Mezei, Kitti[1] – Träger, Anikó[2]

The Digital Equality Turn in the EU: How the AI Act and Platform Regulations Operationalise Non-Discrimination[3]

I. From Opacity to Inequality: Case Studies of Algorithmic Discrimination[4]

Artificial intelligence (AI)[5] poses distinctive challenges for the principle of non-discrimination. While discrimination in traditional legal contexts often depends on explicit or easily identifiable differential treatment, AI systems produce inequality through far more complex and opaque processes. This opacity complicates both detection and legal redress, leading to systemic under-enforcement of equality norms.[6] These abstract concerns become clearer when viewed through concrete examples, which demonstrate how seemingly neutral algorithmic tools can reproduce and even amplify structural bias.

Case studies illustrate the range of discrimination risks. The widely cited COMPAS risk assessment tool in U.S. criminal justice notoriously produced racially biased outcomes. Black defendants were nearly twice as likely as white defendants to be misclassified as high-risk non-recidivists, while white defendants were more likely to be misclassified as low-risk recidivists.

---

[1] Senior Research Fellow, ELTE Centre for Social Sciences, Institue for Legal Studies.
[2] Junior Research Fellow, ELTE Centre for Social Sciences, Institue for Legal Studies.
[3] This paper was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. The research was conducted within the framework of the European Union project RRF-2.3.1-21-2022-00004 (Artificial Intelligence National Laboratory) as well as the Ministry of Innovation and Technology NRDI Office's FK-21 Young Researcher Excellence Program (grant no. 138965).
[4] Algorithmic bias refers to systematic distortions in data or model design that create unequal outputs (Aytekin, F., Bozkaya, B. and Jentzsch, N., 'Algorithmic Bias in the Context of European Union Anti-Discrimination Law' (2023) CEUR Workshop Proceedings 3442, 447–61; Lendvai, G. F. and Gosztonyi, G., 'Algorithmic Bias as a Core Legal Dilemma in the Age of Artificial Intelligence: Conceptual Basis and the Current State of Regulation' (2025) 14(3) *Laws* 41 https://doi.org/10.3390/laws14030041 accessed 21 September 2025.). Algorithmic fairness captures the normative and technical measures—such as fairness metrics or fairness-by-design obligations—meant to mitigate these risks (Hacker, P., Wiedemann, E. and Zehlike, M., 'Towards a Flexible Framework for Algorithmic Fairness' (2020) arXiv https://arxiv.org/abs/2010.07848 accessed 21 September 2025.). Algorithmic discrimination, by contrast, denotes outcomes that legally constitute unequal treatment under EU non-discrimination law, whether through direct, indirect, or proxy discrimination (Wang, Y., Han, C. and Wang, P., 'Algorithmic Discrimination: A Grounded Conceptualization' (2025) 28(2) *Information, Communication & Society* 251–72.).
[5] This paper uses "AI" and "algorithms" in a deliberate way. "AI" refers to the legal category established in EU AI legislation, encompassing machine learning and related techniques. By contrast, "algorithms" is used more broadly to capture decision-making mechanisms at stake—including simpler rule-based or ranking systems—since discrimination risks can arise in both. Accordingly, "AI" is used in regulatory discussions, while "algorithm" is employed when highlighting the technical or socio-legal operation of these systems. Strictly speaking, it is the machine-learned model, rather than the learning algorithm itself, that produces discriminatory outputs.
[6] For clarity in this paper, "non-discrimination" refers to the EU's doctrinal framework of equality directives and case law prohibiting unequal treatment on protected grounds. By contrast, "equality" denotes a broader governance approach, exemplified by digital legislation such as the AI Act, the Digital Services Act, and the Digital Markets Act, which embed systemic fairness obligations even beyond the formal scope of non-discrimination law.

Despite being presented as objective and data-driven, COMPAS effectively reproduced racial disparities and undermined defendants' rights to equal treatment and due process.[7]

More recently, Google's Gemini model illustrates a different pathology: overcorrection bias. In seeking to avoid exclusionary outputs, Gemini generated historically inaccurate and offensive images (e.g., depicting Nazi soldiers as African American). This demonstrates that algorithmic fairness interventions can themselves distort representation, raising fresh challenges for equality law.[8]

In traditional labour markets as well, AI-based systems are increasingly used to streamline and automate processes such as CV screening, performance evaluation, and employee assessment.[9] A well-known example is Amazon's recruitment algorithm, which—trained on data from a historically male-dominated workforce—systematically penalised women applicants by downgrading résumés with indicators of female identity. In this case, historic bias in training data directly translated into gender discrimination, undermining the principle of equality between men and women.[10]

Moreover, food-delivery platforms such as Deliveroo and Uber Eats have become central illustrations of how algorithmic management can generate discriminatory effects in employment contexts. Unlike traditional employment relationships, where managers make direct decisions, these platforms rely on automated rating and assignment systems to allocate work, assess performance, and even terminate contracts.[11]

Between 2021 and 2022, courts in Italy, Spain, and the UK began testing platform algorithms against equality law. In Italy, the Court of Bologna ruled that Deliveroo's rating system unlawfully penalised riders who were absent for protected reasons such as illness, childcare, or

---

[7] Angwin, J., Larson, J., Mattu, S. and Kirchner, L., 'Machine Bias' (2016) *ProPublica* https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing accessed 21 August 2025.
Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G. and Yu, H., 'Accountable Algorithms' (2017) 165(3) *University of Pennsylvania Law Review* 633–705. and Engel, C., Linhardt, L. and Schubert, M., 'Code is Law: How COMPAS Affects the Way the Judiciary Handles the Risk of Recidivism' (2024) 33 *Artificial Intelligence and Law* 383–404.

[8] Da Silveira, J. B. and Alves Lima, E., 'Racial Biases in AIs and Gemini's Inability to Write Narratives About Black People' (2024) 2 Emerging Media 277–287.

[9] Aloisi, A., 'Regulating Algorithmic Management at Work in the European Union: Data Protection, Non-Discrimination and Collective Rights' (2024) 40(1) International Journal of Comparative Labour Law and Industrial Relations 37–70.

[10] Hsu, J., 'Can AI Hiring Systems Be Made Antiracist? Makers and Users of AI-Assisted Recruiting Software Reexamine the Tools' Development and How They're Used' (2020) 57 *IEEE Spectrum* 9–11.

[11] See more Tuomi, A., Jianu, B., Roelofsen, M. and Passos Ascenção, M. P., 'Riding against the Algorithm: Algorithmic Management in On-Demand Food Delivery' in Ferrer-Rosell, B., Massimo, D. and Berezina, K. (eds), *Information and Communication Technologies in Tourism 2023 – Proceedings of the ENTER 2023 eTourism Conference (Springer 2023)* 28–39 https://doi.org/10.1007/978-3-031-25752-0_3 accessed 21 August 2025.

strikes, thereby disadvantaging workers with family duties or union membership.[12] In Spain, the "Ley Riders" law was introduced in response to similar disputes, presuming employment for platform workers and requiring transparency about how algorithms allocate work.[13] In the UK, an Uber Eats courier challenged the company's facial recognition "selfie" checks after repeated lockouts, arguing that the system's higher error rates for racialised individuals amounted to racial discrimination.[14] Backed by the Equality and Human Rights Commission, the case was allowed to proceed before eventually settling with compensation in 2024.[15] Taken together, these disputes highlight a common pattern: algorithmic management systems can appear neutral but often reproduce discrimination, while workers struggle to access effective remedies.

In 2019, Apple and Goldman Sachs faced public and regulatory attention over allegations of gender discrimination in the credit allocation of the newly launched Apple Card. Multiple high-profile users reported that they and their spouses received drastically different credit limits despite sharing finances and having similar credit histories. Women were frequently assigned much lower limits than their male counterparts, even in cases where the women had higher credit scores. The controversy quickly escalated on social media and prompted an official investigation by the New York Department of Financial Services. Although Goldman Sachs denied any intentional discrimination, it acknowledged the existence of disparities in outcomes and attributed them to the opaque decision-making processes of its underlying risk models. The incident highlighted the black-box nature of credit scoring algorithms,[16] where neither applicants nor regulators had full visibility into the factors driving outcomes. Even absent direct

---

[12] Tribunale di Bologna (Labour Section), Judgment No 2949 of 31 December 2020 (*Deliveroo*).

[13] Royal Decree-Law 9/2021 of 11 May on urgent measures to guarantee the labour rights of persons engaged in delivery services through digital platforms (*Ley Riders*); Spanish Supreme Court (Tribunal Supremo, Plenary), Judgment No 805/2020 of 25 September 2020 (*Glovo*).

[14] This claim is consistent with empirical findings such as those of Buolamwini and Gebru, who demonstrated that commercial facial recognition systems systematically underperform in recognising black women's faces, with error rates between 20.8% and 34.7% for this group. These disparities highlight how intersectional algorithmic discrimination emerges when underrepresented minorities in training datasets are disproportionately misclassified, leading to both individual harms, such as exclusion from work opportunities, and structural inequalities that reinforce existing patterns of disadvantage. See Buolamwini, J. and Gebru, T., 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification' (2018) *Proceedings of Machine Learning Research* http://proceedings.mlr.press/v81/buolamwini18a.html accessed 21 August 2025.

[15] Manjang v Uber Eats UK Ltd Employment Tribunal (Oxford), settled March 2024, reported by the Equality and Human Rights Commission, 'Uber Eats courier wins payout with help of equality watchdog after facing problematic AI' (EHRC, 20 March 2024) https://www.equalityhumanrights.com/en/our-work/news/uber-eats-courier-wins-payout-help-equality-watchdog-after-facing-problematic-ai accessed 21 September 2025.

[16] The Court of Justice of the European Union (CJEU) in Case C-634/21 *SCHUFA Holding AG* (7 December 2023) ECLI:EU:C:2023:957; Case C-203/22 *Dun & Bradstreet* (27 February 2025) ECLI:EU:C:2025:117. confirmed that credit scoring falls within EU law's protections against opaque automated decision-making. Both cases show how individuals are often unable to understand or contest algorithmic outcomes and the CJEU responded by strengthening transparency duties under the GDPR, even against claims of trade secret protection.

use of gender as a variable, the system likely relied on correlated proxies (e.g., income history, types of expenditure, or joint accounts) that disproportionately affected women.[17]

From an anti-discrimination perspective, the Apple Card case illustrates the central difficulty of indirect discrimination in algorithmic systems. A neutral-seeming model produced gender-differentiated outcomes that are difficult for claimants to challenge, both because the inputs are obscured by proprietary models and because institutions can defend the practice as a proportionate means of ensuring predictive accuracy. The burden of proof lies heavily with the affected individuals, who lack access to the training data or internal logic of the scoring system,[18] so it creates the evidentiary barrier that Hacker (2018)[19] and others[20] mentioned. The algorithmic opacity may create an "enforcement choke point" in EU anti-discrimination law: claimants cannot realistically access the training data or model logic needed to prove disparate impact, meaning that legal protections remain largely theoretical.

II. From Unequal Data to Doctrinal Gaps: Why AI Challenges EU Anti-Discrimination Law

The roots of discrimination in AI are multifaceted. Discrimination doesn't emerge only from malicious intent, but from a constellation of factors: biased or incomplete training data; design choices in algorithm architecture and objectives; lack of diversity among developers; and deployment in social settings that amplify historic inequalities. Scholars such as Ferrer et al. (2020)[21] and Zajko (2022)[22] emphasize that technical and social dimensions are deeply

---

[17] Osoba, O. A., 'Did No One Audit the Apple Card Algorithm?' (RAND Corporation, November 2019) https://www.rand.org/pubs/commentary/CA907-1.html accessed 21 August 2025.
New York State Department of Financial Services, Report on Apple Card Investigation (March 2021) https://www.dfs.ny.gov/reports_and_publications/202103_report_apple_card_investigation accessed 21 September 2025.

[18] Since existing legal frameworks presuppose that claimants can detect and substantiate discriminatory treatment without granting procedural rights to access underlying evidence, the burden of proof falls disproportionately on affected individuals, who are further disadvantaged by the opacity of AI systems' training data and decision-making processes. See, Grozdanovski, L., 'Non-Discrimination Law, the GDPR, the AI Act and the – Now Withdrawn – AI Liability Directive Proposal Offering Gateways to Pre-Trial Knowledge of Algorithmic Discrimination' (2025) AI and Ethics https://doi.org/10.1007/s43681-025-00754-0 accessed 21 September 2025.

[19] Hacker, P., 'Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies Against Algorithmic Discrimination under EU Law' (2018) 55 *Common Market Law Review* 1143–85.

[20] Hacker, P., Wiedemann, E. and Zehlike, M., 'Towards a Flexible Framework for Algorithmic Fairness' (2020) arXiv https://arxiv.org/abs/2010.07848 accessed 21 September 2025. Wachter, S., Mittelstadt, B. and Russell, C., 'Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI' (2021) 41 *Computer Law & Security Review* 105567 and Wójcik, M. A., 'Algorithmic Discrimination in Health Care: An EU Law Perspective' (2022) 24(1) *Health and Human Rights* 93–103.

[21] Ferrer, X., van Nuenen, T., Such, J. M., Coté, M. and Criado, N., 'Bias and Discrimination in AI: A Cross-Disciplinary Perspective' (2020) arXiv https://arxiv.org/abs/2008.07309 accessed 21 September 2025.

[22] Zajko, M., 'Artificial Intelligence, Algorithms, and Social Inequality' (2022) 16(7) Sociology Compass https://doi.org/10.1111/soc4.12962 accessed 21 August 2025.

intertwined; Kuhlman et al. (2020)[23] show that who builds the model matters just as much as how the model is built. Most machine learning systems learn from past data, which often reflects entrenched social inequalities. Biased training data—whether through sampling errors, historical exclusion, or prejudiced labelling—builds inequality directly into algorithmic logic. Research shows that biased labels can cause models to replicate human prejudice,[24] and that even fairness adjustments cannot fix residual unfairness if the data itself reflects historical exclusion.[25] Labelling and measurement errors further distort fairness metrics.[26] Recent work, such as FAIRLABEL, demonstrates both the pervasiveness of biased labels and methods to correct them.[27] Moreover, Zajko (2020) argues that algorithms are trained on an unequal ground truth, meaning that structural disadvantages across gender, race, class, and disability are already embedded in the data before modelling begins.[28] Lopez (2021) defines different types of bias including "societal bias", where structural inequalities in society are correctly reflected in data—i.e. an unequal ground truth that algorithms reproduce. Machine-learning systems often reproduce what Zajko calls an "unequal ground truth"—the structurally biased social realities encoded in training data—thereby amplifying disadvantage, further embeds disparities.[29] Proxy discrimination is especially problematic: even when sensitive attributes like race or gender are removed, neutral features such as postcode, hobbies, or education can correlate strongly with protected traits, producing indirect discrimination that is both pervasive and difficult to trace. It represents a central challenge in algorithmic fairness: models may appear blind to protected attributes but still discriminate via correlated neutral features. Researchers formalising proxy use,[30] legal analysis of rational proxy discrimination,[31] and empirical work about "fairness

---

[23] Kuhlman, C., Jackson, L. and Chunara, R., 'No Computation without Representation: Avoiding Data and Algorithm Biases through Diversity' (2020) *arXiv* https://arxiv.org/abs/2002.11836 accessed 21 September 2025.
[24] Jiang, H. and Nachum, O., 'Identifying and Correcting Label Bias in Machine Learning' (2019) *arXiv* https://arxiv.org/abs/1901.04966 accessed 21 September 2025.
[25] Kallus, N. and Zhou, A., 'Residual Unfairness in Fair Machine Learning from Prejudiced Data' (2018) *arXiv* https://arxiv.org/abs/1806.02887 accessed 21 September 2025.
[26] Liao, Y. and Naghizadeh, P., 'Social Bias Meets Data Bias: The Impacts of Labeling and Measurement Errors on Fairness Criteria' (2022) *arXiv* https://arxiv.org/abs/2206.00137 accessed 21 September 2025.
[27] Sengamedu, S. H., Pham, T., Ram, P., Shrivastava, A., Tran, T. and Weerasinghe, S., 'FAIRLABEL: Correcting Bias in Labels' (2023) *arXiv* https://arxiv.org/abs/2311.00638 accessed 21 September 2025.
[28] Zajko, M., 'Conservative AI and Social Inequality: Conceptualizing Alternatives to Bias through Social Theory' (2020) *arXiv* https://arxiv.org/abs/2007.08666 accessed 21 September 2025.
[29] Lopez, P., 'Bias Does Not Equal Bias: A Socio-Technical Typology of Bias in Data-Based Algorithmic Systems' (2021) 10(4) *Internet Policy Review*.
[30] Datta, A., Fredrikson, M., Ko, G., Mardziel, P. and Sen, S., 'Proxy Non-Discrimination in Data-Driven Systems: Theory and Experiments with Machine Learnt Programs' (2017) *arXiv* https://arxiv.org/abs/1707.08120 accessed 21 September 2025.
[31] Prince, A. E. R. and Schwarcz, D., 'Proxy Discrimination in the Age of Artificial Intelligence and Big Data' (2020) 105(4) *Iowa Law Review* 1257–1317.

under unawareness"[32] all show that tracing such indirect harm is both technically and legally complex. Furthermore, feedback loops intensify inequality: a biased hiring model that predominantly selects white men will generate new datasets that reinforce the same outcome, locking in structural disadvantage.[33] Recent research shows that biased hiring systems do not just mirror existing inequalities—they can actually magnify them over time through feedback loops. According to Baek and Makhdoumi (2023),[34] for instance, model scenarios where two groups start out with the same underlying skills. Even so, if one group is slightly less likely to be evaluated positively or hired at the beginning, that small difference compounds. Over successive rounds of hiring, the disadvantaged group gets fewer opportunities to demonstrate ability, which in turn leads employers to underestimate their skills even further. The result is a self-reinforcing cycle in which the disadvantaged group is eventually locked out of most opportunities. Hu and Chen (2017) similarly demonstrate that when hiring practices ignore fairness across groups, the labour market can settle into stable but unequal patterns.[35] Once these inequitable "equilibria" form, they tend to lock in disadvantages for certain groups: even if employers are no longer consciously discriminating, the system reproduces past imbalances in ways that are self-sustaining and resistant to change unless targeted corrective measures are introduced.

From a legal perspective, these dynamics map imperfectly onto EU anti-discrimination law.[36] Though direct discrimination—where protected traits explicitly determine outcomes—is relatively rare in AI, indirect discrimination is much more common, as neutral algorithms often

---

[32] Chen, J., Kallus, N., Mao, X., Svacha, G. and Udell, M., 'Fairness Under Unawareness: Assessing Disparity When Protected Class Is Unobserved' (2018) *arXiv* https://arxiv.org/abs/1811.11154 accessed 21 September 2025.

[33] European Union Agency for Fundamental Rights (FRA), *Bias in Algorithms: Artificial Intelligence and Discrimination* (Vienna, 2022) 8–10.

[34] Baek, J. and Makhdoumi, A., 'The Feedback Loop of Statistical Discrimination' (2023) *SSRN* http://dx.doi.org/10.2139/ssrn.4658797 accessed 21 August 2025.

[35] Hu, L. and Chen, Y., 'A Short-Term Intervention for Long-Term Fairness in the Labor Market' (2017) *arXiv* https://arxiv.org/abs/1712.00064 accessed 21 September 2025.

[36] From the perspective of EU law, protection against discrimination is anchored in both the treaties and secondary legislation. Article 21 of the Charter of Fundamental Rights of the European Union sets out a broad prohibition on discrimination covering grounds such as sex, race, ethnic origin, religion or belief, disability, age, and sexual orientation, while Article 19 Treaty on the Functioning of the EU (TFEU) provides the legislative competence to combat such discrimination through EU directives. Building on this foundation, the EU has adopted a framework of anti-discrimination directives. The Racial Equality Directive (2000/43/EC) prohibits discrimination on the basis of racial or ethnic origin across employment, education, social protection, and access to goods and services, while the Employment Equality Framework Directive (2000/78/EC) addresses discrimination in the workplace on the grounds of religion or belief, disability, age, and sexual orientation. Alongside these, the Gender Equality Directives – notably the Recast Directive (2006/54/EC), the Goods and Services Directive (2004/113/EC), and the Self-Employment Directive (2010/41/EU) – establish detailed rules to ensure equal treatment between men and women.

produce disproportionate disadvantages for protected groups.[37] Enforcement is difficult in practice. Victims must often detect and prove disparate impact, but this is nearly impossible when models are opaque and claimants lack access to training data or internal audit trails.[38] Even when patterns of disadvantage are visible, developers may invoke efficiency or predictive accuracy as a proportionate justification for harmful outcomes—which courts under EU law may accept.[39] The result is a gap between the theoretical protective scope of non-discrimination law and its practical enforceability in algorithmic contexts.[40]

Moreover, the EU equality law already struggles to capture intersectional discrimination, with the CJEU showing limited awareness of how overlapping disadvantages operate in practice.[41] This gap becomes even more acute in the AI context, where opacity, structural reproduction of inequality, and the mismatch between legal categories and algorithmic harms make it harder still to address complex, layered forms of discrimination.[42]

Together, these limitations underscore why traditional, reactive, litigation-based enforcement is insufficient and why preventive obligations—such as duties of care in system design, robust dataset governance, enforceable transparency rights, and systemic risk oversight—are essential, as reflected in the EU's AI Act and other digital legislative initiatives.[43]

III. From Fragmented EU Non-Discrimination Law to Preventive Digital Regulation

The EU's non-discrimination framework constitutes a fragmented assemblage of sectoral and ground-specific instruments rather than a single, comprehensive regime. Outside the field of employment, protection against unequal treatment is limited to race and ethnic origin (Directive 2000/43/EC) and sex (Directive 2004/113/EC). By contrast, other grounds recognised in EU

---

[37] Wachter, S., Mittelstadt, B. and Russell, C., (2021).; Weerts, H., Xenidis, R., Tarissan, F., Olsen, H. P. and Pechenizkiy, M., 'Algorithmic Unfairness Through the Lens of EU Non-Discrimination Law: Or Why the Law Is Not a Decision Tree' (2023) CEUR Workshop Proceedings 3442, 805–16 https://doi.org/10.1145/3593013.3594044 accessed 21 September 2025.

[38] Hacker, P. (2018)

[39] Zuiderveen Borgesius, F. J., 'Strengthening Legal Protection Against Discrimination by Algorithms and Artificial Intelligence' (2020) 24(10) *The International Journal of Human Rights* 1572–93 https://doi.org/10.1080/13642987.2020.1743976 accessed 21 September 2025.

[40] Zuiderveen Borgesius, F., Baranowska, N., Hacker, P. and Fabris, A., 'Non-Discrimination Law in Europe: A Primer for Non-Lawyers' (SSRN Working Paper, April 2024) https://ssrn.com/abstract=4786956 accessed 21 September 2025.

[41] Xenidis, R., 'Tuning EU Equality Law to Algorithmic Discrimination' (2020) 27(6) *Maastricht Journal of European and Comparative Law* 742.

[42] Xenidis, R., 'Towards a Legal Response to Algorithmic Discrimination in the Algorithmic Society' (2023) 29(1–2) *European Law Journal* 79–96.

[43] Veale, M. and Zuiderveen Borgesius, F. J., 'Demystifying the Draft EU Artificial Intelligence Act' (2021) 22(4) *Computer Law Review International* 97–112 https://doi.org/10.9785/cri-2021-220402 accessed 21 September 2025.

primary law—religion or belief, disability, age, and sexual orientation—are covered only in employment contexts through Directive 2000/78/EC, but not in the wider provision of goods and services. This asymmetry leaves significant domains of AI-mediated services—including credit scoring, dynamic pricing, and personalised recommendation systems—either partially or wholly beyond the scope of the EU equality acquis when they occur outside the workplace.[44] The European Commission sought to close this gap through the 2008 Horizontal Equal Treatment Directive proposal, yet persistent political deadlock prevented its adoption, and protection across grounds and sectors remains uneven.

Two structural features further complicate the application of EU equality law in digital environments. First, the definition of "services" under Directive 2004/113/EC is linked to remuneration, creating uncertainty as to whether "free" online services financed through advertising or the monetisation of personal data fall within its ambit—even though the Digital Content Directive (2019/770) now recognises personal data as a form of counter-performance in consumer law. Second, the architecture of algorithmic systems—particularly in the form of ranking, targeting, and recommender mechanisms—shapes access to opportunities indirectly, by governing visibility and exposure, rather than through discrete contractual decisions. Such dynamics sit uneasily with traditional conceptions of "access to services," thereby creating regulatory blind spots. These are precisely the spaces in which algorithmic discrimination is most likely to emerge, unaddressed by clear ex ante equality obligations.[45]

In response to such limitations, the EU has pursued complementary regulatory strategies. The AI Act introduces risk-based obligations, including data governance, transparency, and human oversight, for high-risk AI applications such as credit scoring and employment systems. The Digital Services Act (DSA)[46] requires very large platforms and search engines (VLOPs and VLOSEs) to assess and mitigate systemic risks, explicitly including risks of discriminatory impact in recommender systems.[47] The Digital Markets Act (DMA)[48] further addresses structural imbalances by imposing obligations on gatekeeper platforms, such as rules on self-

---

[44] Zuiderveen Borgesius, F. J., (2020).

[45] Directive (EU) 2019/770 of the European Parliament and of the Council of 20 May 2019 on certain aspects concerning contracts for the supply of digital content and digital services [2019] OJ L136/1.

[46] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act) [2022] OJ L277/1.

[47] See about risk-related EU digital regulation: Mezei, K. and Träger, A., 'Risks and Resilience in the European Union's Regulation of Online Platforms and Artificial Intelligence: Hungary in Digital Europe' in Gárdos-Orosz, F. (ed), *The Resilience of the Hungarian Legal System since 2010* (European Union and its Neighbours in a Globalized World, vol 16, Springer 2025) 150–51.

[48] Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act) [2022] OJ L265/1.

preferencing and data use, which indirectly bear on fairness in digital markets. Similarly, the Platform Work Directive (PWD) establishes transparency and accountability requirements for algorithmic management in the gig economy. While these instruments do not expand the substantive scope of EU non-discrimination directives, they signal a preventive, governance-oriented turn that complements the equality acquis in digital contexts.

Crucially, their regulatory form—so-called „acts"[49] with direct effect, alongside directives requiring national implementation—marks a significant departure from the EU's traditional reliance on sectoral equality directives. Whereas anti-discrimination law has long depended on fragmented, transposed instruments whose effectiveness varies across Member States, the new generation of digital regulations establishes uniform, immediately binding obligations on platforms and AI providers throughout the EU. Moreover, these regulations are designed with a marked extraterritorial effect: the AI Act, the DSA, and the DMA apply not only to entities established within the EU but also to third-country providers whose services reach EU users. This regulatory architecture therefore extends the EU's equality-related governance standards well beyond its borders. The result is not merely a complementary layer of protection in digital contexts, but a broader reorientation of EU equality governance—shifting from reactive, litigation-driven enforcement towards preventive, harmonised, and globally resonant regulation.

VI. The AI Act's Response to Algorithmic Discrimination

The AI Act addresses discrimination-related risks through several interlinked provisions targeting different stages of AI system design and use. The general framework is reinforced by Annex III, which designates specific applications as "high-risk" precisely because of their potential to produce discriminatory effects in socially sensitive domains. Credit scoring and creditworthiness assessment are explicitly listed as high-risk applications in relation to access to essential private services [Annex III, point 5(b)]. Recruitment, worker management, and automated termination tools, such as Amazon's hiring algorithm or Deliveroo's and Uber Eats' rating and assignment systems, are classified as high-risk in the field of employment, workers' management, and access to self-employment (Annex III, point 4). Biometric identification systems, such as Uber Eats' facial recognition checks, are designated high-risk in the area of

[49] Scholars have described this broader trend as the "actification" of EU digital law, where regulation increasingly takes the form of directly applicable Acts rather than directives. See, Papakonstantinou, V. and De Hert, P., *The Regulation of Digital Technologies in the EU: Act-ification, GDPR Mimesis, and EU Law Brutality at Play* (Routledge 2024).

biometric identification and categorisation of natural persons (Annex III, point 1). These classifications reflect the recognition that such systems, while often presented as neutral or efficient, directly affect fundamental rights by determining access to employment, income, financial resources, and essential services—precisely the areas where algorithmic discrimination can have the most severe and systemic consequences.

Article 10 of the AI Act addresses discrimination risks at their root: the datasets that train high-risk AI systems. Providers must ensure that training, validation, and testing data are not only relevant to the system's intended purpose but also representative of the populations on which the system will be deployed. This means that datasets must capture the diversity of affected users, including vulnerable groups such as ethnic minorities, older persons, or persons with disabilities, so that these groups are not systematically misclassified or excluded. The provision also requires datasets to be statistically robust and as complete as possible, with careful validation to minimise skewed distributions or gaps that could produce biased outcomes. Importantly, Article 10(4) adds a contextual dimension: datasets must reflect the geographical, linguistic, and cultural environment in which the system is intended to operate—for example, by including regional dialects or behavioural patterns relevant to the target population. In practice, these requirements translate into a duty of bias-checking and data governance long before a system reaches the market.

Articles 14 and 15 extend these safeguards into the operational life of AI systems. Article 14(4)(b) requires human supervisors to monitor for automation bias, the well-documented tendency to defer uncritically to algorithmic outputs, which could otherwise allow discriminatory decisions to pass unchecked.[50] Article 15(4) imposes design obligations on adaptive or continuously learning systems, requiring providers to prevent feedback loops in which biased outputs are fed back into training data, progressively amplifying inequality—for instance, in recruitment systems that only recommend candidates resembling past hires.

Taken together, Articles 10, 14, and 15 of the AI Act establish a three-layer safeguard against algorithmic discrimination by requiring bias-resistant datasets, critical human oversight to counter automation bias, and system design that prevents the structural reinforcement of disparities. Finally, the AI Act acknowledges a key tension in non-discrimination law: bias cannot always be detected without engaging directly with sensitive data such as race, ethnicity, or health status. Article 10 therefore permits the exceptional processing of special categories of

[50] See more, nqvist, L., '"Human Oversight" in the EU Artificial Intelligence Act: What, When and by Whom?' (2023) 15(2) *Law, Innovation and Technology* 508–35 https://doi.org/10.1080/17579961.2023.2245683 accessed 21 August 2025.

data, but only under strict GDPR-level safeguards, including pseudonymisation, limited retention, state-of-the-art security, and deletion once the correction is complete. This balancing act underscores the EU's approach: fundamental rights protection requires not just prohibiting discriminatory outcomes but embedding preventive, context-sensitive duties of care into the design and governance of high-risk AI systems.[51]

The AI Act complements—not replaces—non-discrimination directives by imposing ex-ante duties that operationalise equality concerns in high-risk uses: data governance for training, validation and testing (Article 10), transparency and instructions (Article 13), human oversight (Article 14), logging and record-keeping (Article 12), risk and quality management (Article 9 and 17). It introduces a Fundamental Rights Impact Assessment (FRIA) obligation for public bodies and certain private providers of public services, and for specific high-risk use cases (e.g., credit and insurance). Article 27 requires deployers to identify impacted groups, map risks, and notify market surveillance authorities. This is a systemic, preventive tool aimed exactly at the evidentiary and governance deficits.[52]

Under EU anti-discrimination law, a practice that has a disparate impact can still be upheld if the defendant shows it pursues a legitimate aim and is proportionate. In many cases, courts have accepted efficiency or predictive accuracy as such legitimate aims.[53] An instructive national example comes from the Netherlands, where the life insurer Dazure differentiated premiums primarily on the basis of a customer's postcode, alongside lifestyle factors such as smoking. Although postcode is formally a neutral criterion, in practice it correlates strongly with socio-economic and ethnic segregation, meaning that members of racial or ethnic minorities concentrated in lower-income neighbourhoods were more likely to face higher premiums. This

---

[51] Deck, L. and others, 'Implications of the AI Act for Non-Discrimination Law and Algorithmic Fairness' (2024) *arXiv* https://arxiv.org/abs/2403.20089 accessed 21 August 2025.

[52] Lütz, F., 'The AI Act, Gender Equality and Non-Discrimination: What Role for the AI Office?' (2024) 25 *ERA Forum* 79–95 https://doi.org/10.1007/s12027-024-00785-w accessed 21 August 2025.

[53] Although none of these cases involved AI, they demonstrate the doctrinal logic the CJEU applies when assessing indirect discrimination. See, eg: Case 170/84 *Bilka-Kaufhaus GmbH v Karin Weber von Hartz* ECLI:EU:C:1986:204 (part-time workers excluded from pensions; justification based on business efficiency and cost control); Case C-237/94 *John O'Flynn v Adjudication Officer* ECLI:EU:C:1996:206 (rules disadvantaging migrant workers could be justified by administrative efficiency); Case C-167/97 *R v Secretary of State for Employment, ex p Seymour-Smith and Perez* ECLI:EU:C:1999:60 (two-year qualifying period for unfair dismissal claims disproportionately affecting women upheld as labour-market policy measure); Case C-83/14 *CHEZ Razpredelenie Bulgaria AD v Komisia za zashtita ot diskriminatsia* ECLI:EU:C:2015:480 (electricity meters placed high in Roma districts could in principle be justified by fraud-prevention and efficiency concerns, subject to strict proportionality); Case C-407/98 *Katarina Abrahamsson and Leif Anderson v Elisabet Fogelqvist* ECLI:EU:C:2000:367 (affirmative action struck down where it disproportionately undermined efficiency/merit).

raised questions under the EU Racial Equality Directive 2000/43/EC, which extends the prohibition of direct and indirect discrimination to goods and services, including insurance. Because postcode is not itself a protected ground, the assessment fell under indirect discrimination: a neutral provision liable to put persons of a particular racial or ethnic origin at a particular disadvantage. The Dutch Human Rights Institute acknowledged this disparate impact but ultimately held the practice objectively justified. The insurer was pursuing a legitimate aim – actuarially sound risk assessment—through appropriate means—postcode as a statistically validated proxy for life expectancy—and in a manner deemed necessary, since the alternative of collecting more intrusive individual financial data (e.g., income) would have raised additional privacy and proportionality concerns.[54] This reasoning foreshadows the kinds of trade-offs likely to arise in algorithmic systems: providers may defend biased outputs on the basis that their models are more accurate or privacy-preserving, even if the result entrenches structural disadvantage. The AI Act does not directly amend this proportionality doctrine. Instead, it takes a preventive approach by imposing ex ante obligations: developers of high-risk AI systems must ensure the quality and representativeness of datasets (Article 10), provide documentation and transparency (Article 13), and guarantee meaningful human oversight (Article 14). These requirements mean that providers cannot simply defend an unequal outcome by pointing to accuracy. They must demonstrate that they took active steps to prevent bias in design and deployment.[55] So, the EU non-discrimination law addresses discrimination retrospectively, while the AI Act requires fairness to be built into AI at the design stage; bridging this gap demands translating legal standards into technical requirements through interdisciplinary collaboration.[56]

Moreover, if providers fail to meet these duties, according to the Article 74(13), the market surveillance authorities are empowered to "open the black box": they can require disclosure of technical documentation, training data specifications, and, where necessary, even grant full access to the source code to assess compliance. This regulatory capacity significantly shifts the

---

[54] College voor de Rechten van de Mens (Dutch Human Rights Institute), *Opinion on Postcode-Based Life Insurance Premiums (Dazure Case)* (Utrecht, 2011) https://mensenrechten.nl accessed 21 September 2025.; Zuiderveen Borgesius, F., Baranowska, N., Hacker, P. and Fabris, A., 'Non-Discrimination Law in Europe: A Primer for Non-Lawyers' (2024) *Computers and Society* 5.

[55] As Todorova and others (2023) argue, "the European AI tango", when Europe's AI regulation strategy must constantly dance between preserving innovation and enforcing ethical and legal standards – a dynamic that mirrors the tension between efficiency and fairness in anti-discrimination law. See: Todorova, C. and others, 'The European AI Tango: Balancing Regulation, Innovation and Competitiveness' (2023) *HCAIep Proceedings* 2–8.

[56] Meding, K., 'It's Complicated: The Relationship of Algorithmic Fairness and Non-Discrimination Regulations in the EU AI Act' (2025) *arXiv* https://arxiv.org/abs/2501.12962 accessed 21 August 2025.

balance of power. Instead of claimants struggling to uncover evidence from opaque systems, public authorities are authorised to probe the inner workings of algorithms.

One step ahead, the Article 86 of the AI Act marks a significant step beyond the GDPR by explicitly recognising a "right to explanation" in the context of high-risk AI systems. It entitles any person adversely affected by a legally or similarly significant decision taken on the basis of such a system to obtain from the deployer a clear and meaningful account of two things: (i) the role played by the AI system in the overall decision-making procedure, and (ii) the main elements of the decision itself. In discrimination contexts, this provision directly responds to the enforcement choke point problem: instead of facing a black box, claimants are given a legal entitlement to information that can reveal whether the outcome was shaped by proxy variables or biased data. At the same time, the scope is limited—only Annex III high-risk AI systems are covered, with certain exemptions for law enforcement and where EU or national law provides otherwise. Nonetheless, Article 86 represents the EU's first horizontal, sector-independent recognition that explanations are essential for making anti-discrimination rights enforceable in the age of AI.

In this way, the AI Act has the potential to recalibrate the relationship between efficiency and equality in EU law. Accuracy may still count as a legitimate aim, but it no longer excuses neglect of fairness obligations—and providers face the prospect of direct regulatory scrutiny if they fall short.

IV. From AI to Platforms: How the DSA, DMA, and PWD Extend Equality Governance

Beyond the AI Act, equality and non-discrimination principles also appear—often implicitly — in other major pieces of EU digital platform regulation. These instruments show that concerns about systemic discrimination are increasingly embedded across the EU's digital lawmaking, even outside of explicit human rights legislation.

The DSA introduces a novel regulatory framework for addressing risks arising from the design and operation of VLOPs and VLOSEs. Section 5 requires VLOPs and VLOSEs—defined as services reaching at least 45 million active monthly users in the EU—to identify, assess, and mitigate what the regulation terms "systemic risks" (Article 34). While systemic risks are not exhaustively defined, the regulation provides an illustrative list that explicitly encompasses "any actual or foreseeable negative effects for the exercise of fundamental rights", including the right to non-discrimination under Article 21 of the Charter of Fundamental Rights of the EU. It situates discriminatory outcomes not merely as individual harms but as systemic risks

embedded in platform design, functioning, and use. Practices such as recommender system profiling, targeted advertising, content moderation, and biometric verification are thus placed under heightened scrutiny. Each of these technical and organizational choices can produce indirect or direct discriminatory effects, whether by reinforcing stereotypes, excluding individuals from opportunities, or amplifying unequal visibility across protected groups. To address these concerns, Article 35 imposes a proactive duty on VLOPs and VLOPSEs to implement targeted mitigation measures. These may include redesigning service features that contribute to discriminatory harms, modifying recommender and advertising systems to reduce bias, ensuring that synthetic content is clearly labelled, and strengthening content moderation practices. Importantly, platforms must also provide at least one recommender system option not based on profiling (Article 38), a safeguard designed to reduce the discriminatory effects of inferences drawn from personal or group characteristics. In addition, the obligation to maintain a publicly accessible repository of advertising practices (Article 39) increases transparency around microtargeting, allowing civil society and regulators to monitor whether advertising tools perpetuate inequality. However, this preventive duty risks being reduced to a box-ticking exercise. Even with the compulsory systemic risk audits under Article 37, platforms may comply formally while leaving structural inequalities untouched, unless supervisory authorities develop substantive standards for assessing discriminatory impacts.

A further innovation of the DSA is Article 40, which creates a right of access for vetted researchers to obtain data from VLOPs and VLOSEs. This provision is directly relevant to discrimination because it enables independent scrutiny of algorithms, recommender systems, and their risk profiles. Researchers can examine the logic, design, functioning, testing protocols, and limitations of these systems—effectively opening up otherwise opaque "black boxes." While access is limited to EU-compliant research frameworks and applies only to the very largest platforms, it nevertheless marks a significant step toward overcoming the evidentiary barriers that have long plagued anti-discrimination enforcement. By mandating transparency for research purposes, Article 40 allows discriminatory effects to be systematically documented and challenged, rather than left hidden behind proprietary opacity.[57]

For anti-discrimination law, the DSA is significant in two respects. First, it explicitly acknowledges that algorithmic discrimination can amount to a systemic risk, requiring structural rather than ad hoc remedies. Second, it broadens the accountability framework by linking discrimination not only to individual decision-making but also to the socio-technical

[57] Liesenfeld, A.: The Legal Significance of Independent Research Based on Article 40 DSA for the Management of Systemic Risks in the Digital Services Act. European Journal of Risk Regulation, 2024/16. 184–196.

architecture of platform services. This shift matters because discriminatory effects often arise less from intentional exclusion and more from design choices that embed historical bias or prioritise profit-maximising engagement logics.

The DMA is primarily an economic regulation aimed at curbing unfair practices of large online "gatekeepers", like Google, Meta, Apple, Amazon and more. Although it does not explicitly use the language of discrimination, its emphasis on fairness and equal access indirectly promotes equality. By prohibiting practices such as self-preferencing or discriminatory access to business users, the DMA safeguards equal treatment in digital markets. The DMA further operationalises fairness obligations in algorithmic infrastructures by regulating gatekeeper-controlled ranking systems. Article 6(5) DMA prohibits self-preferencing and requires that rankings be transparent, fair, and non-discriminatory. Unlike the more cumbersome enforcement of Article 102 TFEU—which requires lengthy investigations to prove abuse of market dominance—this provision establishes strict, ex ante duties, thereby enhancing compliance by deterrence. The non-discrimination requirement extends beyond mere transparency: when rankings directly concern individuals or groups, it overlaps with classical equality law prohibitions on discrimination based on protected characteristics; when rankings concern business users or products, it imports a competition-law understanding of non-discrimination akin to FRAND (fair, reasonable, and non-discriminatory) obligations. In this way, the DMA bridges competition and equality law, addressing both anti-competitive bias and systemic risks of exclusion. The purpose here is not to enhance accuracy, but to limit data advantages that could further entrench market power. This creates a structural tension: the AI Act demands richer, more representative datasets to ensure fairness, while the DMA deliberately curtails the accumulation and use of data by dominant platforms to preserve competition. The irony is that the very feature that makes AI models accurate—access to large, diverse datasets—is simultaneously what makes gatekeepers too powerful.[58]

In this sense, the DMA contributes to a broader digital equality framework, ensuring that market power cannot be exercised in a way that disadvantages certain groups of users or competitors. This demonstrates that the DMA, though competition-focused, embeds equality principles by structurally safeguarding non-discriminatory treatment in core digital infrastructures.

The EU's legal framework on equal treatment extends beyond the founding treaties and the Charter of Fundamental Rights to secondary legislation, with particular emphasis on the labour

[58] Hacker, P., Cordes, J. and Rochon, J., 'Regulating Gatekeeper AI and Data: Transparency, Access, and Fairness under the DMA, the GDPR, and Beyond' (2022) *arXiv* https://arxiv.org/abs/2212.04997 accessed 21 September 2025.

market. Employment conditions—such as access to work, remuneration, and promotion—are areas where structural inequalities persist. The rise of digital labour platforms, accelerated by the COVID-19 pandemic, has created new forms of flexible yet precarious employment. These arrangements often exclude traditional guarantees such as minimum wage or sick leave and render workers dependent on platform-based algorithmic management. The PWD seeks to address these vulnerabilities by regulating working conditions in platform employment, with particular regard to algorithmic governance. By creating a presumption of employment where platforms exercise algorithmic control, the PWD ensures that platform workers fall under existing labour law protections, including EU non-discrimination directives. Algorithmic decision-making, while efficient in matching supply and demand, introduces risks of subordination, opacity, and discrimination. Studies reveal that platform algorithms reproduce and exacerbate existing labour market inequalities—such as gender pay gaps, disadvantages for carers, migrants, and workers with limited digital skills—through mechanisms like continuous availability requirements, one-sided evaluation systems, and biased ranking criteria.[59] Article 1(1) of the PWD underlines the need for transparency, fairness, and accountability in platform governance. The PWD on improving working conditions in platform work explicitly addresses the risks of algorithmic management, including discrimination. Article 6 of the text introduces obligations on platforms to ensure transparency and fairness in algorithmic decision-making, while Recital 37 highlights the potential for bias and unequal treatment embedded in algorithmic governance. As discussed earlier, the AI Act classifies employment-related AI systems—such as recruitment tools and algorithmic management platforms—as "high-risk." In parallel, the PWD can be read as a sector-specific complement: it translates the broader preventive logic of the AI Act into the labour context by creating a presumption of employment, mandating transparency in algorithmic decision-making, and ensuring that automated dismissals can be legally contested. Taken together, these instruments embed equality and fairness standards into algorithmic management, ensuring that platform workers enjoy protection comparable to those under traditional managerial practices.

The shortcomings of the EU's existing anti-discrimination directives—limited material scope, ease of justification, and evidentiary choke points—highlight why litigation-based enforcement alone cannot effectively address algorithmic discrimination. Hacker (2018) captures this problem by arguing that what is needed is "equal protection by design": embedding safeguards

[59] Barzilay, A. R. and Ben-David, A., 'Platform Inequality: Gender in the Gig-Economy' (2017) 47 *Seton Hall Law Review* 393, available at SSRN https://ssrn.com/abstract=2995906 accessed 21 August 2025.

against discrimination directly into the technical and organisational structures of AI systems.[60] This is precisely where the EU's new digital regulations (the AI Act, DSA, DMA, and, in part, the PWD) represent a shift. Unlike the older equality directives, which rely on ex post enforcement through individual claims, these instruments establish ex ante duties of dataset governance, transparency, and systemic risk oversight. In doing so, they begin to translate the principles of non-discrimination into the everyday governance architecture of digital technologies themselves, complementing but also moving beyond the traditional equality acquis.

VI. Reframing Non-Discrimination: The AI Act and the Platform Regulations

The analysis undertaken in this paper demonstrates that the AI Act does not establish an entirely new body of anti-discrimination law. Instead, it operationalises and extends the existing acquis of equality and non-discrimination into the algorithmic domain. By classifying certain AI systems as high-risk, prohibiting others outright, and embedding requirements of data quality, transparency, and oversight, the Act translates long-standing legal principles of equal treatment into the regulatory language of risk management and technological governance. In this sense, it functions less as a novel regime and more as the algorithmic articulation of obligations already familiar from EU non-discrimination law.

At the same time, situating the AI Act alongside the DSA, the DMA, and the PWD reveals a broader trajectory that may be described as a "digital equality turn" in EU law. Each of these instruments incorporates, in different forms, concerns about algorithmic fairness, systemic bias, and unequal treatment in digital environments. Read together, they allow us to conceptualise the AI Act and related regulations not only as tools of technological regulation but also as the EU's first genuine attempt at a digital equality law. This reframing highlights the originality of the new digital regulations: while they do not displace existing equality directives, they re-contextualise them for an age in which algorithmic decision-making increasingly mediates access to work, services, and opportunities.

Looking ahead, this reinterpretation has practical consequences for law enforcement and regulatory practice. Courts will need to align the preventive logic of EU digital legislation with the reactive remedies of equality litigation, while supervisory authorities may be called to integrate their monitoring with the systemic risk assessments required under the DSA and the

[60] Hacker, P., (2018)

fairness obligations of the DMA. In this way, the AI Act invites not only doctrinal debate but also institutional innovation in how Europe safeguards equality in the digital era.

Ultimately, the effectiveness of this framework will depend not only on substantive obligations but also on institutional enforcement. Market surveillance authorities will be responsible for inspecting providers, demanding documentation, and even accessing source code to verify compliance. Equality bodies will remain central to investigating and litigating discrimination claims under existing directives. Courts, meanwhile, will continue to apply proportionality analysis in concrete cases, interpreting whether technical compliance under the AI Act also satisfies equality law obligations. This fragmented landscape creates a clear need for coordination mechanisms: without structured collaboration between equality bodies, data protection authorities, and market surveillance authorities, discriminatory systems risk slipping through the cracks. Embedding such coordination into the enforcement architecture would ensure that “digital equality law” is not merely rhetorical but a practical governance framework capable of delivering on its promise.