos\\_access.pdf](https://www.defenseurdesdroits.fr/sites/default/files/atoms/files/836200280_ddd_synthalgos_access.pdf) (Accessed: 9 March 2022).
- Digital Future Society. (2020a). Exploring gender responsive designs in digital welfare. Barcelona, Spain. [PDF] Available at: <https://digitalfuturesociety.com/es/report/exploring-gender-responsive-designs-in-digital-welfare/> (Accessed: 15 May 2022).
- Digital Future Society. (2020b). Towards gender equality in digital welfare. Barcelona, Spain. [PDF] Available at: <https://digitalfuturesociety.com/es/report/hacia-la-igualdad-de-genero-en-el-estado-de-bienestar-digital/> (Accessed: 15 May 2022).

Digital Future Society. (2022). Governing algorithms: perils and powers of AI in the public sector. Barcelona, Spain. [PDF] Available at: <https://digitalfuturesociety.com/es/report/governing-algorithms/> (Accessed: 9 March 2022).

Eguiluz Castañeira, J. A. (2020). Desafíos y retos que plantean las decisiones automatizadas y los perfilados para los derechos fundamentales. Estudios de Deusto, Vol. 68 (2). [PDF] Available at: <https://revista-estudios.revistas.deusto.es/article/view/1954> (Accessed: 9 March 2022).

Eubanks, V. (2018). Automating Inequality. How high-tech tools profile, police, and punish the poor. New York, USA, Ed. St Martin's Press.

European Digital Rights, EDRI. (2021). Civil society calls for AI red lines in the European Union's Artificial Intelligence proposal. [PDF] Available at: <https://edri.org/our-work/civil-society-call-for-ai-red-lines-in-the-european-unions-artificial-intelligence-proposal/> (Accessed: 25 March 2022).

European Union Agency for Fundamental Rights and the Council of Europe. (2018). #BigData: Discrimination in data-supported decision making. [PDF] Available at: <https://fra.europa.eu/en/publication/2018/bigdata-discrimination-data-supported-decision-making> (Accessed: 9 March 2022).

European Union Agency for Fundamental Rights and the Council of Europe. (2020). Getting the future right: Artificial Intelligence and Fundamental Rights. [PDF] Available at: [https://fra.europa.eu/sites/default/files/fra\\_uploads/fra-2020-artificial-intelligence\\_en.pdf](https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf) (Accessed: 9 March 2022).

Gerards, J. and Xenidis R. (2020). Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law. European Commission. [PDF] Available at: [www.equalitylaw.eu/downloads/5361-algorithmic-discrimination-in-europe-pdf-1-975](http://www.equalitylaw.eu/downloads/5361-algorithmic-discrimination-in-europe-pdf-1-975) (Accessed: 9 March 2022).

Gillis, T. and Spiess, J. (2019). Big Data Discrimination. The University of Chicago Law Review, 86: 459-487. [PDF] Available at: <https://chicagounbound.uchicago.edu/uclrev/vol86/iss2/4> (Accessed: 25 March 2022).

Hacker, P. (2018). Teaching fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination Under EU Law. Common Market Law Review 55:1143-1185. [PDF] Available at: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3164973](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3164973) (Accessed: 9 March 2022).

Hao, K. (2019). Making face recognition less biased doesn't make it less scary. MIT Technology Review. [online] Available at: <https://www.technologyreview.com/2019/01/29/137676/making-face-recognition-less-biased-doesnt-make-it-less-scary/> (Accessed: 25 March 2022).

Independent High Level Expert Group on Artificial Intelligence. (2020). Assessment List for Trustworthy AI (ALTAI). European Commission. [PDF] Available at: <https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment> (Accessed: 9 March 2022).

Jiménez, M. (2021). Un buscador para conocer qué algoritmos impactan en tu vida. El País. [online] Available at: [https://cincodias.elpais.com/cincodias/2021/10/07/companias/1633642034\\_800454.html](https://cincodias.elpais.com/cincodias/2021/10/07/companias/1633642034_800454.html) (Accessed: 9 March 2022).

- Kleinberg, J. and Sunstein, C. (2019). Discrimination in the Age of Algorithms. *Journal of Legal Analysis*. [PDF] Available at: <https://academic.oup.com/jla/article/doi/10.1093/jla/laz001/5476086> (Accessed: 9 March 2022).
- Lagioia, F. and Sartor, G. (2020). Artificial intelligence in the big data era: risks and opportunities. *ResearchGate*. [PDF] Available at: [https://www.researchgate.net/publication/345373824\\_Artificial\\_intelligence\\_in\\_the\\_big\\_data\\_era\\_risks\\_and\\_opportunities](https://www.researchgate.net/publication/345373824_Artificial_intelligence_in_the_big_data_era_risks_and_opportunities) (Accessed: 9 March 2022).
- Mahler, T. (2021). Between risk management and proportionality: The risk-based approach in the EU's Artificial Intelligence Act Proposal. *Nordic Yearbook of Law and Informatics*. [PDF] Available at: <https://ssrn.com/abstract=4001444> (Accessed: 9 March 2022).
- Monclús, J. and Martínez, J. M. (2021). Algoritmos y representación legal de la plantilla: Del secreto empresarial... ¿A la cogobernanza? *Cuatrecasas*. [online] Available at: <https://www.cuatrecasas.com/es/spain/articulo/algoritmos-representacion-legal-plantilla-secreto-empresarial-cogobernanza> (Accessed: 9 March 2022).
- National Institute of Standards and Technology, (NIST). (2021). Proposal for Identifying and Managing Bias in Artificial Intelligence. NIST Special Publication 1270. [PDF] Available at: <https://www.nist.gov/artificial-intelligence/proposal-identifying-and-managing-bias-artificial-intelligence-sp-1270> (Accessed: 9 March 2022).
- Oliva León, R. (2021). ¿Es discriminatorio el algoritmo de Deliveroo que "rankea" a sus riders? *Algoritmo Legal*. [online] Available at: <https://www.algoritmolegal.com/tecnologias-disruptivas/es-discriminatorio-el-algoritmo-de-deliveroo-que-ranke-a-sus-riders> (Accessed: 9 March 2022).
- Orwat, C. (2020). Risks of Discrimination through the Use of Algorithms. *ResearchGate*. [PDF] Available at: [https://www.researchgate.net/publication/344815625\\_Risks\\_of\\_Discrimination\\_through\\_the\\_Use\\_of\\_Algorithms](https://www.researchgate.net/publication/344815625_Risks_of_Discrimination_through_the_Use_of_Algorithms) (Accessed: 9 March 2022).
- Pandey, A. and Caliskan, A. (2021). Disparate Impact of Artificial Intelligence Bias in Ridehailing Economy's Price Discrimination Algorithms. *arXiv*. [PDF] Available at: <https://arxiv.org/abs/2006.04599> (Accessed: 9 March 2022).
- Schiek, D. (2016). Revisiting intersectionality for EU anti-discrimination law in an economic crisis - A critical legal studies perspective. *ResearchGate*. [PDF] Available at: [https://www.researchgate.net/publication/312241500\\_Revisiting\\_intersectionality\\_for\\_EU\\_anti-discrimination\\_law\\_in\\_an\\_economic\\_crisis\\_-\\_A\\_critical\\_legal\\_studies\\_perspective](https://www.researchgate.net/publication/312241500_Revisiting_intersectionality_for_EU_anti-discrimination_law_in_an_economic_crisis_-_A_critical_legal_studies_perspective) (Accessed: 9 March 2022).
- Selbst, A. D. and Powles, J. (2017). Meaningful information and the right to explanation. *International Data Privacy Law*, Volume 7, Issue 4, November 2017, Pages 233-242. [PDF] Available at: <https://academic.oup.com/idpl/article/7/4/233/4762325> (Accessed: 9 March 2022).
- Wachter, S., Mittelstadt, B. and Floridi, L. (2017). Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law* (7,2). [PDF] Available at: <https://academic.oup.com/idpl/article/7/2/76/3860948> (Accessed: 9 March 2022).

Xenidis, R. and Senden, L. (2020). EU non-discrimination law in the era of artificial intelligence: Mapping the challenges of algorithmic discrimination. General Principles of EU Law and the EU Digital Order. [PDF] Available at: [https://www.researchgate.net/publication/348382658\\_EU\\_Non-discrimination\\_Law\\_in\\_the\\_Era\\_of\\_Artificial\\_Intelligence\\_Mapping\\_the\\_Challenges\\_of\\_Algorithmic\\_Discrimination](https://www.researchgate.net/publication/348382658_EU_Non-discrimination_Law_in_the_Era_of_Artificial_Intelligence_Mapping_the_Challenges_of_Algorithmic_Discrimination) (Accessed: 9 March 2022).

Zuboff, S. (2020). La era del Capitalismo de la Vigilancia. Madrid, Spain: Ed. Paidós.

Zuiderveen Borgesius, F. J. (2018). Discrimination, artificial intelligence, and algorithmic decision-making. Council of Europe. [PDF] Available at: <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73> (Accessed: 9 March 2022).

Zuiderveen Borgesius, F. J. (2020). Strengthening legal protection against discrimination by algorithms and artificial intelligence. The International Journal of Human Rights. [PDF] Available at: <https://www.tandfonline.com/doi/pdf/10.1080/13642987.2020.1743976?needAccess=true> (Accessed: 9 March 2022).

# Acknowledgements

---

## Authors

**Dolores Morondo Taramundi**, head of research at the Pedro Arrupe Human Rights Institute at Deusto University and member of the European Network of Legal Experts on Equality and Non-Discrimination. Her main research areas are anti-discrimination law, critical theories of law, human rights and legal methodology. Graduate in law, specialising in economics, with a PhD in law from the European University Institute.

**Josu Andoni Eguiluz Castañeira**, researcher focusing on the legal challenges that new technologies (artificial intelligence, automated decision-making, the metaverse, etc.) pose for fundamental rights. Graduate at the Intellectual Property, New Technologies and Data Protection Department at Cuatrecasas. Graduate in law and ICT from Deusto University and Dual Master's Degree in Access to the Legal Profession and Intellectual Property, New Technologies and Data Protection from Ecade Law School.

## Digital Future Society Think Tank Team

- **Carina Lopes**, Head of Digital Future Society Think Tank
- **Tanya Álvarez**, Researcher, Digital Future Society Think Tank
- **Olivia Blanchard**, Researcher, Digital Future Society Think Tank

## Citations

Please cite this report as:

- Digital Future Society. (2022). Algorithmic discrimination in Spain: limits and potential of the legal framework. Barcelona, Spain.

## Contact details

To contact the Digital Future Society Think Tank team, please email:  
[thinktank@digitalfuturesociety.com](mailto:thinktank@digitalfuturesociety.com)



**Digital  
Future Society**