İSTANBUL BİLGİ UNIVERSITY
INSTITUTE OF GRADUATE PROGRAMS
INFORMATION AND TECHNOLOGY LAW MASTER'S PROGRAM

DO REGULATORY EFFORTS IN EUROPE REGARDING ALGORITHMIC DISCRIMINATION CAUSED BY PRIVATE ACTORS OFFER EMANCIPATION FOR WOMEN AND OTHER MARGINALIZED GROUPS?

Deniz ERDEN
118692021

Prof. Dr. Leyla KESER BERBER

İSTANBUL
2022

Do Regulatory Efforts in Europe Regarding Algorithmic Discrimination Caused by Private Actors Offer Emancipation for Women and Other Marginalized Groups?

Avrupa'da Özel Kişi Aktörler Tarafından Sebebiyet Verilen Algoritmik Ayrımcılık ile İlgili Yasal Düzenleme Çabaları Kadınlar ve Diğer Marjinalize Edilmiş Gruplara Kurtuluş Sunuyor Mu?

Deniz ERDEN
118692021

Tez Danışmanı :	**Prof. Dr. Leyla KESER BERBER** İstanbul Bilgi Üniversitesi	(İmza) .....
Jüri Üyeleri :	**Dr. Öğr. Üyesi Mehmet Bedii KAYA** İstanbul Bilgi Üniversitesi	(İmza) .....
	Dr. Öğr. Üyesi Sezen KAMA IŞIK İstanbul Medeniyet Üniversitesi	(İmza) .....

Tezin Onaylandığı Tarih : 19 Ekim 2022
Toplam Sayfa Sayısı : 303

Anahtar Kelimeler (Türkçe)	Anahtar Kelimeler (İngilizce)
1) Algoritmik Ayrımcılık	1) Algorithmic Discrimination
2) Kişisel Verilerin Korunması	2) Personal Data Protection
3) Ayrımcılık yasağı	3) Antidiscrimination Law
4) Kesişimsel feminizm	4) Intersectional feminism
5) Elektronik İşe Alım	5) E-Recruitment

TABLE OF CONTENTS

ABBREVIATIONS

ADM	Automated/Algorithmic Decision Making
AI	Artificial Intelligence
AIA	Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act)
B2B	Business to Business
BLM	Black Lives Matter
CAHAI	Ad hoc Committee on Artificial Intelligence
CFREU	Charter of Fundamental Rights of the European Union
CJEU	Court of Justice of the European Union
CLS	Critical Legal Studies
CoE	Council Of Europe
CRT	Critical Race Theory
DL	Deep Learning
DPL	Data Protection Law
DPD	Data Protection Directive (Directive no. 95/46/EC)
DPIA	Data Protection Impact Assessment
DSA	Proposal for a Regulation on a Single Market For Digital Services (Digital Services Act)
ECHR	European Convention on Human Rights
ECtHR	European Court of Human Rights
EDPB	European Data Protection Board
EDPS	European Data Protection Supervisor
EEOC	The U.S. Equal Employment Opportunity Commission
E.g.	Exempli gratia
Et. al.	Et alia
EU	European Union
EUDPR	Regulation (EU) 2018/1725 on the Protection of Natural Persons with regard to the Processing of Personal Data by the

	EU Institutions, Bodies, Offices and Agencies and the Free Movement of Such Data
FLT	Feminist Legal Theory
GDPR	General Data Protection Regulation
HR	Human Resources
i.e.	Id est
LED	Directive (EU) 2016/680 - Data Protection Law Enforcement Directive
LGBTI+	Lesbian Gay Bisexual Transgender Intersex and More
ML	Machine Learning
NLP	Natural Language Processing
OBA	Online Behavioural Advertisement
OECD	Organisation for Economic Cooperation and Development
TEU	Treaty on European Union / Maastricht Treaty
UNGPs	United Nations Guiding Principles on Business and Human Rights
US	United Stated of America
WMG	Women and Other Marginalized Groups
WP29	The Article 29 Data Protection Working Party

ABSTRACT

Discovery of automated extraction and processing of personal data as a massively profitable economic activity led to the new online public spheres to be intermediated by a few large corporations making use of advanced technologies to collect, manage and analyse vast amounts of data in providing real-time services. This enables significant decisions about individuals made by private companies such as access to employment and credit to be automated. Machine learning algorithms used in these ADM processes have notoriously found to be discriminating against WMG. Some argue that it is easier to fix these systems to overcome bias that is an inherent property in humans, while others argue that they are more likely to perpetuate and cement power relations in the society without regulatory intervention. Europe has become a global trendsetter on tech regulation as well as being at the forefront of democracy and human rights protection. This thesis aims at evaluating the main legislations in the EU that would apply in e-recruitment context which falls under similarly significant category in Article 22 GDPR and likely to result in a high risk for the rights and freedoms of individuals, arguing that transparency and accountability obligations under GDPR could be interpreted in a manner to enhance both *ex ante* and *ex post* protection against risks of algorithmic discrimination and its human rights implications in a democratic society in combination with the EU antidiscrimination law. On the other hand, several remaining loopholes are identified, by critically questioning the role of power relations in shaping underlying values and methods of the applicable law, technologies and business models by taking an intersectional feminist stance.

Keywords: Algorithmic Discrimination, Personal Data Protection, Antidiscrimination, Intersectional feminism, E-Recruitment

ÖZET

Kişisel verilerin otomatik yollarla elde edilmesi ve işlenmesinin muazzam ölçüde karlı bir ekonomik aktivite olarak keşfi, yeni çevrimiçi kamusal alanlara erişimin, veri toplama, veri yönetimi ve analizinde gelişmiş teknolojiler kullanan az sayıda yer ve sosyal ağ sağlayıcı şirket üzerinden gerçekleşmesi sonucunu getirdi. Bu gelişme, gerçek kişiler hakkında özel şirketler tarafından verilen istihdam ve krediye erişim gibi önemli kararların otomasyonunu mümkün kılmakta. Bu otomatik karar alma süreçlerinde kullanılan makine öğrenmesi algoritmaları kadınlar ve diğer marjinalize edilmiş gruplar aleyhine ayrımcılığa neden olmaları ile bilinmekte. Bazıları, insanlara içkin bir özellik olan önyargıyla mücadele için bu sistemlerin düzeltilmesinin daha kolay olduğunu savunurken, diğerleri yasal düzenleme ile müdahale edilmemesi halinde toplumdaki güç ilişkilerini sürdürme ve sabitlemelerinin daha olası olduğunu öne sürmektedirler. Demokrasi ve insan hakları korumasında önde gelen Avrupa, teknoloji düzenlemelerinde global trend belirleyici konumuna gelmiştir. Bu tez, Genel Veri Koruma Tüzüğü'nün 22. maddesinde yer alan "benzer biçimde kayda değer şekilde etkileyen" kategorisine giren ve gerçek kişilerin hakları ve özgürlükleri açısından yüksek bir riske sebebiyet vermesi muhtemel olan elektronik işe alım bağlamında uygulanabilecek olan AB yasal düzenlemelerini değerlendirecektir. Bu değerlendirmede, Genel Veri Koruma Tüzüğünde düzenlenen şeffaflık ve hesap sorulabilirlik yükümlülüklerinin, Avrupa Birliği Ayrımcılıkla mücadele hukukuyla birlikte uygulandığında, algoritmik ayrımcılık riskleri ve insan hakları etkilerine hem *ex ante* hem de *ex post* koruma sağlayabilecek şekilde yorumlanabileceği savunulacaktır. Öte yandan, kesişimsel feminist bir bakış açısından, güç ilişkilerinin bu bağlamda uygulanan hukuk kurallarının temelinde yer alan değer ve yöntemlerin, teknolojinin ve iş modellerinin şekillenmesindeki rolü eleştirel bir sorguya tabii tutularak bu yasal düzenlemelerde baki kalan açıklar belirlenecektir.

Anahtar Kelimeler: algoritmik ayrımcılık, kişisel verilerin korunması, ayrımcılık yasağı, kesişimsel feminism, elektronik işe alım

INTRODUCTION

This thesis focuses on discrimination against women and other marginalized groups (WMG) caused by automated decision making (ADM)[1] and the applicable European law to such discrimination. Machine learning algorithms used in these processes have notoriously found to be discriminating against marginalized groups at the intersections of gender, race, class, ethnicity, sexual orientation, age, disability and other disadvantaging attributes from regulating online speech to access to important economic resources such as employment and credit. Even though they are intended to be fair, objective and neutral, these systems pose the risk of perpetuating and cementing power relations in the society which disadvantage WMG.

Employment context will constitute our lead example as it is where the EU antidiscrimination law offers the highest legal standards as it applies to the broadest set of groups namely ethnicity, gender, religion or belief, disability, age and sexual orientation. In general, discrimination on the ground of gender appears to be the most protected across different areas addressed by the antidiscrimination law.

The new online public spheres are intermediated by a few large corporations who have business models based on automated extraction and processing of personal data. Large social media platforms have proven to be the new public spheres where people look for a job, find other like-minded people, organize and mobilize around causes, in other words where "freedom of opinion, expression

[1] We will also use algorithmic decision making as a term and will refer to it also as ADM which interchangeably means both terms.

and information" and "right to peaceful assembly and association" are exercised;[2] but also they are the spaces of mass corporate surveillance, of constant collecting, processing and sharing of personal data including output data of inferential and predictive analytics by various actors. When vulnerable individuals in need of a job are evaluated with the data collected from this quasi-public sphere, it is likely to have an effect on their individual and collective online activities to challenge the reasons of and to overcome their vulnerabilities in the same space.[3] This may not only cause a chilling effect on socially valuable behaviour[4] in a democratic society but also undermine the democracy by creating a vicious cycle which puts society in the loop.[5] Thus, we identify a disproportionally high risk by ADM processes used by private parties on the fundamental rights and freedoms of marginalized others such as women, LGBTI+ individuals, members of the working class, people in disadvantaged racial and ethnic groups, subordinated people in the developing countries. Thereon, we aim at assessing the sufficiency of the applicable law to prevent from such effect while seeking to reconcile the same groups' interests in data-based solutions within AI-assisted hiring and recruitment context. Can algorithmic systems also act as emancipatory tools or are they destined to be tools of oppression?

Automation in hiring and recruitment was introduced as a tool for combatting human bias and subjective discriminatory decisions. However, marginalized

---

[2] CAHAI, "Ad Hoc Committee on Artificial Intelligence (Cahai): Feasibility Study," Council of Europe CAHAI(2020)23 (2020), https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da.

[3] Anya Prince and Daniel B. Schwarcz, "Proxy Discrimination in the Age of Artificial Intelligence and Big Data," *Iowa Law Review* 105 (2020).

[4] Ibid., 1292.

[5] Kirsten Gollatz, Felix Beer, and Christian Katzenbach, "The Turn to Artificial Intelligence in Governing Communication Online," in *HIIG Workshop Report* (Alexander von Humboldt Institute for Internet and Society (HIIG), 2018).

others have notoriously faced discrimination in employment due to cisgender able-bodied male attributes being taken as norm to describe a good and trustworthy employee based on data that is the product of past and on-going discriminatory practices. This is problematic when all data generated by the candidate knowingly or unknowingly becomes available to infer their[6] attributes and predict effects of these attributes to a possible future performance in the applied position. Moreover, an algorithm may decide not even showing the job advertisement to some groups based on their assessed interests. They might be eliminated even without knowing. AI-assisted hiring and recruitment enable such processes which are likely to work against WMG especially when the algorithm is trained with the historical data seen in the case of Amazon's AI recruiting tool that showed heavy bias against women.

In order to be able to govern ethical, unbiased and non-discriminatory uses of data in AI driven innovations, it is crucial to understand the technological process in which the data flows, conceptions and definitions of gender, sexual orientation, race, class and other interlinked oppressions, as well as why and how they are embedded into technological design. Law, as itself has been criticized as playing a key role in maintaining such oppressions by design.[7] However, law that applies to automated decision making and predictive technologies is yet at its embryological state. This is problematic as such technologies are already showing their effects, but also is an opportunity to find new ways of legal design that would work for protecting the interests of the most oppressed.

[6] In this study we will use the gender neutral pronouns of they/them/theirs where the text calls for a singular pronoun.

[7] Sandra Fredman, Women and the Law (New York: Oxford University Press, 1997); Catharine A. MacKinnon, Toward a Feminist Theory of the State (Cambridge, Massachusetts: New York University Press, 1989).

Moreover, very little of policies, legal[8] and scholarly works, which tackle algorithmic discrimination, have a gender perspective nor apply gender theory. Various types of discrimination against marginalized others have a long history and they are amplified rather than being generated by the new technologies. Therefore there are decades of research on how gender ideology is embedded in social structures such as family, religion, law, economy, and class, as well as in scientific research and technology that have been ignored by some recent studies which try to remove bias from algorithms.[9] Recently, Internet Policy Review published a special issue titled "Feminist data protection" which "is not an established term or field of study" as put in words by its editors,[10] however it is one of the most clear calls to open data protection to discussion about "how it might be understood, critiqued and possibly reimagined in feminist terms."[11] In our opinion, intersectional feminist theory is crucial to fill the gap between regulatory efforts and the real people who require the protection that is aimed to

[8] As an exception, see the report "Human rights, democracy, and the rule of law assurance framework for AI systems: A proposal" by David Leslie et al., "Human Rights, Democracy, and the Rule of Law Assurance Framework for Ai Systems: A Proposal," ArXiv abs/2202.02776 (2022). which was submitted to the Council of Europe in September 2021.

[9] University of Cambridge Leverhulme Centre for the Future of Intelligence's report titled "AI and Gender: Four Proposals for Future Research" points out to the low amount of research which uses gender theory on discrimination caused and biased decisions made by automated decision making despite a bounty of research on these topics. According to the report, law and policy are among areas where there is a lack of research using gender theory in AI research. On the other hand, it can be argued that scholars who are focussing on the discriminatory and marginalizing effects of ADM benefit from feminist theories while not citing these works directly but by citing other works and legislation which were affected by them earlier. We have not conducted additional research on the relations of citations in scholarly works regarding algorithmic discrimination and those works which cite feminist research, such further research could be very beneficial for the field. We believe, claiming a direct link between feminist theory and the law that applies to algorithmic discrimination is nevertheless rare and important.

[10] Jens T. Theilen et al., "Feminist Data Protection: An Introduction," Internet Policy Review 10, no. 4 (2021): 2.

[11] Ibid., 4.

be provided by such regulation. In this respect, this study draws upon and adds on to emerging voices that link intersectional feminist perspectives to regulatory efforts in Europe that focus on bias in algorithmic systems and discrimination harms thereby aiming at an enhanced dialogue in between at a theoretical level, which we find necessary to move to the more practical level of seeking useful ways of governance of algorithmic systems that would work for the most marginalized individuals and groups.

RESEARCH PROBLEM

We will argue that data protection law stands on a crucial spot to offer a legal toolbox in preventing socially deep rooted inequalities to creep into the future in amplified forms while acknowledging its shortcomings and providing an overview of other crucial legal tools such as protection of fundamental rights and freedoms, anti-discrimination law and technology specific laws such as the proposed DSA and AIA of the EU. Thus, we aim at joining the ongoing conversation about mitigating impacts of ADM systems on democratic societies and whether there is need for new regulation in order to ensure creating a just future, with a twist of taking an intersectional feminist perspective.

Thus in this study we aim at asking, discussing and when possible answering the following questions:

- How can algorithmic systems used by private entities put fundamental rights of WMG at risk and cause discrimination?
- How do protection of related fundamental rights, namely right to privacy, non-discrimination, freedom of speech and assembly, interplay? What are the shortcomings of the rights discourse in addressing algorithmic discrimination?
- How and why does anti-discrimination law fall short in addressing algorithmic discrimination?
- Can data protection law (GDPR) address algorithmic discrimination, protect

affected fundamental rights and assist anti-discrimination law? What are the shortcomings of the GDPR?

- Are the new regulations on the horizon namely AIA and DSA likely to be able to close these gaps and overcome these shortcomings?

- What can be done in this current situation? What policies and laws should we promote and what actions should we take for short-term and long-term feminist legal gains? Can we turn these crisis into an opportunity for creating feminist futures?

THEORETICAL ORIENTATION

We acknowledge some problematic aspects of the mainstream legal scholarship that tackle algorithmic discrimination. One of them is taking the concepts of law as granted. We are convinced that joining the discussion on this new phenomena which is in part deep rooted in the past, we need to look beyond the legal concepts and tools of today without overlooking where they are coming from in the first place. We find feminist discussions and critiques of technology and law to offer the best philosophical path to achieve such broad perspective.[12] Our main concern is to understand how algorithmic systems are changing the power relations, if they do at all, to the favour or the detriment of the oppressed.

As emphasized in Michel Foucault's work and in feminist jurisprudence, power does not begin and end with the State. Today, algorithmic systems are mostly developed and deployed by private companies, and while the research has shown how significant their oppressive effect on WMG in the society can be, they do not

[12] Os Keyes finds it impossible to think about ethics and technology without engaging with feminist literature. According to them, this would look like a taxi driver being not interested in engaging with wheels, as they would have most of what they need but wouldn't be able to get to anywhere. See, Os Keyes, interview by Kerry Mackereth and Eleanor Drage, 2022. <https://www.thegoodrobot.co.uk/post/os-keyes-on-avoiding-universalism-and-silver-bullets-in-tech-design> Accessed on 10 October 2022.

share the same level of public scrutiny as public institutions. Thus, this study focuses on risks of ADM by private actors. Moreover, we aim at questioning how these systems affect the most structurally and historically disadvantaged members of the society by following Kimberly W. Crenshaw's historical statement that is when the most disadvantaged enter, we all enter.[13] Therefore, an intersectional feminist approach is strategically chosen.

ADM systems are data-driven technologies and data, as a matter of fact, are biased against those who deviate from the white, male, cis-gender, able norm in the western societies.[14] As Caroline Criado Perez shows, one way this occurs is as a gender data gap which results in the invisibility of WMG and their perspectives to not be taken into account in making any kind of decisions from design of everyday objects to crucial processes to exercise fundamental rights and freedoms, autonomy and human dignity.[15] When we turn to change the societal codes by regulation, there occurs another problem that the law itself also suffers from gender data gap in its design and has not always been developed with women's -

---

[13] Kimberly Williams Crenshaw, "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics," University of Chicago Legal Forum 1989, no. 1 (1989).

[14] Lisa Gitelman, ed. "Raw Data" Is an Oxymoron (Cambridge, Massachusetts: The MIT Press, 2013); Caroline Criado-Perez, Invisible Women : Data Bias in a World Designed for Men (London: Vintage, 2020); Catherine D'Ignazio and Lauren F. Klein, Data Feminism (Cambridge, Massachusetts: Massachusetts Institute of Technology, 2020); Joy Buolamwini and Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in
Commercial Gender Classification," in Conference on Fairness, Accountability, and Transparency, ed. Sorelle A. Friedler and Christo Wilson (Proceedings of Machine Learning Research, 2018); Cathy O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy (Penguin Books, 2016); Simone Browne, "Dark Matters : On the Surveillance of Blackness," (2015); Virginia Eubanks, "Automating Inequality : How High-Tech Tools Profile, Police, and Punish the Poor," (2018); Safiya Umoja Noble, Algorithms of Oppression: How Search Engines Reinforce Racism (New York: New York University Press, 2018).

[15] Criado-Perez, Invisible Women : Data Bias in a World Designed for Men

and of others' who deviate from the norm - needs, life experiences and perspectives in mind.[16] Here then, we face a double knot in our efforts of regulating ADM in an inclusive and non-discriminatory way. Thus it is not surprising that scholars who wrote on ADM and feminist data protection have been calling for new ways of doing law, especially for a collective understanding of rights and redress mechanisms.

Achieving a level of transparency is needed to evaluate whether these systems conform with the applicable law, to give the chance to those who are affected by them to challenge these systems before the courts, therefore to enhance the autonomy and the agency of the affected people. On the other hand, we are informed by the feminist scholarly work on the critique of a liberal law approach that takes a rational, autonomous individual as the norm and leaves it to this hypothetical individual to protect their own rights.[17]

As it has been acknowledged by Shoshana Zuboff[18] and Julie E. Cohen,[19] the new capitalism that is surveillance capitalism or informational capitalism, quite similarly to previous versions, does not play on a level ground. However, this capitalism has its own unique ways of domination and any analyses that takes power relations stemming from digital spheres into account, must acknowledge this broader perspective as well.

---

[16] MacKinnon, Toward a Feminist Theory of the State; Fredman, Women and the Law.

[17] Women and the Law; D. Kelly Weisberg, Feminist Legal Theory : Foundations (Philadelphia: Temple University Press, 1993).

[18] Shoshana Zuboff, The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power, First ed. (New York: PublicAffairs, 2019).

[19] Julie E. Cohen, Between Truth and Power: The Legal Constructions of Informational Capitalism (New York: Oxford University Press, 2019).

RESEARCH METHOD

Methodologically this study follows two intertwining routes: (1) Herbert Buckert's Information Technologies Law methodology,[20] (2) feminist (legal) methodology.[21]

Applying the Information Technologies Law Method, the methodological goal of this thesis is to provide for description and evaluation of the relationship between the information produced by ADM systems and law itself as information from the perspective of WMG. This goal brings two main challenges. First, algorithmic systems such as those used in recruitment and ad delivery are mostly black boxes [22] which makes it really hard to have an insight to how information flows and transforms into information in the form of significant decisions that would affect the lives of individuals. We neither intend nor have the capability to conduct any empirical research on this topic, thus we make use of literature in these areas to support our doctrinal research. We pay special attention to ways the law, especially the data protection law, can be instrumentalized in unlocking these black boxes and providing more transparency into how the personal data flow within these complex systems.[23] In this regard, we challenge the primus inter

[20] Herbert Burkert, "Information Law: From Discipline to Method," Berkman Center Research Publication No. 2014-5, U. of St. Gallen Law & Economics Working Paper No. 2014-02 (2014), htpps://ssrn.com/abstract=2402866.

[21] Katharine T. Bartlett, "Feminist Legal Methods," Harvard Law Review 103, no. 4 (1990); D'Ignazio and Klein, Data Feminism.

[22] Frank Pasquale, Black Box Society : The Secret Algorithms That Control Money and Information, First Harvard University Press paperback edition. ed. (Cambridge, Massachusetts: Harvard University Press, 2016).

[23] See the scholarly discussion on whether a right to explanation exists in the GDPR (Bryce Goodman and Seth Flaxman, "European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"," AI Mag. 38 (2017)., Sandra Wachter, Brent Mittelstadt, and Luciano Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data

pares[24] position of right to privacy and locate right to data protection at the centre of our normative-functional evaluation. This draws upon Gutwirth and De Hert's[25] identification of "privacy as an opacity tool", as opposed to "data protection as a transparency tool" alongside with the critique of privacy in feminist jurisprudence.[26] The second challenge is the standpoint epistemology.[27] As it has been insisted upon by black scholars and activists, discrimination does not occur on just one basis and it is usually more complex and multi-layered than the single-axis way antidiscrimination laws deal with it. Feminism, started with a claim to universality and called for an international sisterhood for all women of the world to gather around feminist values. However, first feminist scholarly works failed to consider that many women's experiences of the world are also determined by other

---

Protection Regulation," International Data Privacy Law 7, no. 2 (2017)., Andrew D. Selbst and Julia Powles, "Meaningful Information and the Right to Explanation," ibid., no. 4; Lilian Edwards and Michael Veale, "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For," Duke Law & Technology Review 16, no. 18 (2017); Margot E. Kaminski, "The Right to Explanation, Explained," Berkeley Technology Law Journal 34 (2018); Gianclaudio Malgieri, "Right to Explanation and Algorithm Legibility in the Eu Member States Legislations," SSRN Electronic Journal (2018); Andrew D. Selbst and Solon Barocas, "The Intuitive Appeal of Explainable Machines," Fordham Law Review 87 (2018); Sandra Wachter, Brent Mittelstadt, and Chris Russell, "Counterfactual Explanations without Opening the Black Box: Automated Decisions and the Gdpr," Harvard journal of law & technology 31 (2018); Bryan James Casey, Ashkon Farhangi, and Roland Vogl, "Rethinking Explainable Machines: The Gdpr's 'Right to Explanation' Debate and the Rise of Algorithmic Audits in Enterprise," Berkeley Technology Law Journal 34 (2018)..

[24] Sandra Wachter, "Privacy: Primus Inter Pares — Privacy as a Precondition for Self-Development, Personal Fulfilment and the Free Enjoyment of Fundamental Human Rights," (2017), https://ssrn.com/abstract=2903514.

[25] Serge Gutwirth and Paul De Hert, "Regulating Profiling in a Democratic Constitutional State," in Profiling the European Citizen: Cross-Disciplinary Perspectives, ed. Mireille Hildebrandt and Serge Gutwirth (Springer, 2008).

[26] Weisberg, Feminist Legal Theory : Foundations.

[27] Sandra Harding, "Rethinking Standpoint Epistemology: What Is "Strong Objectivity"?," in Feminist Epistemologies (Thinking Gender), ed. E. Potter L. Alcott (New York: Routledge, 1993).

dimensions of their identities such as race, class, sexuality, age, religion and geography.[28] Thus, feminist scholarship have had lively discussions on the importance of considering how these intersecting identities can affect the way different women experience oppression and how these intersections may cause amplification of these experiences.[29] This is one of the reasons, why we can benefit from feminist legal theory in questioning how facially neutral systems cause discrimination and how disadvantages are not static and single-axis but dynamic and cumulative. Then we need to also deal with the known epistemological problem about how to acquire knowledge about experiences of others that us ourselves have not experienced. In order to deal with this problem, we combine taking a recipient oriented view in our information-functional evaluation with the theory and methods of intersectional feminism. On this account, we anticipate expectations of the marginalized groups from ADM to its outcomes that may affect them. However, while expectations might differ and clash for different groups, outcomes may also remain unforeseeable and hard to detect. Being informed by intersectional feminist theory and methodology helps with identifying merging expectations.

On this account, intersectional feminist theory and methodologies[30] that are generated in order to cope with the problems identified above are helpful for situating our research as relevant as possible for the most oppressed as gender always intersects with other categories of oppression.[31] This requires to start with

[28] D'Ignazio and Klein, Data Feminism.p.4

[29] Crenshaw, "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics."

[30] Bartlett, "Feminist Legal Methods."

[31] Bianca Prietl, "Big Data: Inequality by Design?" (paper presented at the Weizenbaum Conference 2019 "Challenges of Digital Inequality - Digital Education, Digital Work, Digital Life", Berlin, 2019).

a positionality analysis;[32] in other words, our first step is to situate ourselves as recognized by feminist standpoint theory.[33] By doing so, we acknowledge that our perspectives and experiences shape the outcome of our work, in other words that the knowledge we produce is situated. The writer of this study is a cisgender able-bodied woman. She experiences significant privilege from her ableness, education and by coming from a middle class and well-educated family. These privileges and identifying as cisgender align with dominant group positions. In this regard, we cannot talk about for example poverty, disability or non-binary/queer identities from experience. On the other hand, she experiences disadvantage based on her gender and caring responsibilities as a mother, her age, as well as her immigrant status and her ethnicity in particular within the Turkish immigrant in Germany context. She has felt marginalized on the basis of her religious belief and political views during her earlier life in Turkey. As it was once put by Katharine T. Bartlett "to understand human diversity, however, is also to understand human commonality."[34] In our belief in co-liberation, we acknowledge that the human universal is difference[35] and we commit ourselves to question the neutrality, objectivity and universality of the concepts, legal norms and technologies we tackle, from the perspectives of all the oppressed groups through an effort of educating ourselves on oppressions that we have not experienced first-hand and challenging the role of our group identity in maintaining such oppressions. We take a position that seeks justice for all with no one left behind no matter how "niche" they identify or identified.

---

[32] See D'Ignazio and Klein, *Data Feminism*; Leslie et al., "Human Rights, Democracy, and the Rule of Law Assurance Framework for Ai Systems: A Proposal," 218.

[33] Donna Haraway, "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective," Feminist Studies 14, no. 3 (1988); Harding, "Rethinking Standpoint Epistemology: What Is "Strong Objectivity"?."

[34] Bartlett, "Feminist Legal Methods."p.886

[35] Frank I. Michelman and Kathleen M. Sullivan, "The Supreme Court, 1985 Term," ibid.100, no. 1 (1986).

In order to live up to our commitment, we have not carried out empirical methods, however we have benefited from the empirical research conducted generally upon bias and discrimination identification and mitigation in algorithmic systems, as well as research that cast light in particular on how e-recruitment works. Moreover, we have benefited from the work and experience of those researchers who have experienced forms of oppression that we have not such as racism or heteronormativity. We have mostly used methods of doctrinal research, and support our stance if we take any with the case law and secondary sources such as statistics, documentaries, policy reports, opinions such as those of Article 29 Working Party (WP29)/European Data Protection Board (EDPB) and European Data Protection Supervisor (EDPS), white papers and news about new legislative efforts.

Technology

Individual automated decisions are not always but most of the time based on profiling. We do not focus on ADMs that are not based on profiling as they do not really pose a new kind of discrimination risk. We focus on automated decisions that are based on a profile (struggling single mother) or a score (if the score is less than x then reject) assigned to an individual. These profiles and rules about the calculation of the scores are derived from other individuals' behaviours and other available data and then applied to the individuals that are subjected to ADM[36]. We will be using *algorithm* and *model* as interchangeable terms and we will mostly focus on machine-learned models.

[36] The data upon which these profiles and scores are based, are made available to those who control access to and the flow of information through principles of "surveillance capitalism," a term coined by Zuboff Zuboff, The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.. Thus, talking about surveillance and capitalism requires a questioning of power relations and a focus on privacy as we do, as a result of the very technologies we have chosen to understand.

Legal Framework

Algorithmic discrimination is mostly data-based, thus the legal scholarship tackling the issue has been largely focusing on data protection laws alongside with anti-discrimination laws. The Regulation (EU) 2016/679 (General Data Protection Regulation, hereinafter referred to as "GDPR") is the first piece of legislation to explicitly address algorithmic discrimination and to acknowledge ADM's possible effects on the fundamental rights and freedoms of individuals . While they do not directly mention algorithmic systems, EU antidiscrimination law, consisting mainly of the Recast Directive (Directive 2006/54/EC), the Race Directive (2000/43/EC), the Framework Directive (2000/78/EC), and the Goods and Services Directive (2004/113/EC), is a crucial part of European Union body of secondary law which addresses discrimination in general. The Charter of Fundamental Rights ("CFEU") has the same legal value with the founding treaties of the European Union, and the Convention for the Protection of Human Rights and Fundamental Freedoms ("ECHR") is a direct source of general principles of EU Law as stated in Article 6(3) TEU. Not to mention both documents include the right to equality and non-discrimination, as well as the right to privacy. Also, the CFEU includes a direct right to data protection.[37]

While discrimination, fairness, transparency problems of AI/ML/Big Data technologies have been foci points of scholarly work from various disciplines. They were also addressed by ethical guidelines of the EU and the OECD, where the terminology used including the terms of privacy and fairness are highly contested, context-based and vague that they actually shed little light on what is ethical and what is not.

[37] Evelyn Ellis and Philippa Watson, Eu Anti-Discrimination Law, Second ed. (New York: Oxford University Press, 2012).

Following these efforts drawing on ethics,[38] a couple of important regulatory developments took place on AI/ADM in Europe. Council Of Europe (CoE) founded an Ad hoc Committee on Artificial Intelligence (CAHAI) under the instruction of examining the feasibility and potential elements of a legal framework related to AI, based on the CoE's standards on human rights, democracy and the rule of law and published a Feasibility Report.[39] CAHAI completed its mandate on 3 December 2021. The final deliverable of CAHAI is titled "Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law."

On 21 April 2021, the European Commission published its proposal for a regulation called Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) (hereinafter referred to as "AI Act"). This marks another regulatory effort in Europe that explicitly mentions bias in algorithmic systems and discrimination harms thereby.

As mentioned above some intermediary services, and in particular online platforms have become the new quasi public spheres and Proposal for a Regulation on a Single Market For Digital Services (Digital Services Act) (hereinafter referred to as "DSA") which is amending the E-Commerce Directive (2000/31/EC) also mentions ensuring right to non-discrimination in these platforms as one of its main goals.

STRUCTURE OF THE THESIS

To sum, this study aims at understanding from an intersectional feminist

[38] There is an abundance of AI Ethics Guidelines as it was compiled by the Algorithm Watch available at https://inventory.algorithmwatch.org/ (Accessed on 25 March 2022).

[39] CAHAI, "Ad Hoc Committee on Artificial Intelligence (Cahai): Feasibility Study".

perspective the available solutions to the broad question of how could ADM processes be regulated and governed[40] in order to achieve corporate conduct in the direction of society-aware design and practice so that a disproportionately negative affect to the fundamental rights and freedoms of marginalized groups, such as discrimination in access to services and economic resources is prevented.

In this regard, the first chapter provides an overview of feminist and critical legal theories to law and their relation to technology,[41] a basic introduction to relevant technologies, how bias and discrimination occur in algorithmic systems and technological solutions offered to mitigate the risk of algorithmic bias and discrimination. In order to lay as a ground a slightly more detailed understanding of the technological and corporate processes that we would like to later discuss how the law applies thereto, we focus on the domain specific challenges in AI-assisted human resources management context with special focus on recruitment and job advertisements. Here we describe different phases of AI-based decision-making processes, the types of data and technology used in each phase,[42] as well as trade-offs in corporate decision-making. In doing so, technical process are mapped into phases; and general features, actors, threads and how oppressive practices are embedded within the technical and organizational process for each stage are explored. So that the applicable law and policies can be better identified

---

[40] See Jeanette Hofmann, Christian Katzenbach, and Kirsten Gollatz, "Between Coordination and Regulation: Finding the Governance in Internet Governance," New Media & Society 19 (2017). for the distinction and relation between regulation and governance.

[41] We start with introducing the feminist jurisprudence's arguments about why law itself is gendered in combination with why technology in general and code in particular is also gendered, gender theory will be applied from an intersectional feminist stance of positionality to understand how oppressive practices are embedded within code and law.

[42] First phase is the data collection phase where the input data is gathered. The second phase includes the system design, training of the ML algorithms, inferential and predictive analytics upon the result of which a decision is based. The third phase is where the output data occurs in form of an automated decision which is itself or based on a derived or inferred data or a prediction.

in the following chapters.

The second chapter provides an analysis of the general principles of law in Europe which constitute the basis of non-discrimination principle, the interplay of rights that are under threat from algorithmic discrimination in a democratic society and the responses developed in antidiscrimination law. We start with international framework of fundamental rights and freedoms to introduce the interplay between the related and contested concepts such as democracy, rule of law, privacy and equality. The familiarity in the more philosophical discussions into these concepts will assist our analysis of more practical regulation such as EU antidiscrimination law, GDPR, AI Act and Digital Services Act. We introduce the main concepts and shortcomings of antidiscrimination law when applied to algorithmic systems.

In the third chapter, we delve into the GDPR in seek of remedies for the shortcomings of the EU antidiscrimination law and while doing so we identify overlapping and data protection specific shortcomings. We pay specific attention on transparency that application of GDPR may provide through the right to explanation discussion and the accountability requirements. We advocate that Data Protection Impact Assessments (DPIA) are the best bet for *ex ante* intervention to algorithmic systems with possibly discriminatory effects. We argue that special categories of data/protected attributes approaches of data protection law/antidiscrimination law are limited to address the reality of the oppressed. We speculate that albeit not being enough, a combination of transparency rules, DPIAs in the form of algorithmic human rights risk assessment which also considers societal right at risk and mandate to NGOs may provide the tools for overcoming some of the main shortcomings of the antidiscrimination law. We briefly discuss what do the newly proposed regulations of the EU, namely AI Act and Digital Services Act (DSA) offer to assist and to further this existing legal framework.

Finally, we conclude with imagining an intersectional feminist future in which the stories of algorithmic oppression have a chance of turning into emancipatory

tales. What would the highlights of ethical, regulatory and corporate internal frameworks look like in such times?

CHAPTER 1

1 AUTOMATED DECISION MAKING (ADM) AND DISCRIMINATION

1.1 CODE IS LAW, LAW IS GENDERED

Before Lessig declared that the code was law, and before the discrimination by ADM has become a hot topic as the code was also gendered, feminist political and legal theorists have looked into human decision-making and claimed that law was gendered. We incline toward taking this as our starting point as it lays down the theories on coping with discrimination from the perspective of the WMG.

Alongside with Critical Race Theory (CRT), Feminist Legal Theory (FLT) is a subcategory of Critical Legal Studies (CLS) movement which began in 1970s[43] as the successor of American Realism. According to the CLS scholars, particular understandings of social and political life influence law which is an expression of thereof.[44] They criticized the idea that law and morality consist of neutral principles.[45] These principles may be regarded, at least within the liberal tradition, as general principles of justice, fairness or equality such as John Rawls's idea that each person should enjoy an extensive liberty in a harmony with a like liberty of others; or Ronald Dworkin's idea that each person deserves equal respect simply by virtue of being a person; or the idea that the U.S Constitution offers for each citizen "equal protection under the laws." Within the liberal tradition, it is generally agreed that such principles are true and what moral and legal theory should do is to articulate the correct formulation of them. Instead, Unger offered

[43] See Jefferson White and Dennis M. Patterson, Introduction to the Philosophy of Law : Readings and Cases (New York: Oxford University Press, 1999)..

[44] Ibid., 126.

[45] Jeffrie G. Murphy and Jules L. Coleman, Philosophy of Law : An Introduction to Jurisprudence (Boulder: Westview Press, 1990), 51.

total criticism as total understanding was not anyway possible.[46] According to Unger[47] there are numberless accounts of ways for each to explain the objects and events in the world. Therefore, CLS writers argued that "both legal and moral doctrines are in fact merely after-the-fact rationalizations of a basic and hidden social reality- namely political or power relations and the deep inequalities that such relations inevitably incorporate." Thus, "these function to protect the interests of those having power against those who are powerless by fostering the legitimating illusion that the power inequalities are just in some neutral rational sense".[48] Altman suggested in respect with law and politics that law is the mirror of our political culture and debates in the political arena are reflected to the legal area. Altman also criticized Dworkin's claim, that judges search for the best applicable rule for the case and apply it as subject to the commander rules, because even the process goes like in Dworkin's claim, judges will choose the best rule to apply from the most proper rule to their political views.[49] To summarize, the CLS perspective can be seen as a radical value relativism or scepticism, which holds that "once one fully understands the causal origins of values (as an outgrowth of power relations) one will no longer take value claims seriously"[50] and more importantly without any questions.

In modern democratic societies, the main source of law is the legislature which consists of the elected representatives of the people who are treated equally in their right to vote and to stand in elections. Equal right of "we the people" to govern ourselves is put into practice through periodical transfer of this right to our representatives who make, alter, amend and repeal laws on our behalf. How is this process of law-making similar to the process of algorithmic decision-making? We

---

[46] White and Patterson, Introduction to the Philosophy of Law : Readings and Cases, 127-29.

[47] Roberto Mangabeira Unger, Knowledge & Politics (New York: Free Press, 1975).

[48] Murphy and Coleman, Philosophy of Law : An Introduction to Jurisprudence, 52.

[49] Andrew Altman, "Legal Realism, Critical Legal Studies, and Dworkin," Philosophy and Public Affairs 15, no. 3 (1986): 231.

[50] Murphy and Coleman, Philosophy of Law : An Introduction to Jurisprudence, 52.

could start from its lack of diversity before moving onto lack of democratic .

In almost every country in the world, majority men are overrepresented in parliaments.[51] Caroline Criado Perez shows how this translates into a data gap that leaves the needs and life experiences of WMG out of law-making. Wherever there is a quota program for example to increase the number of women in the legislature, the results show more women in the parliaments which translate into more policies and laws are promoted directed to specific women's problems and beyond such as education and care in general; and the possible effects of new laws on women are more likely to be taken into account during the legislative process.[52] When we flip the coin though, underrepresentation of WMG in law-making continues to be one of the reasons that law, the output of the legislative process, is continuously gendered. However, this gendered-ness is disguised under a veil of objectivity and neutrality.[53] As argued by Catherine MacKinnon, while women's needs are being discussed under affirmative action plans, in reality, "every quality that distinguishes men from women is already affirmatively compensated" as they are taken granted to constitute the structure and values of the society.[54] For instance "biographies of men define workplace expectations and successful career patterns"[55] while women have to campaign for pregnancy or childcare to be recognized as normal parts of human life and face criticism for demanding

---

[51] Melanie Hughes, "Diversity in National Legislatures around the World," Sociology Compass 7 (2013).

[52] Criado-Perez, Invisible Women : Data Bias in a World Designed for Men 272.

[53] Hilary Charlesworth, Christine Chinkin, and Shelley Wright, "Feminist Approaches to International Law," American Journal of International Law 85, no. 4 (1991); Duncan Kennedy, A Critique of Adjudication (Harvard University Press, 1997); Catharine A. MacKinnon, "Feminism, Marxism, Method, and the State: Toward Feminist Jurisprudence," Signs 8, no. 4 (1983); Theilen et al., "Feminist Data Protection: An Introduction."; Fredman, Women and the Law; Bartlett, "Feminist Legal Methods."

[54] Catharine A. MacKinnon, Feminism Unmodified : Discourses on Life and Law (Cambridge, Mass.: Harvard University Press, 1987).

55 Ibid., 36.

privileged treatment by that.

In this line, one of the feminist strategies have been closing the gender data gap in decision-making positions by making possible that more women stand in elections and enter legislative office, as well as employing more women and breaking the glass ceiling by bringing more women in managerial positions. Similarly, to fight with bias in algorithmic systems more diverse teams in technology companies and more data on WMG are often suggested. Although it is a very important step as Wajcman puts it that the underrepresentation of women in technoscience affects how the world is made, one must also be careful while identifying the problems as data gap or bias to avoid falling into the trap of assuming easy fixes (such as collecting more data and diversifying the team behind the algorithmic systems or legislations) to structurally deep systematic problems.[56] Moreover, one must be cautious for not falling victim to white feminism,[57] which is to say injecting gender diversity with an assumption that including cisgender heterosexual able-bodied Western, white women would bring in the experiences of all the oppressed and solve all their problems which ends up being many times the case in forming diversity and inclusion boards in tech companies.[58] Neither non-intersectional feminism would solve the hostility of the law and/or the code against the women of colour and their bias problem in general, nor would individualised solutions automatically help bringing system level change.[59]

Law as a means to the political emancipation of women is actually paradoxical as feminists also reject the methodology and structure of legalism as they reinforce

---

[56] Judy Wajcman, Technofeminism (Cambridge: Polity, 2004).

[57] Cate Young, 2014, https://www.cate-young.com/battymamzelle/2014/01/This-Is-What-I-Mean-When-I-Say-White-Feminism.html

[58] Abeba Birhane and Olivia Guest, "Towards Decolonising Computational Sciences," Kvinder, Køn & Forskning 29 (2021).

[59] Ibid.; Theilen et al., "Feminist Data Protection: An Introduction."

and reproduce dynamics which play a role in women's subordination.[60] In this regard, a similarity can be drawn between legalism in law and the problem in ADM about algorithmic models being only as good as the data they had been fed, also known as "the garbage in, garbage out" problem. Once the input is biased, the outcome decision which takes its references from that biased data will also show bias. Thus, Scales suggested that feminist lawyers should speak openly about the politics of neutrality in law as sexism has not been just a legal mistake.[61] According to Martha Minow, justice has to be engendered which is achieved when judges (here we could read it also as law-makers, regulators, data scientists or coders) "admit the limitations of their own viewpoints, when judges reach beyond those limits by trying to see from contrasting perspectives, and when people seek to exercise power to nurture differences, not to assign and control them."[62] Similar to techno-pessimists, many feminist lawyers have been doubtful about law's capacity to be a tool in bringing radical system change as it is a tool of hegemonic control.[63] Feminists have also been critical about equal rights struggle as it bore the risk of confining their struggles to litigation and lobbying which would then loose the creativity needed to imagine a different system for the future. On the other hand, Olsen suggests that despite they have often appeared as rights achievements they were both resulting from real struggle, so it counts as feminists have gained ground in the legal arena as more are likely to come.[64] In this regard

[60] Margot Stubbs, "Feminism and Legal Positivism," Australian Journal of Law and Society 3 (1986): 82.

[61] Ann C. Scales, "The Emergence of Feminist Jurisprudence: An Essay," Yale Law Journal 95 (1986).

[62] Martha Minow and Donald C. Langevoort, "The Supreme Court, 1986 Term," Harvard Law Review 101, no. 1 (1987).

[63] Douglas E. Litowitz, "Gramsci, Hegemony, and the Law," BYU Law Review 2000 (2000).

[64] Frances Olsen, "Statutory Rape: A Feminist Critique of Rights Analysis," Texas Law Review 63, no. 3 (1984).

law is acknowledged as power/knowledge,[65] as in the case of technology not only shaping the world we live in, but also being shaped by it.[66] Furthermore, according to Polan,[67] it is not so much that laws must be changed, it is patriarchy that must be changed. Actions taken within the legal system cannot by themselves eliminate patriarchy, which is a pervasive social phenomenon. As male supremacy is not only represented in the legal arena, "legal efforts to end women's subordinate status cannot effectively challenge or cripple patriarchy unless they are undertaken in the context of broader economic, social and cultural change."[68] Looking through the other side of the coin, as long as law reinforces male power, elimination of patriarchal power and social control cannot be achieved.[69]

There are many feminisms but they all unite on the belief that the societies we live in are patriarchal, shaped and defined by men. Beyond this, feminist concerns are as diverse as women's experiences. As its most known motto "personal is political" shows feminist theory values personal experiences of women even as the starting point of theory exercised through the consciousness raising method. Feminist legal theory shares these features in its inquiry into law's complicity in maintaining women's subordination alongside with possibilities of instrumentalizing law for change.[70] FLT also shares with CLS scholarship a focus upon "the ways law legitimates, maintains, and serves the distribution and

---

[65] Annie Bunting, "Feminism, Foucault, and Law as Power/Knowledge," Alberta law review 30 (1969).

[66] Wajcman, Technofeminism.

[67] Diane Polan, "Toward a Theory of Law and Patriarchy " in Feminist Legal Theory : Foundations, ed. D. Kelly Weisberg (Philadelphia: Temple University Press, 1993), 425; ibid.

[68] Ibid., 425.

[69] Janet Rifkin, "Toward a Theory of Law and Patriarchy," Harv. Women's L.J. 3 (1980).

[70] Elizabeth Marie Schneider, "The Dialectic of Rights and Politics: Perspectives from the Women's Movement [1986]," New York University Law Review 61 (1986); Robin West, "Introduction to the Research Handbook on Feminist Jurisprudence," in Research Handbook on Feminist Jurisprudence, ed. Robin West and Cynthia Grant Bowman (Edward Elgar Publishing Limited, 2019).

retention of power in society."[71]

Feminist legal scholarship alongside with women's movements have been effective in using the main arguments of liberal political theory that are dominant in western democracies for bringing change into law and women's lives. In this regard, three liberal values have been the most important: (i) Individualism, (ii) equality, (iii) public/private distinction.[72] These main liberal values and commitments also constitute the main lines of criticism for law's complicity in women's subordination.

First of all, the main source of value in liberal thought is individuals who trust their own judgments and challenge authority. In this regard, individual freedom and autonomy come forth as main values. According to Dunn, the idea of autonomy in the sense of choosing freely for oneself sits at the core of democracy's power and appeal.[73] Autonomy as a value contains the "importance of self-determination and the pursuit of one's own understanding of the good life without societal or state-based censorial control".[74] In the ideal liberal society, common good is achieved through autonomous individuals making good choices for their lives. For liberal feminism, this commitment to the freedom and capacity of every individual to follow their own path without being bound with assigned characteristics of a group, provided fertile ground for challenging exclusion of women as a group from the public life. If women are free individuals too, they

---

[71] Heather Ruth Wishik, "To Question Everything: The Inquiries of Feminist Jurisprudence," Berkeley Journal of Gender, Law and Justice 1 (1985).

[72] Which is criticised for reflecting white middle class women's experience while the experiences of black men and women were not reflecting this distinction. See Crenshaw, "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics."

[73] John Dale Dunn, "Democracy: The Unfinished Journey, 508 Bc to Ad 1993," Contemporary Sociology 22 (1993).

[74] West, "Introduction to the Research Handbook on Feminist Jurisprudence."

could not be deterministically assigned a domestic role based on their "nature" as mothers and caregivers. However, the strong emphasis on personal autonomy also feeds to the overreliance on consent in regulating relations in the private realm. Social control mechanisms that perpetuate the subordination of women usually act upon their consent rather than apparent coercion.[75] Be it consumers, workers or women, allowing consent to bad deals, may go a long way to legitimising power relations that impose them in the first place.

Secondly, capacity for reason, an autonomous will and entitlement to be treated with dignity and respect are universally shared by all individuals. Thus, on these grounds all individuals are treated the same, but other than that as everyone is determined by their individual attributes and choices, no group identity can justify unequal and worse or better treatment under anti-caste principle. As a result of liberal feminist struggle, this understanding of formal equality has become the norm for liberal democracies in which women cannot be treated differently than men based on stereotypical assumptions based on their nature.[76] However, women and men are not the same in their biology and their acculturation related thereto. Overemphasis of liberal theory and law on formal equality, compromises antidiscrimination norms as they do not take into account women's attributes, experiences, and problems that stem from their differences from men under the veil of gender neutrality. Susan Mendus argued that women's underrepresentation in positions of power as opposed to their overrepresentation amongst the unemployed, low paid and the part-time workforce in all the democratic states around the world for example, may be the evidence for a deep gender bias in democratic liberal theory itself which actually fails in putting its anti-caste principle into practice as a result of taking male experiences as the neutral norm, deviance from which is labelled as disadvantage. Thus, removing disadvantages for creating a level playing field can also be interpreted as a strategy of

[75] Ibid.

[76] Ibid.

assimilation for all the differences from the norm, which is a losing game from the beginning for the WMG.[77] The ideal of democracy as the free market-place of ideas is based on diversity, as it has been recognized by John Stuart Mill, that pursuing equality must start with recognizing differences.[78]. Finally, one must be alert about both difference and equality being value-laden terms.[79]

Third, public authority of the state is acknowledged as a potential source of oppression, while the private realm of intimate, familial, and commercial relations is perceived as a place for free and autonomously self-determined individuals interacting with each other. In order to individual freedoms and rights to be exercised to the full-capacity, the private realm has to be protected against the intrusion of public authority. Based on this liberal commitment, reproductive rights to contraception and abortion are taken as women's choice issues, as well as LGBTI+ communities' sexuality, marriage and parenthood are to be protected as matters of the private. On the other hand, as law would try not to intervene much in domestic sphere in liberal theory, women, children, and sexual minorities have struggled in to receive the protection of state against violence they usually experience within private and intimate spheres such as marital rape or child abuse.[80]

Lawrence Lessig had warned against leaving the cyberspace unregulated to the hands of private companies which would give them chance to wield too much power and to control the society contrary to the libertarian ideals of the early Internet by using the code to regulate through the very architecture in which social interactions take place, to decide which values to flourish and which of them to be

---

[77] Susan Mendus, "Loosing the Faith: Feminism and Democracy," in Democracy the Unfinished Journey: 508 Bc to Ad1993, ed. John Dunn (New York: Oxford University Press, 1992).

[78] John Stuart Mill, "On Liberty," in On Liberty, ed. David Bromwich, George Kateb, and Jean Bethke Elshtain, Rethinking the Western Tradition. (New Haven: Yale University Press, 2003).

[79] Mendus, "Loosing the Faith: Feminism and Democracy."

[80] West, "Introduction to the Research Handbook on Feminist Jurisprudence."

excluded such as in the case that unprofitableness of anonymity replacing it in the very architectures of the web services with the identifiability to enable commerce.[81] One of the major liberal value claims had come from John Perry Barlow in the early days of the Internet. In "A Declaration of the Independence of Cyberspace," he denied consent to the regulation of the cyberspace where unwritten codes were providing order which made it possible "for anyone, anywhere to express his or her beliefs, no matter how singular, without fear of being coerced into silence or conformity." Barlow declared building a new World where identities had no bodies and all could "enter without privilege or prejudice accorded by race, economic power, military force, or station of birth."[82] Interestingly he forgot to mention gender. But women[83] entered the cyberspace with their bodies. From revenge porn[84] to rape threads as a response to their activism, their bodies have been violated and voices have been silenced in the privately controlled public sphere of the cyberspace.[85] Now that states are also counting on the codes of algorithmic systems by private companies to provide protection to WMG, Lessig's warnings and intersectional feminist critics of law and technology for being value-laden while appearing to be objective has to be kept in mind.

---

[81] Lawrence Lessig, Code 2.0 (CreateSpace, 2009).

[82] John Perry Barlow, "A Declaration of the Independence of Cyberspace," (1996), https://www.eff.org/cyberspace-independence.

[83] Alongside with transgender and gender non-conforming people.

[84] Danielle Keats Citron and Mary Anne Franks, "Criminalizing Revenge Porn " Wake Forest Law Review 49 (2014).

[85] Brandeis Marshall, "Algorithmic Misogynoir in Content Moderation Practice," (2021), https://eu.boell.org/sites/default/files/2021-06/HBS-e-paper-Algorithmic-Misogynoir-in-Content-Moderation-Practice-200621_FINAL.pdf; Amnesty International, "Toxic Twitter – Women's Experiences of Violence and Abuse on Twitter," https://www.amnesty.org/en/latest/news/2018/03/online-violence-against-women-chapter-3/.

1.2 ALGORITHMIC DECISION-MAKING PROCESS

In this part, we will first introduce the technologies used in ADM processes. Then we will look into the meaning of algorithmic bias and discrimination and identify how they occur in ADM processes. Finally, we will have an overview of where these new processes are located within the ongoing power relations at the societal level through the lens of old and new ways of social sorting, as well as surveillance and profiling technologies that run on big data.

1.2.1 Algorithmic Decision-Making (ADM)

Algorithmic decision-making systems (also referred to as automated decision-making systems) use algorithms to make decisions or to assist human decision-makers in their decisions by making recommendations or showing probabilities.[86]

Algorithms are "a specific form of instruction that leads to the solution of a mathematical problem."[87] They are encoded procedures for solving a problem by transforming input data into a desired output, based on specified calculations".[88] According to Alpaydın, an algorithm is "like a recipe for a dish."[89]

There are a number of algorithm types which can be used in ADM processes. First of all, in ***rule-based algorithms***, rules and subrules are expressed in an "if this then that" pattern and programmed into computer instructions by technical

---

[86] Jessica Wulf, Automated Decision-Making Systems and Discrimination: Understanding Causes, Recognizing Cases, Supporting Those Affected (Berlin: AlgorithmWatch, 2022), https://algorithmwatch.org/en/autocheck/. 5.

[87] Ibid.

[88] Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot, eds., Media Technologies: Essays on Communication, Materiality, and Society (United States of America: The MIT Press, 2014), 167.

[89] Ethem Alpaydın, *Machine Learning, Revised and Updated Edition* (Cambridge, Massachusetts: The MIT Press, 2022), 16.

experts and data scientists with the help of substance experts (such as lawyers, marketing teams, finance professionals etc.). In these algorithms human involvement is high as rules are pregiven. They are rather rigid or static and not well suited for dynamic environments and tasks while being predictable and easier to explain. A traffic rule where there is no value judgment required can be translated into an algorithm easily as: if someone's speed in a given road exceeds 80 km per hour, then this person has to pay a 100 € fine.[90]

Second, ***machine learning (ML) algorithms*** are distinguished by their ability to learn and adapt, evolve and improve autonomously as they can change the rules according to the new input and optimize outputs for any input without needing to be specifically programmed to do so. They are used like the rule-based algorithms in making automated decisions or creating more complex decision trees, but also in pattern recognition, profiling or predictive calculations on which again decisions can be based. There are various techniques (classification, clustering, regression, association) used in ML suitable for solving different problems. In identifying spam emails, an algorithm needs to learn which e-mails fall under the categories of spam and not spam using *classification techniques*. While classification techniques use a labelled dataset, *clustering techniques* are more fitting for unstructured and unlabelled datasets[91] as they form clusters by identifying strong commonalities and correlations among situations, objects or persons, as well as detecting anomalies.[92] For example, an algorithm tasked with identifying credit card fraud may use clustering techniques by finding out

[90] European Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law* (Publications Office, 2021), 33.

[91] See <https://www.javatpoint.com/clustering-in-machine-learning> Access date: 15 September 2022.

[92] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 33.

anomalies or a customer segmentation[93] algorithm can identify customers with similar interests and economic status. *Regression techniques* are used in predicting the future as opposed to classification techniques which refer to the past.[94] Credit risk of their customers may be calculated by financial institutions through the comparison of their personal data with all the other available data to be able to estimate the likelihood of repayment or default.[95] A recruitment tool may also use a similar technique to calculate for example the probability of candidates to stay more than 3 years in the company. Algorithms that suggest those who have bought the same book with you also liked this book or those who have watched these series also watched those use *association techniques*. These algorithms use affinity grouping,[96] which is that of mining a large database for such correlations which are independent events happening frequently in a way suggesting future behaviour.[97]

Thus, machine learning can be defined as "the ability of computer algorithms to learn from data and make predictions for new situations, and improve automatically through experience."[98]

The learning process may require involvement of humans (supervised learning)

---

[93] Clustering models are also known as segmentation. See Ana Canhoto and James Backhouse, "General Description of the Process of Behavioural Profiling," in *Profiling the European Citizen: Cross-Disciplinary Perspectives*, ed. Mireille Hildebrandt and Serge Gutwirth (Springer, 2008), 51.

[94] Ibid.

[95] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 34.

[96] Canhoto and Backhouse, "General Description of the Process of Behavioural Profiling." See also <https://tdwi.org/articles/2016/03/16/machine-learning-techniques-methods.aspx> Accessed on: 16 September 2022

[97] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

[98] Nadezhda Purtova, "The Law of Everything. Broad Concept of Personal Data and Future of Eu Data Protection Law," *Law, Innovation and Technology* 10, no. 1 (2018): 53.

or be completely machine-led (unsupervised learning). In *supervised learning* the algorithm is fed data that is labelled by humans to belong a particular class or category and given positive or negative feedback according to whether it recognized the data to go under the correct category or not. After this training phase, the algorithm can be deployed to operate in the real world with unlabelled data to perform the task it is trained for.[99] In *unsupervised learning*, the algorithm is asked to discover correlations and patterns by itself in a large unlabelled training dataset. Here the feedback is given on the output for the algorithm to correct itself to fit the expectations, however how it produced the output can be hard for humans to understand which is known as the black box problem.[100] Another form of learning is *reinforcement learning,* here the algorithm is given an aim to achieve while interacting with an environment and is not given feedback immediately, the human developers act more like a critic this time than a teacher as Alpaydın describes it. This way the algorithm looks back in its actions/decisions to calculate which of them brought it closest to achieve the aim as early as possible and this way it optimizes. A program learning to play chess or a robot trying the way out of a maze can be examples to where reinforcement learning would be used. The feedback is whether the program won the game or broke out of the maze and it comes at the end of a sequence of actions and decisions.[101]

The current-state-of-the-art in machine learning is deep learning which uses

[99] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 34-35.

[100] Ibid., 35.

[101] Alpaydın, *Machine Learning, Revised and Updated Edition*, 161-62.

neural networks as opposed to Bayesian statistics[102] and feature engineering[103] used in classical ML.[104] Deep learning algorithms imitate the way the human brain operates. They do not require to be trained on selected dataset, nor are they given feedback by developers on their output.[105]

Use of cookies and other tracking technologies generate voluminous user data, enabling devices such as smart phones also collect and share data and data coming from this internet of things (IoT)[106] also add up on the Big Data that is available to be processed through different techniques briefly introduced above. Different algorithms may interact with each other, the output of one algorithm can become

---

[102] As explained by Alpaydın, sometimes the prior knowledge or beliefs of the data scientists may suggest possible values of parameters such as in the case of flipping a coin, one would expect the probability of heads to be close to ½. So here, the data scientists combine their prior knowledge and/or belief with the data they have to calculate such parameters or to get rid of parameters that are making the model too complex for its task. See. Ibid., 101.

[103] Also known as "feature extraction" or "feature processing" or "feature construction" or "feature transformation" as mentioned in David Lehr and Paul Ohm, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning " *U.C. Davis Law Review* 51, no. 2 (2017): 700. In this regard, feature engineering is the creation of inputs, often by obtaining some information from unstructured data or by transforming existing input variables (ibid.) that are important for predicting the specific target and addressing the specific question(s) the data scientists have at hand to develop a model. In order to do so, data scientists need to apply domain knowledge to decide how to use the information they have in a way to create new input variables that can predict the target better and improve the performance of the model (See. <https://towardsdatascience.com/what-is-feature-engineering-bfd25b2b26b2> Accessed on 17 September 2022).

[104] Hadi Asghari et al., "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making.," *Alexander von Humboldt Institute for Internet and Society* (2021).

[105] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 36.

[106] See Ayça Atabey, *Is Google at Odds with the Gdpr? Evaluation of Google's Personal Data Collection on Mobile Operating Systems in Light of the Principles of Purpose Limitation, Data Minimisation, and Accountability* (On İki Levha, 2020).

the input of the other.[107] The combination of various models using different techniques can be utilized for profiling. In this case, while models using clustering and association techniques would create profiles and assign individuals to them, regression techniques can be used to predict future behaviour[108] or condition (financial, health etc.) of these individuals. The more complex the autonomous systems become, we are more tended to call them artificial intelligence (AI) as they do not seem close to human thinking at all. However, this resemblance to intelligence usually is reflected in a limited area and understood narrow in the sense of intelligence.[109]

One can derive a definition for ADM also from GDPR Article 22. In this regard, it is a decision based on automated processing including profiling that produces significant effects for the subject of this decision. The GDPR definition is narrower than this definition we adopt as it applies to decisions "solely" based on automated processing and seeks for "legal or similarly significant effects." Because the use of decision making algorithms may be effective on important decisions that might not solely based on ADM and even though "discrimination" is acknowledged as a legal or similarly significant effect,[110] because the legal discussion based on European data protection law can be limiting for the purposes of this study, these two elements are left out from the definition.

On the other hand, 'artificial intelligence system' (AI system) is defined in the Proposal for a Regulation of the European Parliament And Of The Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) as "software that is developed with one or more of the techniques and approaches

---

[107] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 38.

[108] Ibid., 34.

[109] Ibid., 39.

[110] GDPR Recital 75

listed in Annex I[111] and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with."

This term encompasses algorithmic systems that we will tackle in this study. However, we agree with the AlgorithmWatch report stating that AI as a term brings to mine responsibility by being associated to closeness to human intelligence,[112] which can be misleading. Thus, we rather use the term algorithmic/automated decision making (ADM) interchangeably.

1.2.2 Algorithmic Bias and Discrimination

ADM systems are socio-technical systems, meaning that not only the historical and societal context in which they are developed shape them, but also in turn they shape societies they operate in and affect specific groups and individuals therein.[113]

First of all, there is a distinction in the terminology between algorithmic bias and algorithmic discrimination. Algorithmic bias is not always concerned with fairness and generally refers to any systemic error which maybe statistical, cognitive, societal, structural, or institutional, that affects the outcome of an algorithmic system.[114] When bias is used in relation to fairness, then it refers to a

---

[111] The list goes as follows: "(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; (b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems; (c) Statistical approaches, Bayesian estimation, search and optimization methods."

[112] Wulf, *Automated Decision-Making Systems and Discrimination: Understanding Causes, Recognizing Cases, Supporting Those Affected* 5.

[113] Ibid.

[114] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 47.

systemic error which advantages privileged groups and disadvantages unprivileged groups[115] whom we refer as WMG. On the other hand, algorithmic discrimination, which is our main concern, refers to ADM systems resulting in discriminatory outcomes as defined in a legal system. In this regard, algorithmic discrimination is unjustified unfavourable treatment of protected groups or on the basis of protected grounds as specified in law or the disadvantage experienced by them and/or on these grounds caused by algorithmic systems.[116] In this regard, algorithmic fairness covers a broader area than algorithmic discrimination. However, taking fairness into consideration in algorithmic systems that goes beyond the notions of equality and non-discrimination as prescribed by law is left to the good will and interest of the developers or deployers of algorithmic systems, whereas algorithmic discrimination can be enforceable by law despite the difficulties that we will discuss throughout this study.

WMG have been historically and structurally discriminated against in access to employment as it will be detailed later in Section 1.3.1.

If discrimination is embedded in the human relations, why are there particular focus on algorithms and their developers with respect to concerns of discrimination? If discrimination of WMG in the society is this deep rooted, how can we expect computer scientists to fix it? To be able to answer these commonly heard questions, we need to have a closer look at (1) why and how machine learned algorithms discriminate, (2) how it is different to discrimination by humans and (3) how we can mitigate the harms of machine bias?

Some authors favour algorithmic discrimination over discrimination by humans because in certain types of algorithmic systems discrimination is identifiable,

[115] Ibid.

[116] Ibid.

fixable and more transparent compared to discriminatory human judgements.[117] Other scholars, on the contrary, argue that algorithmic discrimination is much worse than of humans as we do not know how to detect it and it doesn't come intuitive.[118]

One of the major differences of machine learning is that correlations that humans could not be able to find out would appear in the models and bias their predictions in the detriment of certain traditionally disadvantaged groups or completely unnoticeable groups to the human eye. For those who were not shown a job advertisement or were rejected among thousands of applicants at an earlier stage, it is particularly hard to be aware of discrimination based on complex patterns and correlations, which would be more subtle than human discrimination.[119] For example, a model of a screening company identified having the name Jared and playing high school lacrosse as strong signals of success.[120]

While humans usually have an intuition for discrimination from other humans, it may be really hard to become aware of it when machines discriminate. On the other hand, because the machines are more likely to come up with some mathematical/statistical reasoning to their discrimination, both the employers and the potential and current employees may be more likely to trust their judgements

---

[117] Goodman and Flaxman, "European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"."

[118] Sandra Wachter, Brent Mittelstadt, and Chris Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai " (2020); Wachter, Mittelstadt, and Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation."

[119] Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai ".

[120] Miranda Bogen and Aaron Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias," (2018).

to be fair, a concept known as automation bias.[121] Another difference is that machines are always more conservative than humans as they cannot imagine a different future from and are bound with the data they were fed with, which is always going to be data that reflects a past and constrained reality. Most of the fairness-aware algorithms cannot handle complex sensitive attributes and intersectional discrimination. Moreover, as a result of limited opportunities for error correction and possibilities of reinforcing feedback effects, many algorithms bear the risk of worsening bias and inequality. Finally, machines are capable of carrying the bias onto and discriminating on a larger scale than humans can.

Here below, we will have a brief but closer look into the stages of an algorithmic decision-making process to identify in which stages and how bias and discriminatory outcomes may creep in the system.

1.2.2.1 Project Design

This stage involves identifying the need for an ADM process, defining a particular objective for the algorithm to be used, deciding which type of algorithm can deliver the best outcomes for the goal, how the output of the algorithm will be used, deciding whether to develop the algorithm in-house or to procure or a combination of both, building a team of data scientists and domain experts accordingly, assessing the data needed to build, train and test the algorithm, whether such dataset is available in-house or publicly, whether it will be necessary to collect or generate data or to buy a dataset.[122]

Biased and/or non-diverse groups of people making these decisions can be one of the reasons that leads to discriminatory results at this stage. But also defining a

[121] Ibid.

[122] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 38.

business problem in a way that a ML algorithm can predict or estimate something that is measurable,[123] so that this problem can be solved is a very complex task. If supervised learning is chosen to be used, then an outcome variable has to be defined first. If classification techniques are being used, this can be as simple as a binary yes/no, spam/not spam, or if multi-label classification is needed for example for an autonomously driving car to detect the traffic light it may look like Green / Yellow / Red / Black (if the traffic light is out of service).[124] In regression techniques, where a continuous outcome such as second-hand car prices, credit score[125] or sales of a company for the next month[126] is being predicted, this would be a numeric value.[127] In any case, first, the abstract goal of the project should be translated into what to predict. Second, this prediction goal has to be specified as an outcome variable. In a human resources scenario as exemplified by Lehr and Ohm, HR department of a company may want to predict good employees to recruit or to promote. How can goodness be measured? What kind of outcome variable would allow the HR team to measure it? As it is so abstract and subjective, maybe goodness cannot be measured at all, which would call for choosing a proxy for it. What would it be? Tenure, overtime working hours, performance scores, sale numbers? All of these would reflect qualities that structurally-burdened good employees might not be able to show. After all, beyond being hard to measure, a good employee is a construct that has been shaped in the form of an able-bodied cis-gender male with no domestic burden. Here, firstly, subject matter knowledge

---

[123] Lehr and Ohm, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning " 672.

[124] This example is taken from Kareem M. Metwaly et al., "Car - Cityscapes Attributes Recognition a Multi-Category Attributes Dataset for Autonomous Vehicles," *ArXiv* abs/2111.08243 (2021): 14.

[125] Alpaydın, *Machine Learning, Revised and Updated Edition*. According to Alpaydın, credit score can be defined as a classification problem as well, i.e., high risk/low risk.

[126] <https://www.seldon.io/machine-learning-regression-explained> Accessed on 18 September 2022.

[127] Lehr and Ohm, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning ".

may come into play suggesting that tenure is a better indicator than sale numbers. Secondly, the outcome variable an algorithm can work with and the output that it may produce varies according to the type of algorithm. If the output of an algorithm is going to become the outcome variable of another model, or the decision-makers have a reason to stick with a type of algorithm, they may specify their outcome variable accordingly. Thirdly, they may choose the outcome variable based on difficulty and the cost of measuring, going for easier and more affordable options.[128] All of these decisions may have an effect once the model is developed and deployed on its accuracy, fairness and explainability, as well as leading to discrimination.

One example that the bias at this stage with the downstream effect of discrimination is the child benefits fraud algorithm. This occurred in Netherlands, where an algorithmic system designed to detect incorrect applications and fraud in child benefits was deployed in 2013 by tax authorities. According to the report of Amnesty International[129] the fact that the citizenship of parents was chosen as a data field and non-Dutchs were automatically considered higher risk to commit fraud based on subject matter knowledge that is informed by stereotypes, caused many low-income families disproportionately from ethnic minorities falsely being accused, being subjected to harsh investigations and had their child benefits suspended. Even before any data was collected or the model was developed, the design choice of adding this data point led the self-learning algorithm to assign non-Dutch status a high fraud risk score.

[128] Ibid., 675.

[129] Amnesty International, Xenophobic Machines: Discrimination through Unregulated Use of Algorithms in the Dutch Childcare Benefits Scandal, (London: Amnesty International, 2021), https://www.amnesty.org/en/documents/eur35/4686/2021/en/.

1.2.2.2 Development

In this stage, developers write the computer codes that will build the algorithm. If a rule-based model is being used the ADM process is broken down to different steps and translated into code as rules and decisions, variables that the system should use are chosen. If a self-learning model will be developed, learning technique and technologies for data analysis are chosen, target variables are defined and translated into code, training and test data are prepared. Then the algorithm is trained, given feedback, fine-tuned, validated and certified if possible, so that it is ready to be deployed in the real world.[130]

1.2.2.2.1 Data

Algorithms run on data. They learn from data, they process input data, they produce output data and they are as good as their data. While an algorithm is being developed, collecting enough data[131] is the first goal and this data would normally need cleaning.[132] Then, the data set at hand is partitioned into two sets of data called training data and test data. Even how this portioning is done, can have effect how it will perform for particular individuals and groups.[133] Basically, a model would learn with the training data by being exposed to right answers and then the resulting predictive model (a.k.a. the algorithm) gets applied to new, unclassified data called 'test data.'[134] When the developer is convinced it is ready, the algorithm is deployed to function in making real life decisions informed from the training.

---

[130] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 39.

[131] Lehr and Ohm, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning " 678.

[132] Ibid., 681.

[133] Ibid., 684-88.

[134] Ibid., 684.

The notorious “garbage in garbage out” concept comes from the fact that the biased data used in training or as input would lead the model to discriminate against particular groups.

First of all, in supervised learning, human analysts label the training data, so that it can correctly and accurately calculate the desired output. These humans may have implicit bias towards a group[135] or lack of knowledge and/or first-hand experience due to their privileged social status about what the societal reason behind the predictive power of an input to the desired output is. In both cases, the final model may show bias towards disadvantaged groups in the society such as WMG.

WMG are more likely to be underrepresented if not completely excluded, overrepresented due to oversurveillance,[136] or have more erroneous information in datasets, all of which may result in less accuracy for these groups in a systematic way which would lead to unfair decisions and discrimination harms.[137] For example, being excluded from the public life women’s lives had not been recorded; which constitutes one of the reasons for women’s needs not been reflected in the making of general rules and organization of public and private institutions. Today, we see this phenomenon when being less datafied[138] causes algorithmic discrimination to particular groups. Here, non-representative data collection hampers the generalizability of the model, meaning that the training/test

---

[135] Philipp Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," *Common Market Law Review* 55 (2018): 6.

[136] Sonia K. Katyal, "Private Accountability in an Age of Artificial Intelligence," *The Cambridge Handbook of the Law of Algorithms* (2020).

[137] Solon Barocas and Andrew D. Selbst, "Big Data’s Disparate Impact," *California Law Review* 104 (2016).

[138] See D’Ignazio and Klein, *Data Feminism*; Criado-Perez, *Invisible Women : Data Bias in a World Designed for Men*

data do not represent the real population,[139] also known as sampling bias.[140] Ohm and Lehr suggest that the subject matter knowledge may help the analyst to include the population that is not represented in the data through non-random sampling.[141] Being attuned to such problems though would require having social scientists, lawyers and a team of analysts from diverse backgrounds.

Another very well-known problem is historical bias where past discriminatory decisions or prevalent biases deeply rooted in the society inform decisions of future, for example about whom to employ, predicting people similar to those who were deemed successful in the past to be good employees.[142] In this case, if the employer is happy with the hiring decisions they made based on the algorithm's prediction without knowing or caring about which talents they may have missed, these decisions continue to inform the system, creating a feedback loop that result in self-fulfilling prophecies.[143] Another example is, deeply structured and historical discrimination expressed in language itself as implicit human biases creeping in ML systems through which stereotypes such as women belong to domestic work or not good at technical tasks like coding or maths are reproduced.[144]

---

[139] Buolamwini and Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification."

[140] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

[141] Lehr and Ohm, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning " 704.

[142] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

[143] Ibid., 43; Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

[144] Tolga Bolukbasi et al., "Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings" (paper presented at the NIPS, 2016); Aylin Caliskan, Joanna J. Bryson, and Arvind Narayanan, "Semantics Derived Automatically from Language Corpora Contain Human-Like Biases," *Science* 356, no. 6334 (2017).

In supervised learning techniques[145] the target variable that the model is trained to predict is known and labelled by programmers (e.g. knowing what a spam mail is and labelling it as such so that the machine can learn what is spam). In other words, because the model will accept the labelled training data as ground truth, bias that the training data may contain will cause biased outcomes.[146] Hacker defines ground truth as "the best available approximation of reality, expressed in data."[147]

On the other hand, unsupervised machine learning algorithms seek for correlations/patterns in the training data to predict the target variable. In this case, unequal ground truth, meaning that some groups in the society correlate with some capacities or risks more than other groups in the dataset, may cause proxy discrimination (a.k.a. statistical discrimination)[148] many times even without any intention in this direction. The correlation captured by the model may be statistically true that for example young women's careers are more likely to be interrupted due to pregnancy and childcare compared with young man. In this case there is even causality between uninterrupted tenure and gender. A similar correlation would be the distance of an employee's home to the workplace to predict premature resignation,[149] in which case a model trying to weed out such cases may end up discriminating against candidates from a certain race or ethnic minority as zip code is known to correlate with such group memberships. This is also the result of a very well-known example of traditional proxy discrimination called redlining. In the U.S., financial institutions used to refuse extending

[145] According to Lehr and Ohm, most of the algorithms that are used in hiring and other areas of HR are developed with supervised learning techniques. See Lehr and Ohm, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning ".

[146] Pauline Kim, "Data-Driven Discrimination at Work," *William & Mary Law Review* 48 (2017).

[147] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 6.

[148] Ibid.

[149] Bogen and Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias."

mortgage to African Americans due to explicit and implicit racial bias. After such practice was banned by law, they started to draw a redline on the maps to exclude areas with high investment risk, which also happened to be the areas where predominantly African American communities lived.[150] Here, zip code becomes a proxy for race, in other words replaces race as a direct parameter. While proxy discrimination by humans is in most cases intentional as in the redlining example, algorithmic systems are likely to cause unintentional proxy discrimination. Prince and Schwarcz give another example that a job requires over 180 cm height to be performed safely, because the employer do not have access to the information on the height of applicants, they weed out traditionally female names as height is highly correlated with sex. Such proxy discrimination may be carried out by a model similarly and more sophisticatedly by finding out sex as a proxy for high performance at the job that requires tall people and using Netflix viewing data to predict the sex of candidates, turning Netflix behaviour into direct proxy for sex.[151] However, a model produce a target variable in the form of a score (a number indicating to job performance, financial capacity or likelihood to click) through processing of highly granular data which do not include any protected grounds against discrimination such as sex, race, age, disability, however, the score may still be correlating to some protected grounds leading to indirect proxy discrimination.[152]

Correlations and proxy discrimination are understood as one of the biggest challenges to address algorithmic discrimination.[153] For example a model that is trained on the past job performance data of a company may correlate being woman with low job performance, not because there is causality between gender and job

---

[150] Prince and Schwarcz, "Proxy Discrimination in the Age of Artificial Intelligence and Big Data."

[151] Ibid., 1280.

[152] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 7.

[153] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 44.

performance but because women have been evaluated worse than men for the same performance due to implicit and explicit bias of past supervisors and the company culture.[154] Findings of ML algorithms are correlations that do not necessarily suggest a causal link between the input variables and the output, however they would serve for confirming and cementing already existing biases and stereotypes as seen in the previous example, as they do not reason their findings. This problem aggravates as these models are also very good at correlating with proxies for the data that are not available, one of the reasons for which can be legal rules against discrimination.[155]

Prince and Schwarcz offer a taxonomy of proxy discrimination consisting of two types of direct proxy discrimination, namely causal and opaque; and indirect proxy discrimination. In causal proxy discrimination, there is one variable causally linked to the target variable and in the case that it is missing from the dataset, the algorithm would find any available facially neutral classifier that partially proxy this causal variable. For example when a gene is solely responsible for developing or not developing a genetic disease and if the genetic data is legally not available for an insurance company's use, the algorithm may proxy the missing gene data with membership to a Facebook group for people with the disease. Thus, a facially neutral classifier (membership in a Facebook group) proxies for the suspect classifier (genetic data) which causally predicts the target variable (highly likely/highly unlikely to develop the disease). In the opaque proxy discrimination, causative variable may either not be fully understood to be quantified or it is hard to quantify. For example, a study showed that young women drive more safely than young men, which forms a causal link between gender and insurance claims (target variable). However, gender (suspect classifier) cannot be directly used in

---

[154] Kim, "Data-Driven Discrimination at Work."

[155] Prince and Schwarcz, "Proxy Discrimination in the Age of Artificial Intelligence and Big Data."; Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

differentiating insurance premiums, so that the model uses online shopping behaviour (facially neutral classifier) as proxy to gender which is a causal proxy for care level (unquantifiable or unavailable variable). Indirect discrimination can be better understood with the following example. This time, driving faster (quantifiable or available variable) is causally linked to being involved in more accidents therefore more insurance claims (target variable). If it is known that men (suspect classifier) drive faster and men are more likely to drive red sports cars (facially neutral classifier), the algorithm may issue higher premiums for red sports car drivers which would affect male customers worse.[156] Here, driving a red sport car gains its predictive power from being correlated to being man which is correlated to driving faster. Being men is not directly predictive, instead it provides one of several potential ways of assessing the likelihood of some true causative factor (driving fast) that is both quantifiable and potentially available but is not involved in the dataset on which the model was trained. This is similar to the example above, where Netflix watching behaviour correlated with job performance because it was a proxy for gender which was an imperfect predictor for height as gender is a proxy for sex and sex correlates highly with height which was a causal predictor of good job performance but was not available to the model.[157] This phenomena that the target variable considerably correlating with membership in a protected group, thus such membership can be taken as indirect proxy is also called redundant encoding[158] which is one of the reasons algorithmic models reproduce structural discrimination.[159]

---

[156] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 7.

[157] Prince and Schwarcz, "Proxy Discrimination in the Age of Artificial Intelligence and Big Data," 1280.

[158] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 7.

[159] Raphaële Xenidis and Linda Senden, "Eu Non-Discrimination Law in the Era of Artificial Intelligence: Mapping the Challenges of Algorithmic Discrimination," (2020), 20.

Prince and Schwarcz suggest that (i) not including data for legal and/or ethical reasons due to discrimination and fairness concerns is not a good remedy as it will cause proxy discrimination, (ii) the more causal data which directly predict the target variable are available, the less indirect proxy discrimination will occur, (iii) because there can be many ways of predicting a target variable all these three forms of proxy discrimination are expected to happen within the same ADM process.[160]

From the time target variable is defined to the functioning of the model after deployment, data plays a crucial role in its fairness. On this journey of data from the very beginning we placed data related issues with discriminatory effects in the development phase as it is the most defining for such outcome. In sum, if the target variable is disadvantageously defined for some groups, even if the training data and the test data is not biased, they would produce biased outcomes. Target variable might be carefully defined, however pre-existing human biases[161] in the data may bias the rules and the outcome. The dataset used to train and test the algorithm may be not representative for some groups, some input variables maybe more predictive for one group that the others,[162] some predictive data being missing from the dataset and reliance on correlations may lead to proxy discrimination[163] which may mask discriminatory intents,[164] make it nonintuitive[165]

[160] Prince and Schwarcz, "Proxy Discrimination in the Age of Artificial Intelligence and Big Data."

[161] See Katyal, "Private Accountability in an Age of Artificial Intelligence." Where a variety of human biases are explained.

[162] Lehr and Ohm, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning " 703.

[163] Prince and Schwarcz, "Proxy Discrimination in the Age of Artificial Intelligence and Big Data."

[164] Barocas and Selbst, "Big Data's Disparate Impact," 692.

[165] Selbst and Barocas, "The Intuitive Appeal of Explainable Machines," 1091.

for humans (developers, users, those who are discriminated against, courts etc.)[166] to grasp unintentional discriminatory effects, or create a veil of objectivity for the discriminatory effect through the predictive accuracy of the model.[167] All of these are capable of reproducing structural inequality patterns, in a greater scale and speed[168] than traditional human biases resulting in discriminatory decisions.[169]

1.2.2.2.2 Model

At this stage the data scientist first selects suitable algorithms. They may choose to implement a pre-programmed algorithm or program their own by taking into account the outcome variable, whether it can translate the acceptability of false negatives and positives for the stakeholders into the model, explainability, overfitting, opportunities for tuning, resource limitations of the environment they will be deployed such as processing power, time, and memory space to run. The analyst can choose which algorithm should further to testing and be deployed based on the performance of the models taking into account the accuracy of the model during the training phase.[170]

Lehr and Ohm explain that training phase starts with choosing tuning parameters, which would allow changing the internal operation of the model. Objective function, which is the mathematically expressed version of algorithm's

---

[166] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*; Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai ".

[167] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 23.

[168] It is called "the scale and speed challenge" in Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

[169] Ibid.

[170] Lehr and Ohm, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning " 688-95.

goal, is one of the parameters generally available for tuning, that is also suggested by Lehr and Ohm to be useful in mitigating discriminatory outcomes. Another relevant tuning parameter is bias-variance trade-off which allows for fixing overfitting and underfitting. Underfitting occurs when the error rate of the model both in training data and test data is high, overfitting happens when the error rate is low in training data and high in test data. In both cases the model will not perform well with the real world data, in other words the model cannot generalize.[171] Because data is often noisier for minority groups than for others,[172] an overfitting algorithm could generate less accurate predictive rules for minority groups than for others. The analysts should take this into consideration at model selection.[173]

Kim[174] explains that the more variables one has, the more data is needed in order to keep a decent level of accuracy in the model, which could be costly. For this reason, one of the important parts of machine learning processes is determining which variables to include or exclude, also referred to as feature selection.[175] Variables which seem to not add value in solving the problem and are therefore excluded from the data could more likely be related to those who are farther than the mainstream. The Gender Shades study[176] showed how facial recognition algorithms were significantly less accurate for darker skin women than lighter skin

---

[171] Ibid., 697.

[172] See D'Ignazio and Klein, *Data Feminism*, 33-39; Joost Kappelhof, "Survey Research and the Quality of Survey Data among Ethnic Minorities," (2017); Criado-Perez, *Invisible Women : Data Bias in a World Designed for Men*

[173] Lehr and Ohm, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning " 704.

[174] Kim, "Data-Driven Discrimination at Work."

[175] See Lehr and Ohm, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning " 700.

[176] Buolamwini and Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification."

women and most accurate for lighter skinned man. Although, making decisions based on protected attributes could be in many jurisdictions illegal, it is also known that excluding a protected attribute may cause more harm to the protected group. An example is "omitted variable bias" that race is omitted for some good reasons and the model finds out a negative relation between prior military service and job performance based on the general data, while actually for African Americans there has been a positive relation. This model might end up disadvantaging African Americans more than it would have if race was not excluded.[177]

Data on which the analysts work may not exactly reflect the reality of the population about whom the algorithm will make decisions. Such failure to generalize can arise from non-representative training data as explained above, as well as giving in to overfitting (selecting a bagging algorithm may help overcome overfitting), favouring low bias in bias-variance trade-off (if the training/test data and the real world is different and analyst leaned towards less bias rather than variance the model would fail to generalize, so that fixing this again requires the awareness of the analyst about such unfair distribution of data), improper assessing of accuracy during training. Lehr and Ohm suggest while problems in data leads to discriminatory results, focusing on this stage of algorithmic processes may help with mitigating many risks that would lead to discrimination. Human in the loop also would be better placed in this stage to intervene. According to these scholars, waiting for the output to be produced/decision to be made leaves the human in the loop rubber stamp the conclusion of the algorithm.

Finally, everything could be done right during the training, a measurable objective criterion could be chosen as the target variable, and the protected attributes could be omitted, but the algorithm may still discriminate by finding a neutral factor which correlates highly both with the target variable and a protected class (proxy discrimination). One machine learning model, for instance, found out

[177] Kim, "Data-Driven Discrimination at Work."

that the distance an employee lives from the workplace is a very strong indicator for the length of employee tenure. However, in the US context zip codes are known to act as proxy for race.[178]

1.2.2.3 Deployment

When an algorithm is deployed to carry out the task to achieve the objective set in the project design stage, it will be fed or encounter with input from the real world and generate output accordingly. At this phase of ML, correlations/patterns discovered during the training phase are turned into rules on which the model predicts outputs. This output may take the form of a decision which can be generated automatically without any human involvement or take the form of a recommendation or prediction to assist human decision-makers.

This stage also involves monitoring of the ADM process for any unwanted outcomes due to rule-based or supervised-learning algorithms being prone to become outdated because the training data does not match the real world anymore or self-learning and deep-learning algorithms may generate unexpected outcomes as they continuously learn from the new input.[179] One extreme example to that is Microsoft's chatbot Tay ending up reflecting the worst of humanity within 16 hours it engaged in conversations with Twitter users until it was taken down by Microsoft.[180] Monitoring at this stage involves making necessary adaptations to fix the algorithm and to be able turn the kill switch on.

In this stage data continues to play a crucial role. Incorrect information may

---

[178] Ibid.

[179] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 40.

[180] <https://techcrunch.com/2016/03/24/microsoft-silences-its-new-a-i-bot-tay-after-twitter-users-teach-it-racism/> Accessed on 19 September 2022.

also result in incorrect and discriminatory decisions after the model is deployed even if the model is functioning correctly.[181] This incorrect information may also be the output of another algorithm, for example an inference about a person may not be true and lead to them being assigned under a profile which would then be used for another algorithm to deliver job ads. The wrong assignment to a profile may cause a person to be excluded from otherwise fitting opportunities or delivered irrelevant or bad offers.

Although, the source of opaqueness and difficulties in explaining is not this stage of ML, this is the stage the algorithms start making decisions about individuals or groups of people that affect their lives. However, many of these individuals will not really understand how these decisions about them are being made. Thus, at this stage opting for fairness would also mean developing processes to make sure those who are affected are properly informed. We will return to the transparency and explainability challenge[182] later.

In cases when the algorithm is not automatically making a decision but supports human decision-makers, automation bias and anchoring may debase and cripple having a human-in-the-loop.[183]

In sum, different stages of an ADM process bear different risks dependent on the context. We could identify project design, data generation, model development, and deployment stages, which operate in cycles and the output result often become the input again in real-time self-learning algorithms. In the project design phase, biased and/or non-diverse teams, as well as how these teams define the problem, in other words decide on the target variable plays a role in bias

[181] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

[182] Ibid., 45.

[183] Ibid., 42.

creeping in. In the data generation stage of development phase, un-representative datasets and historical bias may cause reproducing structural inequality patterns and proxy discrimination. In the model design stage, target variables/outcome definition and feature selection is where data scientists are involved and their personal opinions or limitations may cause bias.

In this regard, at the deployment phase, correlations should not be interpreted in a manner they indicate to causalities and automation bias should be avoided. Statistics driven implementation may neglect special situations and individual issues. Then these models may end up producing self-reinforcing results which are also called self-fulfilling prophecies especially when the system repurposes its own outcomes and uses these inferred data as input for further decisions. Thus continuous monitoring of the ADM processes after deployment is crucial to intervene before such tendencies turn into broad societal problems.

1.2.3 Surveillance and Profiling

Is data the new oil? An analogy which has been at use especially in the business circles[184] allegedly since it was coined by Clive Humby in 2006.[185] Referring to data (including and especially personal data) as a natural resource of great value to be exploited for profit is reminiscent of colonialist strategy of calling "the new world" no man's land.[186] Moreover, referring to data as raw material veils the work that has to be put to generate it and to make sense of it which makes it value-laden

[184] See "The world's most valuable resource is no longer oil, but data" <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data> Accessed 12 May 2022 and "Personal Data: The Emergence of a New Asset Class" <https://www.weforum.org/reports/personal-data-emergence-new-asset-class> Accessed 12 May 2022.

[185] <https://ana.blogs.com/maestros/2006/11/data_is_the_new.html> Accessed 10 June 2022.

[186] Cohen, Between Truth and Power: The Legal Constructions of Informational Capitalism.

and is also the reason why raw data is an oxymoron.[187]

Big Data is at the core of the new economic growth[188] and at the verge of expanding even more as smart city, smart home and IoT Technologies are expected to translate many offline human activities into mass amounts of data as well. Big Data is defined usually with four Vs. Volume relates to the massive amounts of data collected, the Variety relates to the various sources from which the data are collected, the Velocity is about how analysis of data can unwrap in real time and the Veracity relates to the quality and accuracy of data which could be achieved through the analytical process.[189]

One important aspect of Big Data is its relationality to other data, in other words its value comes from making connections with other data and deriving patterns, that it is fundamentally networked.[190] Big Data is a kind of information asset that is hard to exploit with traditional methods.

Artificial Intelligence,[191] in this regard, is seen as a key for unlocking the value of these assets. One of the technical mechanisms that is used for this purpose is

---

[187] Gitelman, "Raw Data" Is an Oxymoron.

[188] https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/646117/EPRS_BRI%282020%29646117_EN.pdf

[189] Tal Z. Zarsky, "Incompatible: The Gdpr in the Age of Big Data," Seton Hall L. Rev. 47 (2016): 998-99; Information Commissionar's Office (ICO), "Big Data, Artificial Intelligence, Machine Learning and Data Protection," 20170301 Version: 2.0 (2017), https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf. Rob Kitchin and Gavin McArdle, "What Makes Big Data, Big Data? Exploring the Ontological Characteristics of 26 Datasets," *Big Data & Society* 3, no. 1 (2016).

[190] danah boyd and Kate Crawford, "Six Provocations for Big Data," Computer (Long. Beach. Calif) 123 (2011).

[191] Here we purposefully use the hyped term.

machine learning.[192] In other words, Analysing Big Data by making use of AI techniques such as machine learning is called “big data analytics” and it enables enhanced insight and decision-making. Another commonly used term is ‘algorithmic decision-making’ defined by Yeung as “the use of algorithmically generated knowledge systems to execute or inform decisions”.[193] Any method of digital data processing follows the IPO (input, processing, output) model – data enter a system as input, are processed, and then leave it as output.

Instead of the new oil, Pendergrast calls big data the next big cheap with a note that it is cheap to the capitalist. Otherwise, there are concerning social costs of the cheap labour needed from the manufacturing of the hardware to labelling images for training the models, and the environmental costs of the ever rising need of energy to keep the databases and high processing power of machine learning models running.[194] On the other hand, branding personal data as something free like an abundant natural resource is a capitalist tactic similar to those used in the early capitalist days.[195] It starts with creating a human/nature binary, that anything separated or excluded from human becomes exploitable which have expanded to humans by colonizing and enslaving indigenous people and women.[196] Like America was not *terra nullius* when the colonists arrived but also the native American people did not have an understanding of ownership similar to the

---

[192] Information Commissionar’s Office (ICO), "Big Data, Artificial Intelligence, Machine Learning and Data Protection".

[193] Karen Yeung, "Algorithmic Regulation: A Critical Interrogation," Regulation & Governance 12, no. 4 (2018): 506.

[194] Kelly Pendergrast to Real Life, 25 November 2019, 2019, https://reallifemag.com/the-next-big-cheap/; Emma Strubell, Ananya Ganesh, and Andrew McCallum, "Energy and Policy Considerations for Deep Learning in Nlp," *ArXiv* abs/1906.02243 (2019); Karen Hao, "We Read the Paper That Forced Timnit Gebru out of Google. Here’s What It Says.," MIT Technology Review, https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/.

[195] Cohen, Between Truth and Power: The Legal Constructions of Informational Capitalism.

[196] Pendergrast The Next Big Cheap.

colonists,[197] private and social lives of internet and connected device users used to belong to them but were not appropriated as commodity. Ownership on personal data, whether it can be owned at all, if yes to whom does it belong (the identified or identifiable person to whom the data relates to or the entity who collected and made sense of it; if the latter is possible then when does this ownership begin), and what are the legal and societal consequences are all very important questions. There is an ongoing scholarly debate on it;[198] however, this study attempts neither to dig deeply into this discussion nor to seek answers to any of the questions raised by it. The reason for it is that we are sceptical that putting a price on it would normalize the commodification of personal data, which would allow immediate needs of the data subjects to decide on the final value of data protection overriding its intermediate value as an instrument which ensures their autonomy. Autonomous individuals constitute the very basis of the liberal democratic system to the extent the whole system would collapse in the absence of autonomic

[197] Cohen, Between Truth and Power: The Legal Constructions of Informational Capitalism.

[198] Francesco Banterle, "Data Ownership in the Data Economy: A European Dilemma," (2018); Tommaso Fia, "An Alternative to Data Ownership: Managing Access to Non-Personal Data through the Commons," Global Jurist 21, no. 1 (2021); Patrik Hummel, Matthias Braun, and Peter Dabrock, "Own Data? Ethical Reflections on Data Ownership," Philosophy & Technology 34, no. 3 (2021); Václav Janeček, "Ownership of Personal Data in the Internet of Things," Computer Law & Security Review 34, no. 5 (2018); Gianclaudio Malgieri, "'Ownership' of Customer (Big) Data in the European Union: Quasi-Property as Comparative Solution?" (2016); Barbara Prainsack, "Logged Out: Ownership, Exclusion and Public Value in the Digital Data and Information Commons," Big Data & Society 6, no. 1 (2019); Nadezhda Purtova, "Property in Personal Data: A European Perspective on Instrumentalist Theory of Propertization," Research Policy - RES POLICY (2010); "Do Property Rights in Personal Data Make Sense after the Big Data Turn?: Individual Control and Transparency," ISN: Property Protection (Topic) (2017); Jeffrey Ritter and Anna Mayer, "Regulating Data as Property: A New Construct for Moving Forward," (2018); Malgieri, "'Ownership' of Customer (Big) Data in the European Union: Quasi-Property as Comparative Solution?."

capabilities.[199] Moreover, making the choice of disclosing some of their personal data by some may disadvantage others who had not chosen to do so.[200]

As Zuboff calls it, surveillance capitalism was born to turn the human experience into surplus value.[201] Haggerty and Ericson addressed the emergence of a surveillant assemblage through the convergence of surveillance systems of state and non-state institutions with different purposes and targets.[202] Couldry call it annexation of life to capital under data colonialism.[203] Digital platforms as the pioneers of data colonialism, encourage their users to share more of their lives, thoughts, interests and to engage socially and politically with others online. In all areas of human production and movement of goods, data driven logistics capture and appropriate another previously unavailable data and unknown information derived from it. Finally, through wearable devices and special applications, both voluntary and involuntary self-tracking plays an important role in appropriation[204] of especially intimate human experiences at scale.[205] More to come on this as there

---

[199] The concept of autonomous subjects is also a deeply problematic issue that has been well discussed in the feminist literature Catriona Mackenzie and Natalie Stoljar, Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self, vol. 17 (Oxford University Press, 2000); Fredman, Women and the Law.. See also Steffen Böhm, Ana C. Dinerstein, and André Spicer, "(Im)Possibilities of Autonomy: Social Movements in and Beyond Capital, the State and Development," Social Movement Studies 9, no. 1 (2010)..

[200] Antoinette Rouvroy and Yves Poullet, "The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Inportance of Privacy for Democracy," in *Reinventing Data Protection?*, ed. Serge Gutwirth, et al. (New York, NY: Springer, 2009), 50.

[201] Zuboff, The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.

[202] Kevin D. Haggerty and Richard V. Ericson, "The Surveillant Assemblage," The British journal of sociology 51 4 (2000). According to the writers, Foulcault's Panopticon analogy is outdated.

[203] Nick Couldry and Ulises A. Mejias, "Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject," Television & New Media 20, no. 4 (2019): 338.

[204] Ibid.

[205] Anastasia Siapka and Elisabetta Biasin, "Bleeding Data: The Case of Fertility and Menstruation Tracking Apps," Internet Policy Review 10, no. 4 (2021).

are many projects working on translating more of the offline world into digital.[206] According to Kadiri, such datafication does not only perpetuates but also builds on racism.[207]

Datafication is closely related to surveillance that is experienced disproportionately by women and marginalized groups. Common to be heard as opposed to surveillance concerns, "you don't have to fear, if you've got nothing to hide" argument is otherizing in itself as it implies one would be in need of hiding only for objectively and universally accepted wrongdoings. In reality though, it is more complex than that for everyone as privacy's interplay with other rights and that surveillance may disrupt the exercise of these other rights and cause disintegration of democracy in general.[208] In addition to that, it is much harder for WMG to not fear as they are usually targeted for the very basics of their identities.[209] Women grow up and live their lives with the inherent understanding of the male gaze or the feeling of being watched in almost all public places they enter. This has been even more deeply established as black people have been deemed to be commodity and their deterrence was punished by law. Browne refers to lantern laws,[210] which required black slaves to carry a lantern when they are out

---

[206] See Natasha Dow Schüll, "Data for Life: Wearable Technology and the Design of Self-Care," *BioSocieties* 11, no. 3 (2016); Natali Helberger, "Profiling and Targeting Consumers in the Internet of Things – a New Challenge for Consumer Law" (2016).

[207] Aisha P.L. Kadiri, "Data and Afrofuturism: An Emancipated Subject?," *Internet Policy Review* 10, no. 4 (2021).

[208] Daniel J. Solove, "'I've Got Nothing to Hide' and Other Misunderstandings of Privacy," 44 (2007); Theilen et al., "Feminist Data Protection: An Introduction."

[209] Toby Beauchamp, "Artful Concealment and Strategic Visibility: Transgender Bodies and U.S. State Surveillance after 9/11," Surveillance and Society 6 (2009); Nicole Shephard, "Big Data and Sexual Surveillance" (2016); Maya Ganesh, Jeff Deutch, and Jennifer Schulte, Privacy, Visibility, Anonymity: Dilemmas in Tech Use by Marginalised Communities (2016); Browne, "Dark Matters : On the Surveillance of Blackness."; Shmyla Khan, 2017, https://www.privacyinternational.org/news-analysis/3376/surveillance-feminist-issue.

[210] Another example to how law has been complicit in the oppression of WMG.

in the dark as a racial surveillance practice, as an example to such laws. The angst of being watched online is reminiscent of these shared experiences of women and other marginalized groups and it is closely related to online harassment, and cause chilling effects which lead to WMG to self-censor their speech and adapt behaviour to conform with the norms imposed on them. The white gaze and/or male gaze objectify through racializing, sexualizing and/or heteronormalizing their target in an act of subordination by making them visible and forcing them to adapt their behaviour or else, may they face violence.[211] Moreover, oversurveillance is the reason why some populations[212] may be overrepresented in some datasets.

Moreover, with the availability of the technologies of today, it is no longer that only the suspect populations are under attack. The systematic data-based surveillance,[213] also known as dataveillance, differs from surveillance in the sense that it is not motivated to monitor for a specific purpose and involves continuous tracking all the meta data and data related to every population without following a pre-set and specific purpose[214] as Snowden revelations brought into bright light.

According to Couldry, a life that can be continuously tracked is "a dispossessed life" and "recognizing this dispossession is the start of resistance to data

[211] See Marshall, "Algorithmic Misogynoir in Content Moderation Practice"; Christina Dinar, "The State of Content Moderation for the Lgbtiqa+ Community and the Role of the Eu Digital Services Act," (2021), https://eu.boell.org/sites/default/files/2021-06/HBS-e-paper-state-platform-moderation-for-LGBTQI-200621_FINAL.pdf.

[212] See Eubanks, "Automating Inequality : How High-Tech Tools Profile, Police, and Punish the Poor." For the oversurveillance of the poor and how it is related to structural disadvantages.

[213] Shephard, "Big Data and Sexual Surveillance."

[214] José van Dijck, "Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology," surveillance and society 12 (2014): 205.

colonialism."[215] One significant point Haggerty and Ericson make is how unlike the underclass proles in Orwell's novel 1984, the ordinary people in our real life dystopia have not been excluded from the interest of surveillance. Foucault had already shown the disciplinary surveillance was specifically targeting populations that deemed to need disciplinary control via soul training. However, Haggerty and Ericson discuss that today's surveillant assemblage is not so much about soul training, although control and profit are among its desires. Even though they accept the fact that the targeting of surveillance is differential in its controlling and discriminating effects, let's say between WMG and the mainstream populations, they argue the new rhizomatic surveillance leaves no one outside of the sight of the gaze and is capable of keeping the powerful under scrutiny.[216] As Foucault put it, "where there is power, there is resistance."[217] Browne narrates the act of staring back at the white gaze as resistance may start with a look right back in the eyes of the oppressors.[218] Thus, sousveillance[219] or synopticism[220] have been seen as a way of challenging the power relations by watching the watchers. However, in the absence of full access and understanding of these assemblages, as well as in the lack of consequences of the exposed wrongdoings of the powerful,[221] it is not

[215] Couldry and Mejias, "Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject," 345.

[216] Haggerty and Ericson, "The Surveillant Assemblage."

[217] Michel Foucault, The History of Sexuality. Volume 1: An Introduction (New York: Pantheon Books, 1978).

[218] Browne, "Dark Matters : On the Surveillance of Blackness."

[219] Steve Mann and Joseph Ferenbok, "New Media and the Power Politics of Sousveillance in a Surveillance-Dominated World," Surveillance & Society 11, no. 1/2 (2013).

[220] Thomas Mathiesen, "The Viewer Society:Michel Foucault's `Panopticon' Revisited," Theoretical Criminology 1, no. 2 (1997).

[221] "Why filming police violence has done nothing to stop it" <https://www.technologyreview.com/2020/06/03/1002587/sousveillance-george-floyd-police-body-cams/> Accessed 15 May 2022.

sufficient[222] to bring balance to the distribution of power.[223] This is the framework that provides a critical eye not only on the concept of privacy but also on our demands for transparency.

More than two decades after 9/11 and almost one decade after Snowden revelations, it is clear that being watched does not have the same disciplinary effect on all, and accountability is not the direct result of being caught in the act, at least not for everyone. Annanny's[224] definition of algorithmic systems as "assemblages of institutionally situated computational code, human practices, and normative logics that creates, sustains, and signifies relationships among people and data through minimally observable, semiautonomous action"[225] helps nuancing our stance. Thus watching the watchers to be an effective strategy, one should be able to watch the assemblage as a whole.

This extends to making room for intervention in every aspect of any data controller's whole organization and the whole lifecycle of their algorithms including when they are only forming in the heads of business strategists or computer scientists. We shall also look into the whole system of machine learning led surveillance and dataveillance, as regards to the interconnectedness and distributedness of various data collection and processing operations by various public and private organisations exchanging and integrating of data at the level of

[222] Mike Ananny and Kate Crawford, "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability," New Media & Society 20, no. 3 (2018).
[223] See "Chapter 9: Governments Strike Back" in Zeynep Tufekci, Twitter and Tear Gas: The Power and Fragility of Networked Protest, (United States of America: Yale University Press, 2017), https://www.twitterandteargas.org/downloads/twitter-and-tear-gas-by-zeynep-tufekci.pdf.
[224] Mike Ananny, "Toward an Ethics of Algorithms:Convening, Observation, Probability, and Timeliness," Science, Technology, & Human Values 41, no. 1 (2016).
[225] Ibid., 7.

machines.[226] According to Solove, the disinterestedness of these machines on who they are watching, combined with the lack of knowledge and participation of those who are being watched, creates a vulnerability reminiscent of Josepf K., while the harms are reminiscent of those caused by bureaucracies such as "indifference, errors, abuses, frustration, and lack of transparency and accountability."[227] We need to make an intervention here before we continue with this valuable contribution, that dataveillance in this sense is a little bit of white man's problem. As WMG are more subject to surveillance for exactly because of who they are or more correctly put, which group they are affiliated with. Thus they actually know by whom and why records are being kept or likely to be used about them, for them it is more of a "how?" question.[228]

According to Hildebrandt concerns related to the effects of this picture drawn by Solove that is in the form of conforming to the expectations of these watchers indicates to spontaneous normalisation in Foucault's panopticon. On the other side of the coin, there is another normalisation process addressed by Lessig,[229] that is customisation of online options based on observed patterns in the behaviour of web users, turning into personalised profiles which define what individual users see, the behaviour developed and observed within this closed system of options is being fed back into the profiles, in this cycle probable preferences are being

---

[226] Mireille Hildebrandt, "Profiling and the Identity of the European Citizen," in *Profiling the European Citizen: Cross-Disciplinary Perspectives*, ed. Mireille Hildebrandt and Serge Gutwirth (Dordrecht: Springer Netherlands, 2008), 306.

[227] Joseph K. is the protagonist of Franz Kafka's novel in "the Trial." See. Solove, "'I've Got Nothing to Hide' and Other Misunderstandings of Privacy," 766; "Privacy and Power: Computer Databases and Metaphors for Information Privacy," *Stanford Law Review* 53 (2001).

[228] Bartlett calls it feminist knowing. Here we do not refer to a hierarchically higher epistemological ground as it was criticized by Haraway. We refer to how it looks from where they stand, thus we refer to the standpoint epistemology.

[229] Lessig, *Code 2.0*, 220.

reinforced in a way it shapes those who are observed (or profiled).[230] Especially for dynamic machine learning models that run in the real world on real data, these cycles may turn into self-fulfilling prophesies as we mentioned in the previous section.

Hildebrandt defines profiling as "the process of 'discovering' correlations between data in databases that can be used to identify and represent a human or nonhuman subject (individual or group) and/or the application of profiles (sets of correlated data) to individuate and represent a subject or to identify a subject as a member of a group or category."[231]

As we mentioned above, most online services follow an advertising business model, which is mostly based on tracking and profiling the web users in order to deliver personalized advertisements. There are a couple of tracking tools that enable such monitoring such as spyware, web bugs, hidden identifiers, web tracking pixels and tracking cookies which are installed in the terminal equipment of the user.[232] Another remote tracking mechanism is known as fingerprinting. It is based on the fact "many technological devices possess subtle but measurable variations" which may be unique enough to identify the user.[233] Many websites use fingerprinting algorithms, which collect and combine information about the browser, network, device language, keyboard layout, time zone, version of the operating system, in order to identify the user. This allows for tracking and identifying the user in a high certainty even if cookies are disallowed on the

[230] Hildebrandt, "Profiling and the Identity of the European Citizen," 306-07.

[231] "Defining Profiling: A New Type of Knowledge?," 19.

[232] Proposal for Regulation on Privacy and Electronic Communications (E-Privacy Regulation), Recital 20.

[233] Peter Eckersley, "How Unique Is Your Web Browser?" (Berlin, Heidelberg, 2010), 1.

browser or blocked by the user.[234] Cookies are small textfiles which enable ad networks to recognize the user in their repeating use of the web site or any other partner web site of the ad network. These are third party cookies as normally only the website who installed the cookie can read it, however, browsers may install cookies which monitor first party cookies installed by the web sites themselves or third parties may partner with the websites so that they share the information collected by their first party cookies with their partners. Cookies collect and retain information about the users' device, browser settings, IP address, location, what the users have read or watched, how long they stayed on the web site, which ads they clicked and so on.[235] All of this information is attached to the user's cookie ID which is sent by the website to a supply-side platform (SSP), who sends it to the ad exchange where a real-time bidding takes place between demand-side suppliers (DSPs). DSPs represent advertisers and when a bid is up, they compare the cookie ID and attached information with their databases to find connections to campaigns they are running for the advertisers, they place their bid according to this evaluation and the winner gets to display their campaign in the advertisement space reserved on the web site.[236] All of this happens in milliseconds[237] and involves ADM.[238]

---

[234] Matt Burgess, The Quiet Way Advertisers Are Tracking Your Browsing, 2022. <https://www.wired.com/story/browser-fingerprinting-tracking-explained/> Acessed on 10 September 2022.

[235] Karolina Iwańska, Behavioural Advertising 101, 2020. <https://medium.com/@ka.iwanska/behavioural-advertising-101-5fee17913b22> Accessed on 10 September 2022.

[236] Article 29 Working Party, "Opinion 2/2010 on Online Behavioural Advertising " WP 171 (2010): 6.

[237] Ezgi Eren et al., "Increasing Fairness in Targeted Advertising: The Risk of Gender Stereotyping by Job Ad Algorithms.," ed. N. Birner, et al. (2021).

[238] Article 29 Working Party, "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679," WP 251 (2018), https://ec.europa.eu/newsroom/article29/items/612053

On social media platforms, the platform itself plays a central role in targeting users with advertisement. There are three main aspects of this process, ad targeting by advertisers, ad delivery by platforms and ad display to users. First of all, in ad targeting, advertisers are enabled by the platform's advertisement interface to choose their audience based on selection categories such as demographic and attributes that are derived and inferred from age, gender and location information, profile information, activities on the platform and information retrieved from third parties; personal information, which means the advertiser may upload exact target audience person by person; and similar users to their choices which is determined on the basis of the classification made by the platform, such as lookalikes on Facebook. Advertisers may monitor the performance of their campaign via tools made available by the platform.[239] Once a user is using the platform, all the live ads fitting the profile of the user is gathered and a relevance score is computed by the platform, then a real-time bidding process runs as described above for the web sites, however the relevance score[240] is also taken into consideration in choosing the winning bid, thus a higher bid may lose to a lower bid with a higher relevance.[241] This is one of the strategies through which platforms keep their users engaged. The winner is displayed to the user, which determines what the user will see on the platform, which gives the platforms massive control over the perception of the users about the world and themselves, which may shape them and their futures without them being aware of it as we discussed above.

---

[239] Eren et al., "Increasing Fairness in Targeted Advertising: The Risk of Gender Stereotyping by Job Ad Algorithms.," 6; Muhammad Ali et al., "Discrimination through Optimization: How Facebook's Ad Delivery Can Lead to Biased Outcomes," *Proceedings of the ACM on Human-Computer Interaction* 3 (2019).

[240] As well as estimated action rates in the case of Facebook, which refers to the performance of the ad in engagement based on the general users, as opposed to relevance score is the likelihood of a particular user's engagement with the ad. "Discrimination through Optimization: How Facebook's Ad Delivery Can Lead to Biased Outcomes."

[241] Ibid.

There has been several studies which found out how this process of determining what users will see online, can be informed by stereotypes about gender, age, race and so on.[242] However, both the technological and organisational opaqueness of the process have left the researchers without a proof about how this exactly happens.[243] We will get back to the discriminatory effects of this process below when we discuss job ads as a part of our lead example.

Recently the industry giants are showcasing their commitment to privacy due to user expectations and regulatory pressure by making tracking of users harder such as Apple's ATT (App Tracking Transparency) and Google's FloC (Federated Learning of Cohorts)[244] which was then announced to be replaced with a new approach called Topics.[245] According to some authors, these can be shown as examples to when big private companies are left to self-fix, power relations are left untouched if not shifted more towards themselves. What happens is that not only now in tracking and profiling industry becomes more dependent on Apple and Google by the cut of third party cookies, it also promotes a false promise about privacy. If privacy is about levelling the power relations by leaving room for the

---

[242] Wulf, *Automated Decision-Making Systems and Discrimination: Understanding Causes, Recognizing Cases, Supporting Those Affected* ; Latanya Sweeney, "Discrimination in Online Ad Delivery," *Commun. ACM* 56, no. 5 (2013); Anja Lambrecht and Catherine Tucker, "Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of Stem Career Ads," *SSRN Electronic Journal* (2016); Amit Datta, Michael Carl Tschantz, and Anupam Datta, "Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination," *ArXiv* abs/1408.6491 (2014).

[243] Eren et al., "Increasing Fairness in Targeted Advertising: The Risk of Gender Stereotyping by Job Ad Algorithms.."

[244] Caitlin Fennessy and Nicole Sakin, "How Google and Apple Are Shaking up Adtech," (2021), https://iapp.org/news/a/how-google-and-apple-are-shaking-up-adtech/.Accessed on 15 October 2022.

[245] Joseph Duball, "Changing Course: Google Maneuvers toward New User-Tracking Norms," (2022), https://iapp.org/news/a/changing-course-google-maneuvers-toward-new-user-tracking-norms/. Accessed on 15 October 2022.

individuals to be in control of their informational self-determination, this is not happening here. Even though, the profiles are not "personalized" or personal data is not being shared with third parties; the very first party is shaping users' environment by offering them personalized services nevertheless and this is clearly an exercise of power over them. Just because users privacy is protected, it does not mean they are in control,[246] while data protection is exactly aiming at that.

1.2.4 Mitigating Algorithmic Discrimination (FAT/ML)

In this section we will introduce the technical solutions to mitigating algorithmic bias which may or may not result in discrimination. There is a very lively and ever progressing scientific debate on these topics, thus we are not able to provide the whole scope of technical solutions here, especially the most up-to-date state-of-the-art, neither have we the capability to assess their technical appropriateness or correctness.[247] However, this introduction is important to

---

[246] Reuben Binns et al., "Third Party Tracking in the Mobile Ecosystem," in Proceedings of the 10th ACM Conference on Web Science (Amsterdam, Netherlands: Association for Computing Machinery, 2018).

[247] As example to the related conversation in the scientific community see Schwartz Reva et al., "Towards a Standard for Identifying and Managing Bias in Artificial Intelligence," (Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, 2022); Bruno Lepri et al., "Fair, Transparent, and Accountable Algorithmic Decision-Making Processes," *Philosophy & Technology* 31, no. 4 (2018); A. Koene, L. Dowthwaite, and S. Seth, "Ieee P7003tm Standard for Algorithmic Bias Considerations" (paper presented at the 2018 IEEE/ACM International Workshop on Software Fairness (FairWare), 29-29 May 2018 2018); P. Phillips et al., "Four Principles of Explainable Artificial Intelligence," (NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, Gaithersburg, MD, 2021); Moritz Hardt, Eric Price, and Nathan Srebro, "Equality of Opportunity in Supervised Learning," *ArXiv* abs/1610.02413 (2016); Matthew Joseph et al., "Rawlsian Fairness for Machine Learning," ibid.abs/1610.09559; Reuben Binns, "Fairness in Machine Learning: Lessons from Political Philosophy," *Decision-Making in Computational Design & Technology eJournal* (2018); X. Ferrer et al., "Bias and Discrimination in Ai: A Cross-Disciplinary Perspective," *IEEE Technology and*

understand the difference in terminology for the scientific community and the legal community as fairness, accountability and transparency are used in both groups but they may mean something completely different or refer to a different aspect of the problem.

1.2.4.1 Fairness

As introduced above, law has been complicit in perpetuating inequalities. Both in law and in machine learning there are rules to achieve fairness. What kind of fairness/equality do we want to achieve with anti-discrimination laws and rules is a vital question as the chosen notion of equality or fairness would affect the outcome.[248] We will come back to how the antidiscrimination law uses these notions of equality in employment related cases and what it means for inclusive AI but before that let's have a look at plurality of fairness definitions and metrics in machine learning.

Equality of opportunity in ML fairness would be satisfied for the gender aspect of the preferred "hired" label, if qualified job applicants "are equally likely to be hired irrespective of" their gender. For example, if the qualified men and women both have a 50% chance of being hired, even though the percentage of women

---

Society Magazine 40, no. 2 (2021); S. Verma and J. Rubin, "Fairness Definitions Explained" (paper presented at the 2018 IEEE/ACM International Workshop on Software Fairness (FairWare), 29-29 May 2018 2018); Rebekah Overdorf et al., "Questioning the Assumptions Behind Fairness Solutions," *ArXiv* abs/1811.11293 (2018); Alice Xiang and Inioluwa Deborah Raji, "On the Legal Compatibility of Fairness Definitions," ibid.abs/1912.00761 (2019); James Foulds et al., "An Intersectional Definition of Fairness," *Journal Name: IEEE International Conference on Data Engineering* (2020); ibid.

[248] Sandra Wachter, Brent Daniel Mittelstadt, and Chris Russell, "Bias Preservation in Machine Learning: The Legality of Fairness Metrics under Eu Non-Discrimination Law," *SSRN Electronic Journal* (2021).

hired is significantly lower than men, this metric is satisfied[249] as it only requires the ratio of true positive decisions (qualified and hired) to match the false negative decisions (qualified and rejected). However other metrics would not be satisfied. For example, *equalized odds* wouldn't be because more women are rejected than men. It would be satisfied provided that regardless of a job applicant's gender, being qualified would result in equally as likely to be hired, and the same would apply to rejections for those who are unqualified.[250] This metric requires the ratio of true positive decisions (qualified and hired) to match the false negative decisions (qualified and rejected), as well as true negative decisions (unqualified and rejected) to match false positive decisions (unqualified and hired) across populations.[251] According to Wachter et. al, the dependence on such bias preserving metrics on the ground truth data makes such evaluation particularly difficult because distribution of groups may shift between training and deployment stages.[252] Another reason is that ground truth data may not always be available or correct for all populations. As the ground truth data can also be generated as the model runs in the real life, it becomes even more unexpected for the negative decisions to be possible to be weighed in order to check whether these metrics are satisfied, as this would require hiring seemingly unqualified applicants to produce false positives to satisfy equalized odds.[253]

If at the end of the day less women were admitted than men, this time *demographic parity* would not be satisfied as bias transforming metrics require matching decision rates as opposed to bias preserving metrics such as the two

---

[249] See <https://developers.google.com/machine-learning/glossary#k-means> Accessed 17 June 2022.

[250] See <https://developers.google.com/machine-learning/glossary#k-means> Accessed 17 June 2022.

[251] Wachter, Mittelstadt, and Russell, "Bias Preservation in Machine Learning: The Legality of Fairness Metrics under Eu Non-Discrimination Law."

[252] Ibid.

[253] Ibid.

examples above, which require matching error rates across groups.[254] Irrespective of the general qualification of the group, if the same percentage of men and women were hired, "demographic parity" would be satisfied whereas both "equality of opportunity" and "equalized odds" would not be.[255]

"Conditional demographic parity," if conditioned on performance score, would require both women and men scored over a threshold to receive positive decisions for promotion, while both some women and men scored under the threshold to be also promoted.[256] Here, Wachter et al indicate to another problem as the performance score is a proxy for what is actually wished to be predicted such as being a good manager. This is exactly how social bias enters the system[257] as time and time again managers (due to vertical segregation prone to be disproportionately male) give higher performance score to male employees[258] or the very components of the score is based on parameters better fitting male lives such as flexibility, availability, longer working hours,[259] or unable to measure female leadership based on soft skills.[260]

Another crucial point while evaluating fairness metrics is whether they are fit for the context. In case of the employment context, demographic or conditional demographic parity would be fair in its societal effects as well, if it leads to giving

---

[254] Ibid., 29.

[255] For more machine learning definitions see <https://developers.google.com/machine-learning/glossary#k-means> Accessed 17 June 2022.

[256] Wachter, Mittelstadt, and Russell, "Bias Preservation in Machine Learning: The Legality of Fairness Metrics under Eu Non-Discrimination Law."

[257] Ibid.

[258] Kim, "Data-Driven Discrimination at Work," 1301. See also ingroup bias and outgroup bias in Sonia K. Katyal, "Private Accountability in the Age of Artificial Intelligence," *UCLA L. REV.* 66 (2019): 80-81.

[259] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 56.

[260] Kim, "Data-Driven Discrimination at Work."

a chance to those who normally would have been rejected but may prove to be a good asset on the job and change the ground truth for future hires. On the other hand, a bank giving loans to false positives in order to satisfy these metrics would lead to disadvantaged group members having loans that they are not able to pay back and sucked into a spiral of debt and poverty adding up to structural disadvantages,[261] which is not a societal result to hope for.

As it is seen there are many ways of achieving one form or another form of fairness in machine learning. In this regard, the developers of the model must take the context and the use case into consideration while choosing which fairness definition to use. In this case again, domain experts'AW opinions and the inputs from the gender theory on equality, fairness and justice obtained via social scientists, as well as data produced by and for disadvantaged groups[262] may help choosing the fairness metric and sensitive attributes, as well as to identify the risks better.

However, there is still the problem that legal requirements, feminist concerns and calculatable ML fairness do not match each other most of the time. Also, it is really hard to consider and achieve intersectional fairness.[263] One thing to keep in mind is that the unequal results or discrimination against a group at the results of application might be resulting from the choice of fairness notion and how it fit to the context, as well as how we defined sensitive attributes. We will later refer to fairness metrics in relation with antidiscrimination law and data protection law.

---

[261] Wachter, Mittelstadt, and Russell, "Bias Preservation in Machine Learning: The Legality of Fairness Metrics under Eu Non-Discrimination Law."

[262] D'Ignazio and Klein, *Data Feminism*.

[263] At least one study claimed to achieve intersectional fairness. See Foulds et al., "An Intersectional Definition of Fairness."

1.2.4.2 Accountability

It is important that ADM systems work as they are intended (*ex ante*) and allow for review and oversight after a decision is made (*ex post*).[264] Thus, accountability embraces fairness and transparency to be built-in from the beginning to ensure designing a process that does not cause harm such as discrimination and that is open to scrutiny. As Kroll et. al put it, it is about ensuring ADM processes are in accordance with procedural and substantive goals that make it acceptable. Accordingly, procedural goal would be first to achieve procedural irregularity, meaning that a rule, a goal or a policy is applied to every decision by an ADM process, in other words each individual decision follows the same rule. Second procedural goal is that the ADM system should be open to scrutiny, such as being transparent or explicable. The substantive goal, on the other hand, would be how the ADM process interplayed with substantive principles and policy choices such as fairness, whether it is morally, ethically, legally correct and consistent.

We will discuss transparency more in detail below. Alongside transparency, audits may verify the way a model works but they treat the decision process like a black-box that its input and output may be observed while the inner workings remain secret. This is not sufficient as it doesn't help justifying a decision, in the substantive sense of accountability. Software verification, cryptographic commitments, zero-knowledge proofs, fair random choices are suggested to be used alongside with transparency and audits as a means to providing for procedural regularity. Kroll et. al suggest that applying all of these technical tools may provide greater accountability than transparency in the sense of disclosing the source code. In this regard, systems can be designed to publish commitments as to what they will do, after deployment whether the actual actions of the system match with the commitments can be ensured through zero-knowledge proofs, which can

[264] Joshua Kroll et al., "Accountable Algorithms," University of Pennsylvania Law Review 165 (2017).

be made public. This ensures that the decisions were not arbitrary but a product of predetermined policy.[265]

Kroll et. al indicate to a remaining clash between what computer scientists and law-and policymakers understand from accountability, similar to the difference in terminology in fairness that we introduced above. For computer scientists who are building an automated system as it will be run by machines who do not have capacity for interpretation that is informed from substantive principles, behaviour of the system should be properly defined and specified in detail. Thus computer scientists must design explainable, reviewable systems that are open to *ex post* oversight and justification. Law makers, however, as a result of various reasons design rules that are open to interpretation, where the gaps left to case by case judgements of the courts that take into consideration specificities of each conflict. First of all, law making is a political process, which means lack of majority support for a policy goal may lead to a compromised vague language, such as it is seen for example in the GDPR's making process explicit prohibition of algorithmic discrimination or granting a right to explanation are left to non-binding Recitals[266] which are causing debates among lawyers. Second, laws are made with long-term effect in mind, thus a level of flexibility is intended to compromise for not being able to consider all the possibilities. Another reason for the laws to remain technology-neutral. Third, the lawmakers may be able to identify the problem but may not be sure about the best method of solving it. By drafting a text that is open to interpretation, they leave room for experimentation.[267] According to Kroll et. al,

---

[265] Ibid.

[266] See Wachter, Mittelstadt, and Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation."

[267] Kroll et al., "Accountable Algorithms," 695-99. For a discussion on how the vagueness of the law is not a bug but a feature, see Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai ". On the difficulty of translating data protection law into code see Bert-Jaap Koops and Ronald E. Leenes, "Privacy

lawmakers, however, must take into account that in the case that policy goals are not well specified in laws, it means that filling in the details related to ADM systems which will take effect on large populations are left to computer scientists, who are unlikely to have substantive knowledge of the policy goal and unlikely to be held responsible for the large societal effects.[268]

1.2.4.3 Transparency (Explainability)

While transparency has been seen as a prerequisite to achieve accountability of algorithmic systems, Kroll et.al. argued that full disclosure of the source code is not necessary which may have its own downsides such as being illegible to lay persons and not being able to fully predict a program's behaviour.[269] Transparency is not even always wanted because (i) it may lead to gaming of the system, reveal sensitive information, be prohibited by law; (ii) cannot sufficiently verify if the systems were not designed with accountability in mind, (iii) randomness of the system may be desired, but the results may be irreducible, (iv) especially for models running in real-time the rule may become obsolete the moment it is made.[270]

On this account, the procedural goal seems to shift to explainability rather than full transparency. As in Franz Kafka's "The Trial" when the protagonist Joseph K. is not provided with any explanation as to the reason of his arrest, the grounds

---

Regulation Cannot Be Hardcoded. A Critical Comment on the 'Privacy by Design' Provision in Data-Protection Law," *International Review of Law, Computers & Technology* 28 (2014). How purpose limation principle allows for dividing legal responsibilities under the GDPR until a granular level that allows for automation, see Maximilian von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Iii," *European Data Protection Law Review* 7, no. 3 (2021).

[268] Kroll et al., "Accountable Algorithms," 700.

[269] Ibid.

[270] Ibid., 658-60.

upon which a decision about him would be based, who prosecutes him, how does the process work and how it can be challenged or reversed;[271] the black box ADM processes which are not explainable would undermine autonomy, human dignity and agency.[272] Thus, explainable AI (XAI) has become an important field of work that may mitigate such negative effects of ADM processes as well as discrimination harms they may cause.[273]

As mentioned above, instead of a pure technical process ADM systems can be better understood and addressed as assemblages.[274] They do not only consist of various actors (individual humans, institutions, organisations) and technologies but also they are dynamic in the sense that they develop in different stages and change over time.[275] Although, we agree with the argument that looking into this complex socio-technical ecosystem[276] does not necessarily provide understanding of its functioning and this may lead to giving false hopes of accountability which would then lead to even less accountability.[277]

---

[271] Andrew D. Selbst, "An Institutional View of Algorithmic Impact Assessments," *Harvard Journal of Law & Technology* 35, no. 1 (2021); Danielle Keats Citron, "Technological Due Process," *WASH. U. L. REV.* 85, no. 6 (2007); Solove, "Privacy and Power: Computer Databases and Metaphors for Information Privacy."; ibid.; Daniel J. Solove, *The Digital Person: Technology and Privacy in the Information Age* (New York, USA: New York University Press, 2004); "Privacy and Power: Computer Databases and Metaphors for Information Privacy."; *The Digital Person: Technology and Privacy in the Information Age*; Mireille Hildebrandt, "Profiling and the Rule of Law," *Identity in the Information Society* 1, no. 1 (2008).

[272] Selbst and Powles, "Meaningful Information and the Right to Explanation."

[273] Asghari et al., "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making.".

[274] Ananny, "Toward an Ethics of Algorithms:Convening, Observation, Probability, and Timeliness."

[275] Asghari et al., "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making.".

[276] Ibid.

[277] Ananny and Crawford, "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability."

It is also true that explaining such an assemblage should be concerned with questions of who needs the explanation for what purpose and how can the explanation be tailored for this target audience and purpose. Such a contextual approach also calls for the involvement of the user who is receiving the explanation in the evaluation of XAI. This is true if we compare the kind of explanation a consumer whose loan application was refused needs to challenge such decision with the explanation the developers of the model need in order to fix bias in the ADM system.[278] In order to challenge such a loan refusal decision, a counterfactual explanation[279] might be sufficient.

Recently, in a case before the Finnish National Non-Discrimination and Equality Tribunal, the investigation brought out to light that the applicant being a woman or living in an urban area or having Swedish as his mother tongue (instead of Finnish) negatively affected the loan to be extended to him. The case resulted in such practice based on pure statistics instead of the financial status being prohibited as the counterfactual explanation indicates to differential treatment based on protected grounds.[280] We will discuss how sufficient such an explanation to challenge the automated decision under antidiscrimination law and data protection law later. However, it is a good example for what counterfactual explanations provide to a consumer who would like to challenge an ADM that concerned them.

When the need for explainability in ADM is being discussed, we often encounter with the black box problem. However, not all types of ADM have to operate as black boxes. Asghari et.al distinguish three general types of ADM

---

[278] Asghari et al., "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making.".

[279] Wachter, Mittelstadt, and Russell, "Counterfactual Explanations without Opening the Black Box: Automated Decisions and the Gdpr."

[280] Carsten Orwat, "Risks of Discrimination through the Use of Algorithms. A Study Compiled with a Grant from the Federal Anti-Discrimination Agency" (2020), 38-39.

systems with different levels of explainability: (i) symbolic reasoning which can be explained in a way that humans can easily understand but the performance is much lower such as expert systems, (ii) classical machine learning can be inspected well enough by humans in some models such as decision trees, some other models require higher efforts to become interpretable for humans, and finally (iii) deep learning is where most of the problems discussed as the black box problem come from.[281] These systems work with a high performance as they are able to catch patterns that humans cannot because humans do not think that way, which is the exact reason that they remain uninterpretable.[282]

This approach also indicates to another related problem in machine learning known as the accuracy vs interpretability trade-off. However, decisions made about or effect humans have to be justifiable and contestable in order to be legitimate.[283] Thus such a trade-off is not acceptable for important decisions such as hiring for a job or extending credit. On the other hand, a study found out that accuracy is favoured over explainability by the recipients of the decision in healthcare related scenarios.[284] However, as long as the accuracy is not hundred per cent, which is not likely, then the recipient of a wrong decision or their family would most likely demand an explanation also in a healthcare scenario. According to the suggestion of Asghari et.al, in such a case a domain expert (medical doctor of the wrong ADM's recipient) or a NGOs or representatives of special interests groups might receive the explanation, then help the recipient to understand and/or

[281] Asghari et al., "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making.".

[282] Ibid.

[283] Ibid.

[284] Sabine N van der Veer et al., "Trading Off Accuracy and Explainability in Ai Decision-Making: Findings from 2 Citizens' Juries," Journal of the American Medical Informatics Association 28, no. 10 (2021).

to contest.[285]

Despite there are ways such as natural language explanations or heatmaps that try to explain the inner workings and behaviour of complex black box models, according to Asghari et.al, the best way of providing a meaningful level of transparency is to ensure "the highest amount of transparency at the input level."[286] This would require documenting the processes of data gathering, data preparation such as datasheets, architectural choices such as an explanation as to why a large and complicated model had to be used, as well as choices made during testing and deployment.[287] Kroll et. al distinguish between static methods such as reading the source code and dynamic testing, which may hint the future behaviour of a program but are limited to number of inputs that can be tested, not allowing for every possible output to be observed. Kroll et. al favour white box testing as in addition to consideration of input and output in black box testing, it also shows which inputs are more predictive of the output, or played a role in reaching the output. While being simple, logging is well-known and it also allows for dynamic *ex post* review.

Furthermore, it must be noted that "online" machine learning systems that interact with their environment in making real-time decisions can update their predictive model after each decision, as they incorporate them into their training data. In such cases, how they interact with their environment as well as the time of interaction also have to be known.[288]

---

[285] Asghari et al., "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making.".

[286] Ibid., 17.

[287] Ibid.

[288] Kroll et al., "Accountable Algorithms."

1.3 A LEAD EXAMPLE: ALGORTIHMIC DISCRIMINATION IN AI ASSISTED HUMAN RESOURCES & E-RECRUITMENT

Justice, equality and privacy are all context dependent concepts,[289] meaning that they may take different definitions or may be understood differently dependent on the context and the human subject experiences within that context. As many aspects of life is becoming automated, algorithmic discrimination is likely to appear in a variety of contexts raising questions of justice, equality and privacy.

While the technological aspects that are introduced above remain at the heart of the algorithmic discrimination problem, the effects such discrimination produce may differ wildly based on the context. We are in the opinion that addressing all the contexts at once leads to missing nuances of each of them. Therefore, even though within the text we will be referring to examples that have or may occur in different contexts, we prefer to have a lead example that is nuanced enough to be referred to in our legal analysis of the EU law applicable to algorithmic discrimination.

Alongside employment, assessment of creditworthiness has been mentioned both in the GDPR as it may have discriminatory effects and in AI Act as a high risk AI system. In a very recent example from Hesse in Germany, a woman chose the "purchase on account" option, which required a credit check, during her online shopping for clothes, her suspicious rejection led her to call the external credit agency the online store was working with to assess creditworthiness. At this call she learned that women around the age of 40 would be given an inadequate score

[289] W. B. Gallie, "Essentially Contested Concepts," *Proceedings of the Aristotelian Society* 56 (1955); Deirdre K. Mulligan, Colin Koopman, and Nick Doty, "Privacy Is an Essentially Contested Concept: A Multi-Dimensional Analytic for Mapping Privacy," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374 (2016).

due to being "divorced and therefore destitute."[290] Algorithmic credit discrimination as a solely automated decision is more expected in low value online credits. However use of alternative data sources and more complex machine learning systems by financial service providers[291] may also cause similar effects due to automation bias.

Housing is a protected domain under US antidiscrimination law and housing ads on Facebook has been under scrutiny.[292] ProPublica bought ads on Facebook and managed to exclude as many people belonging to protected categories such as African Americans, mothers of high school kids, people interested in wheelchair ramps, Jews, expats from Argentina and Spanish speakers, interested in Islam, Sunni Islam and Shia Islam, soccer moms, people interested in American sign language, gay men and Christians to be excluded from their ad campaign, and they were able to perform redlining by excluding zip codes where predominantly minorities resided.[293]

---

[290] Wulf, *Automated Decision-Making Systems and Discrimination: Understanding Causes, Recognizing Cases, Supporting Those Affected*

[291]Mikella Hurley and Julius Adebayo, "Credit Scoring in the Era of Big Data," *Yale J.L. & Tech* 18, no. 1 (2017); Nikita Aggarwal, "The Norms of Algorithmic Credit Scoring," (2020); Federico Ferretti, "Not-So-Big and Big Credit Data between Traditional Consumer Finance, Fintechs, and the Banking Union: Old and New Challenges in an Enduring Eu Policy and Legal Conundrum," *Global Jurist* 18, no. 1 (2018).

[292] On 21 June 2022, Facebook/Meta and Department of Housing and Urban Development signed a settlement. Facebook/Meta agreed to a couple of measures including Facebook/Meta platforms to stop allowing lookalike audience option in housing ads and developing a Variance Reduction System, which will balance the audience fitting the choices of advertisers and actual populations. <https://about.fb.com/wp-content/uploads/2022/06/June_2022_HUD_Settlement_Agreement.pdf> Accessed on 25 June 2022.

[293] Julia Angwin, Ariana Tobin, and Madeleine Varner, "Facebook (Still) Letting Housing Advertisers Exclude Users by Race," https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin; Julia Angwin and Terry Parris Jr., "Facebook Lets Advertisers Exclude Users by Race," https://www.propublica.org/article/facebook-lets-advertisers-

In education, there are examples where ADM is used in Belgium, in order to place elementary and secondary children in schools, which so far seems to be making the system fairer.[294] On the other hand, the results of A-levels algorithm in the UK, which assigned test results to students who could not take the test due to Covid-19 lockdown in the country based on the test results from previous years, had to be cancelled as a result of widespread protests. The algorithm assigned 40 percent lower results than those assessed by teachers, it advantaged students from private schools and disadvantaged students from historically underperforming schools.[295]

In the public sector distribution of welfare benefits is an area where there are many examples of ADM. One prominent example is Dutch child benefits fraud algorithm that we introduced above. In her book "Automating Inequality," Virginia Eubanks gives examples from public health insurance, matching unhoused people with housing opportunities based on vulnerability assessment, to a predictive tool to prevent child abuse, all of which concern the poor and exacerbate the disadvantage they experience.[296] In Poland, Ministry of Labour introduced an ADM that determine the "distance" from the labour market and the willingness to enter or re-enter the labour market based on assessment of 24 characteristics including age, gender and degree of disability, information on periods of childcare and care of persons requiring assistance. It build a three-tier system, in which being assigned to Profile III would receive less support. The system found to be likely to cause both direct and indirect discrimination based on

---

exclude-users-by-race; Orwat, "Risks of Discrimination through the Use of Algorithms. A Study Compiled with a Grant from the Federal Anti-Discrimination Agency."

[294] Fabio Chiusi et al., eds., *Automating Society Report 2020* (Germany: AlgorithmWatch gGmbH & Bertelsmann Stiftung, 2020), 34-37.

[295] Ibid., 280; Orwat, "Risks of Discrimination through the Use of Algorithms. A Study Compiled with a Grant from the Federal Anti-Discrimination Agency."

[296] Eubanks, "Automating Inequality : How High-Tech Tools Profile, Police, and Punish the Poor."

gender.[297] In Austria a very similar system was introduced in job centres, which is also found to discriminate against women and those with caring responsibilities, who are again predominantly women, and limit their chances.[298]

Predictive policing tools have been shown to lead to more police control in areas where disadvantaged racial groups live.[299] Another infamous example is COMPAS system (Correctional Offender Management Profiling for Alternative Sanctions), which was designed to help judges in determining the risk of recidivism by calculating the probability of re-arrest based on crime statistics and public documentation. COMPAS resulted in false positives disproportionately for black defendants in contrast with high false negatives for white defendants.[300]

Why do we think human resources is a good lead example? First of all it is a very well documented field of discrimination. The most of EU Antidiscrimination Law instruments and therefore the case law of the CJEU are employment related, which is a crucial social good, exclusion from or subordination within may affect life chances of individuals and groups gravely. On the other hand, it covers all the technologies that may lead to discrimination. Moreover, an employer may be a public body or a private entity. Furthermore, human resources tools are used on and by platforms, large corporations, but also small and medium sized entities (SMEs), they may be developed in house but also there is a market for any

---

[297] Orwat, "Risks of Discrimination through the Use of Algorithms. A Study Compiled with a Grant from the Federal Anti-Discrimination Agency," 42-43; Chiusi et al., *Automating Society Report 2020*, 186.

[298] Doris Allhutter et al., "Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective," *Frontiers in Big Data* 3 (2020).

[299] Céline Castets-Renard, "Human Rights and Algorithmic Impact Assessment for Predictive Policing," in *Constitutional Challenges in the Algorithmic Society*, ed. Amnon Reichman, et al. (Cambridge: Cambridge University Press, 2021).

[300] Orwat, "Risks of Discrimination through the Use of Algorithms. A Study Compiled with a Grant from the Federal Anti-Discrimination Agency."

company to incorporate such technologies into their processes, all of which bring a completely different set of problems that may occur in building and deploying an ADM process.

Finally, women have been historically discriminated in the realm of paid work,[301] which offers a rich feminist analysis of the sources including intersectional discrimination, individual and social effects, and desired remedies, as well as involving activism. Despite improvements, equal access to this important economic resource has not yet been achieved. Though not sufficiently adequate, there are non-discrimination rules in many national and international legal frameworks. Meanwhile, the use of artificial intelligence (AI), more specifically machine learning algorithms, especially in the context of important decision-making processes such as hiring for a job, has emerged as a novel threat that gives rise to concerns of discriminatory harms that already exist to be reinforced and exacerbated in novel ways and scale.

Automation of hiring processes is an interesting one also because it requires many negative decisions to be made. A bank will extend a loan to anyone who meets the requirements as there is not a set amount of money that they will need to distribute among applicants,[302] at least at the scarcity level of a vacancy. Hiring decision for one open position with many applications though, will inevitably involve many rejections. Thus, e-recruitment should not be understood as one ADM but as a series of decisions which make use of different technologies. In this regard different stages of an ADM process has to be analysed contextually in order to understand why bias keep creeping in these systems and cause discriminatory harms. Moreover, one of the most important parts of a hiring process is making

---

[301] See, Fredman, *Women and the Law*; S. Walby, *Theorizing Patriarchy* (Basil Blackwell, 1990), 25-60.

[302] Save requirements of allocative efficiency of consumer credit markets and internal governance of banks.

the public aware of a job vacancy,[303] which is in today's world mostly done online and fall within the framework of online delivery of advertisements which is targeted to related profiles.

On the other hand, automation in making human resources (HR) decisions such as hiring and promoting has been introduced as a tool for combatting human bias and subjective discriminatory decisions which have notoriously followed systemic patterns against disadvantaged groups. However, the novelty of threat by machine learned algorithms comes from being data-driven[304] and data, as feminist and other critical scientists have long shown, is not neutral. Thus, in addition to introducing new forms of bias, automated decisions based on data are also likely to perpetuate and amplify existing bias in the society even when they mean well. There is a seemingly growing ecosystem in which several vendors offer diversity-and-inclusion-as-a-service solutions, promising that their de-biased models will lead to more inclusive decisions and diverse teams. Some studies have already looked into how such technologies work.[305] I will benefit from these studies in order to understand the challenges that these technological solutions are likely to tackle, leave unsolved, perpetuate or amplify.

---

[303] Paul Post and Rikki Holtmaat, "A False Start: Discrimination in Job Advertisements," *European Gender Equality Law Review*, no. 2/2014 (2015).

[304] Kim, "Data-Driven Discrimination at Work."

[305] Manish Raghavan et al., "Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain: Association for Computing Machinery, 2020); Prasanna Tambe, Peter Cappelli, and Valery Yakubovich, "Artificial Intelligence in Human Resources Management: Challenges and a Path Forward," *California Management Review* 61, no. 4 (2019); Max Langenkamp, A. Costa, and Chris Cheung, "Hiring Fairly in the Age of Algorithms," *ArXiv* abs/2004.07132 (2020); Javier Sánchez-Monedero, Lina Dencik, and Lilian Edwards, "What Does It Mean to 'Solve' the Problem of Discrimination in Hiring? Social, Technical and Legal Perspectives from the Uk on Automated Hiring Systems," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (Barcelona, Spain: Association for Computing Machinery, 2020); Bogen and Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias."

To sum, employment context offers a wide range of opportunities to nuance our understanding of algorithmic discrimination and the law applies to it.

1.3.1 Discrimination in Employment

Algorithms learn from data which is inevitably historical. As feminisms have articulated there are a number of reasons why women have historically been excluded from, segregated and underpaid in employment. First of all, women's confinement to the private sphere of family as a patriarchal strategy of control over women's bodies have played a significant role in women's exclusion from the public sphere.[306] The household was men's private life, where they took a break from writing the history in the public sphere and where women cared for the children and performed reproductive housework.[307] These were told to be women's virtues and they were not to be paid for. In contrast, anything related to technology has been understood as skilled job. However, from inventing the first agricultural tools to programming the first electronic general-purpose computer, ENIAC, women's achievements in technology have been left unacknowledged and taken as part of their natural skill set such as knitting or sewing, which was then claimed not to be much of a skill.[308]

Sandra Fredman compiled many court decisions rejecting women access to higher education, to public representation, as well as those approving of discriminatory hiring practices, all of which cite to women's delicate nature that belongs to the private sphere where they should be focusing on the care of children and the organization of the household. Here, law claims to bring out a neutral fact

---

[306] Walby, *Theorizing Patriarchy*.

[307] Margareta Kreimer, "Labour Market Segregation and the Gender-Based Division of Labour," *European Journal of Women's Studies* 11, no. 2 (2004).

[308] See <https://www.history.com/news/coding-used-to-be-a-womans-job-so-it-was-paid-less-and-undervalued> and <https://spectrum.ieee.org/untold-history-of-ai-invisible-woman-programmed-americas-first-electronic-computer> Accessed on 10 September 2022.

for all women which is based on traditional assumptions and reflects blindness to all non-bourgeois women's experiences.[309] On the other hand, within the working class, male workers deliberately worked under trade unions in order to keep women in domestic duties, which is another reason for women's subordination to both places, one reinforcing the other.[310] Exclusion also came in the form of protective legislation, which restricted women from working during pregnancy and breastfeeding. The way these laws were designed was a translation of the "male breadwinner - women caregiver" model reflecting the traditional gender division of workforce into the labour market. This weaker position of women in the labour market reduced their bargaining power and led them to concentrate in lower wage industries, which also has been used as a strategy for justifying to pay women lower wages than men which would otherwise be contrary to equality principles if they were doing the same job, working at the same factory base. Feminist research has demonstrated that women's exclusion from technology was also a consequence of the male domination of skilled trades which developed during the Industrial Revolution.[311] Today, gender division of labour is still connected to association of technology, masculinity and the notion of what constitutes as skilled work.

Another phenomenon of the gendered labour market is occupational segregation divided into two categories. Horizontal segregation refers to the tendency of women and men to be employed in different occupations, while vertical segregation refers to the tendency for women and men to be employed in different positions within the same occupation or company[312] (a.k.a. the glass ceiling). Today we can observe this trend in companies where human resources

---

[309] Fredman, *Women and the Law*.

[310] Heidi Hartmann, "Capitalism, Patriarchy, and Job Segregation by Sex," *Signs* 1, no. 3 (1976).

[311] Judy Wajcman, "Feminist Theories of Technology," *Cambridge Journal of Economics* 34, no. 1 (2009).

[312] Helina Melkas and Richard Anker, "Occupational Segregation by Sex in Nordic Countries: An Empirical Investigation," *Int'l Lab. Rev.* 136 (1997): 342.

and call-centres are mostly occupied by female employees, whereas technical and sales departments are mostly male dominant. Unsurprisingly, the latter are better routes to top managerial positions. Moreover, the requirements of these positions such as an uninterrupted career, being available every day on a full-time basis and anytime for travels do not fit with family responsibilities. Occupational segregation is also responsible for the tendency called feminization of occupations, that the entry of women leading to a downgrade of the pay and status of an occupation. [313]

1.3.2 The Hiring Funnel

Employers, in growing numbers, use machine learning techniques which detect patterns in the data available to them (training data) in order to build "hiring models" with predictive outcomes.[314] The use of predictive analysis and automated decisions is not only appealing to employers whose will is to achieve inclusion and diversity. Hiring processes are usually time consuming and the more time they take the more company resources are allocated to these tasks. In addition, predictive analysis is thought to help in making better hiring decisions which may result in hiring better fitting team members, preventing theft, and increasing employee tenure.[315]

---

[313] Harriet Bradley et al., *Myths at Work* (Polity, 2000). See also Diane Elson, "Gender Justice, Human Rights, and Neo-Liberal Economic Policies," in *Gender Justice, Development, and Rights*, ed. Maxine Molyneux and Shahra Razavi (Oxford University Press, 2002). For a reverse example on how men's entry into an occupation led its value to rise, see. The First 1940s Coders Were Women–So How Did Tech Bros Take Over?
<https://www.history.com/news/coding-used-to-be-a-womans-job-so-it-was-paid-less-and-undervalued>

[314] Hilton, Cisco, PepsiCo, Amazon, Ikea to name a few. See Bogen and Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias," 3.

[315] Ibid., 6.

It is important to consider that there is never a single decision to hire or not to hire an applicant, but a cumulative series of small decisions taken at different steps like announcing an open position, evaluating qualifications or the short listing for the interview process. There are various vendors in the market specializing on these different steps that an employer may choose for different steps of the hiring process. Looking into this variety may help understand how automation in hiring works. The "hiring funnel"[316] is commonly used to show how the number of candidates is reduced through sourcing, screening, interviewing and the selection stages.[317]

During the **sourcing stage**, an applicant pool is created. Some vendors[318] offer solutions for analysing the language of job descriptions to make them more appealing for certain applicant groups, including a "gender tone meter" which would show whether the text is off-putting for one gender. Employers also exploit all available ad technology that enables behavioural targeting and microtargeting specific demographic groups. Another process is matching, which is exercised through so-called recommender systems.[319] We will get into the details of how this stage produces bias and discriminatory effects more in detail in Section 1.3.4.

At the **screening stage,** applicants are assessed in order to determine whether they meet required qualifications or top performer traits. While candidates deemed to be unqualified are rejected, the strongest applicants move to the next stage. Chatbots using natural language processing (NLP) would engage in conversations with applicants at this stage to decide whether they are qualified for the job based on the employer's predefined criteria. There are other assessment tools that offer self-assessment surveys, games and interactive activities to candidates, in order to

[316] See www.upturn.org/reports/2018/hiring-algorithms/

[317] Bogen and Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias."

[318] textio.com/products/recruiting

[319] Bogen and Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias."

explore their personality traits and to compare them to the desired traits of the employer, usually based on top performing current or past employees. Here it is very possible all the already existing bias that may reflect in the current employee pool and performance indicators will be informing the assessment tool. These machine learned assessment tools may not produce causal relations at all and many of them are informed from psychology and behavioural science research, which itself have been found bias and unrepresentative many times.[320]

At the **interviewing stage**, usually video interviews are analysed through various technologies such as speech recognition, facial and vocal analysis, in order to predict future job performance of applicants based on the employer's performance metrics derived again from current successful employees. This reoccurring method of taking current employees as example is to be critical about when a company is aiming at change. For example, facial impressions and their meanings may differ across different cultures. Smiling is seen intelligent in Germany and China, but unintelligent in Iran.[321] In this case, if the model was trained on the data acquired from the top performers of a German company with mostly German employees, it may find out a pattern of smiling. In the beginning this might look like a neutral criterion, however, an applicant from Iran might hold herself from smiling during the interview with fear of appearing unintelligent and therefore get a negative score. Also, in many cultures having eye-contact with the opposite sex especially for women can be seen as a negative attribute and an algorithm taking eye-contact as a positive attribute for engagement and confidence may discriminate against women from non-western cultures. Alongside with the facial recognition systems, these interviewing tools make use of speech recognition software, which carry bias risk the same way it was explained above for NLP.

---

[320] Ibid.; Criado-Perez, *Invisible Women : Data Bias in a World Designed for Men*

[321] Kuba Krys et al., "Do Only Fools Smile at Strangers? Cultural Differences in Social Perception of Intelligence of Smiling Individuals," Journal of Cross-Cultural Psychology 45, no. 2 (2014).

The **selection stage** is directed to shortlisted applicants where background checks would be run. Some of the new tools offered in this area predict potential childcare providers or a candidate's risk of engaging in toxic behaviours such as sexual harassment by analysing publicly available content produced by the candidate in social media or blog posts. Some hiring tools at this stage help employers to predict the likelihood of a candidate to accept a specific offer. By adjusting salary, bonus and other benefits; they can tailor an offer to a level that it cannot be refused. In this case for example, the model can learn from the existing gender pay gap and successfully perpetuate it by offering lower salaries to women.[322]

1.3.3 Fairness in e-hiring

There are various technical challenges with possible negative societal outcomes that may arise from automating the hiring process, stages of which we have introduced above. First of all, most of the time the outcome variable will not be measurable. Second, data that is required to train and test the models are non-existent, insufficient or biased. Third, accuracy of the model is prone to be the result of self-fulfilling prophecies.

As to the first challenge, measuring what makes a good employee or a best job performance, or best cultural fit[323] are more complex issues than identifying spam emails. Good is subjective and not measurable, and good employee is not evident.[324] As We discussed above, it requires data scientists to choose more measurable proxy variables such as higher sales, shorter production time, or longer

---

[322] Bogen and Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias."

[323] Natalie Sheard, "Employment Discrimination by Algorithm: Can Anyone Be Held Accountable?," *UNIVERSITY OF NEW SOUTH WALES LAW JOURNAL* 45, no. 2 (2022): 627.

[324] Barocas and Selbst, "Big Data's Disparate Impact."; Kim, "Data-Driven Discrimination at Work."

tenure, or some numerical expression defining them (monthly sales to be over *x*) which are themselves may not be objective or fair to all applicants. Furthermore, any of the selected criteria will be a very limited way of defining good.[325]

On the other hand, letting the algorithm to find correlations to predict the outcome variable may not be desirable in the employment context. As Kim explains, unlike social scientists who have a hypothesis and looking for causal relations in the statistical data to prove it, thus making informed choices about which variables to include as the results will be very sensitive to those choices, data mining relies on massive data sets in capturing correlations without being inspired from a theory to justify the results.[326] For instance, one resume screening company's algorithm discovered having the name Jared and playing high school lacrosse to be strong indicators of success.[327] Decisions based on such correlations are hard to explain and harder to justify, thus cannot be easily claimed to be fair even if these success indicators were not acting as proxies to being man or white. Thus, in employment context, it is suggested to opt for developing causal models,[328] not only because of possible legal consequences but also because they may lead to loosing employee trust and motivation.[329]

Second challenge is about the data that is needed to develop useful models. When an employer decides to build a model to discover good performer traits in order to use it in future hiring processes, they may face a couple of data problems. First of all, there can never be enough data that records everything that is done by a good performer at a workplace, job definitions are usually broad, and many tasks

---

[325] Barocas and Selbst, "Big Data's Disparate Impact."

[326] Kim, "Data-Driven Discrimination at Work."

[327] Bogen and Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias," 8.

[328] Joshua R. Loftus et al., "Causal Reasoning for Algorithmic Fairness," *ArXiv* abs/1805.05859 (2018).

[329] Tambe, Cappelli, and Yakubovich, "Artificial Intelligence in Human Resources Management: Challenges and a Path Forward," 37.

are performed in teams. Moreover, recorded evaluations may be biased as a result of general stereotyping and cognitive biases of team members or supervisors that affect evaluations.[330] Also, choosing more quantifiable measures for performance in order to be objective, over softer skills such as leadership or collaboration may introduce further biases.[331] Furthermore, important events such as hiring and firing are rare to produce sufficient data for training a model. Also, most of the time companies do not retain detailed data on candidates who were not hired. So, the machine learning algorithm would be trained on the limited dataset of previous hires which is most likely already reflecting historical bias and therefore misrepresentative of the future applicant pools. In one known example, Amazon's hiring algorithm discriminated against female applicants. It was trained on job performance of previous employees which were mostly white men, therefore they constituted also the best performers. The algorithm not only gave higher result to white men, but also ruled out candidates based on attributes associated with being woman, including ruling out those who took courses in women's studies.[332]

Moreover, in most companies data that may inform ML algorithms are not kept in one place to be analysed and discovered for new patterns. First, it is very common to use tools from different vendors for different tasks which are not compatible with each other. Second, politics within companies may give rise to departments refusing to share their data with other departments.[333] All in all, small data problem makes it even more attractive to work with vendors who would be able to train algorithms on big data as they have access to data from many employers. However, then employers may not be able to assess how similar the training and test data used in development of these models to their employees and

[330] Ibid.

[331] Kim, "Data-Driven Discrimination at Work," 876.

[332] Tambe, Cappelli, and Yakubovich, "Artificial Intelligence in Human Resources Management: Challenges and a Path Forward," 16.

[333] Ibid.

companies, thus whether the model will perform with similar accuracy. Such problems occur in different solutions offered by vendors to automate different stages of hiring process as explained above.

Another point is, because the employer would not be able to know how the rejected applicants would have performed and those employed with the help of an algorithm might perform just good enough, discriminating models may turn into self-fulfilling prophecies.[334]

On this account, the limitations of the very technologies that are used to collect data on a candidate in order to produce a decision of qualified/not qualified can itself be problematic. In the screening stage some companies make use of Chat Bots which use national language processing (NLP) in order to automate their time consuming initial interview process. NLP trained on real life data has shown bias along the lines of gender and race, as well as having a hard time coping with dialects and those who speak English as a second language.[335]

There is also a wide range of companies who provide diversity and inclusion solutions to employers in Europe. Despite providing services to a rather global market, most of the vendors are US based, meaning they are coming from a completely different antidiscrimination jurisdiction. The US law for example requires “the highest-passing and lowest-passing demographic categories” to follow a 4/5th rule, which is easily machine translatable and doesn’t exist in the EU law. As discussed in the recent scholarly work, EU anti-discrimination law is rather contextual.[336] Even if, general fairness is sought, as we have already

---

[334] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

[335] Rachel Griffin, "Tackling Discrimination in Targeted Advertising," Verfassungsblog, https://verfassungsblog.de/targeted-ad/.

[336] Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai ".

discussed above different fairness metrics would lead to different results but may not end up providing meaningful diversity and inclusion. Many of the offers in the market are also limited to single-axis understanding of identification which disregards intersectional discrimination.[337] Also, the very fact that these data-driven solutions have to be informed by the past in order to perform in the future, make their progressiveness questionable as the populations to which training and test data are related may be completely different than the population the employer will deploy the model upon. Overall, if the models reduce the expenses of hiring, perform at an acceptable level of accuracy, and do not significantly violate the law, there is not much reason to think vendors and the employers would seek for further mitigation of bias.[338]

Even though there is a growing market of Diversity & Inclusion Technologies which promise to eliminate human biases, use of machine learned algorithms run the risk of codifying inequalities while providing a veil of objectivity. When predictions, numerical scores, or rankings are presented as precise and objective, it may affect the decision of human decision makers, also known as automation bias. Courts also may be influenced by this objective presentation of business decisions[339] and structural inequalities might be left unchallenged.

1.3.4 The Role of Platforms and Algorithmic Job Ads Delivery

Online platforms have become the new public sphere, therefore they have also become places where people may passively encounter or seek for new opportunities. In any case, there is a difference from the offline world when one

---

[337] Sánchez-Monedero, Dencik, and Edwards, "What Does It Mean to 'Solve' the Problem of Discrimination in Hiring? Social, Technical and Legal Perspectives from the Uk on Automated Hiring Systems."

[338] Kim, "Data-Driven Discrimination at Work."

[339] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

may encounter a poster which advertises a new position or go to a networking event. An online platform is a mediated environment in which one cannot simply walk like a flaneur/flaneuse on the street or read through their newspaper on stick in a coffee house and bump into seemingly irrelevant things that catch their interest. Here, everything they see is optimized for relevance to them based on their past interaction with this environment as well as other data that is used to build their profile. Eren et. al pointed out the risks of gender stereotyping and algorithmic discrimination which may arise therefrom in their study,[340] which also suggested some solutions to this problem.

As we discussed earlier, the dominant business model of online services is behavioural advertising, which is based on data-driven technologies of surveillance mostly conducted by cookies and their own algorithmic profiling methods in the case of platforms. Probably the most sublime way of discrimination stems from job ads delivery, for which platforms make decisions whether to show an opportunity to a user or not based on the inferred data about them.[341]

Many employers rely on social media and job board platforms, such as Facebook, Instagram, LinkedIn, Indeed, ZipRecruiter, Xing, Monster, and CareerBuilder, in order to reach suitable candidates. These platforms act as recommender systems making use of ADM to recommend matching candidates to recruiters and employers to candidates. Prior choices and behaviour of other similar users are some of the information points that are taken into consideration in order to run predictive analytics (LinkedIn and Xing). They examine thousands of profiles that would fit a job vacancy in order to recommend the fitting candidate

[340] Eren et al., "Increasing Fairness in Targeted Advertising: The Risk of Gender Stereotyping by Job Ad Algorithms.."

[341] Ibid.

to the recruiter.[342] General search engines such as Google and individual websites and mobile apps also allow employers to display job ads, targeting works the same way it is done through ad networks and real-time bidding as we described above in Section 1.2.3.

Both on social media platforms and general web advertising, job ads are being targeted to profiles that are inferred from users' behaviour and other available information, allowing recruiters to include or exclude based on desired or not desired shared attributes.[343] There are many examples where female and male jobseekers were delivered these advertisements differently, such as male users receiving more ads promoting high-paying jobs or female users being less likely to be shown an ad for STEM (science, technology, engineering and math) jobs in a well-known example.[344]

As mentioned above, like in housing ads, also in job ads Facebook lets employers to target or exclude users based on numerous attributes including directly based on protected grounds or proxies for them. Facebook, Google, LinkedIn all also provide targeting based on predicted similarity. For instance lookalike tool (changed to Special Ad Audience tool) of Facebook allows employers to upload their current employee pool, by making use of Custom Audience made from customer list option,[345] to reach at similar people. Thus, an

---

[342] Alina Köchling and Marius Claus Wehner, "Discriminated by an Algorithm: A Systematic Review of Discrimination and Fairness by Algorithmic Decision-Making in the Context of Hr Recruitment and Hr Development," *Business Research* 13, no. 3 (2020).

[343] Bogen and Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias."

[344] Lambrecht and Tucker, "Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of Stem Career Ads."

[345] Meta Business Help Center- About Customer List Custom Audiences <https://www.facebook.com/business/help/341425252616329?id=2469097953376494> Accessed on 10 October 2022. In order to improve match rate Facebook suggests providing Email, Phone Number, Facebook App User ID, Facebook Page User ID, First Name, Last Name.

employer may target people between the ages of 18-35 or in the category young & hip or by uploading their current employees who are all younger than 35 and reach at the similar result of excluding older users. Similarly it applies along the lines of gender and race.[346]

Job matching platforms and recommender systems may rely on content-based filtering or collaborative filtering. The first assesses what a user seems to be interested in by looking into their online activities, while the latter predicts what the user is likely to be interested in based on the inferred interests of people similar to the user. Both of these filtering processes are capable of keeping women, for example, away from managerial positions. First, biased clicking behaviour of recruiters in favour of male candidates may cause algorithms to learn recommending male job seekers not only to the recruiter with the biased behaviour but also to other recruiters who are seeking for candidates for similar roles. Likewise, the behaviour of some woman job seekers such as refraining from applying to more senior jobs due to self-doubt despite having many years of experience may also lead to similar effects that not only she would be shown more junior jobs but other woman job seekers with similar job experience who are eager to lend a senior position may be shown junior positions because similar women click more on junior positions.[347] Thus intermediated bias is perpetuated in collaboration with internalized bias, overall causing an amplified bias against women in recommender systems.

---

[346] Pauline T. Kim and Sharion Scott, "Discrimination in Online Employment Recruiting," *Saint Louis University law journal* 63 (2018).

[347] Bogen and Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias," 21-25.

CHAPTER 2

2 THE RELATIONSHIP OF INTERNATIONAL HUMAN RIGHTS FRAMEWORK AND EUROPEAN ANTI-DISCRIMINATION LAW TO ADM

In this chapter, we will first establish the relationship between right to privacy and non-discrimination, with other rights they protect, especially in relation to WMG. Then we will discuss the applicability and shortcomings of the EU antidiscrimination law in addressing algorithmic discrimination.

2.1 INTERNATIONAL HUMAN RIGHTS FRAMEWORK

Within the context of this thesis, substantive rights and freedoms found in international human rights frameworks which are the most relevant to the protection of minorities and marginalized individuals from the adverse effects of ADM are "right to privacy, identity and autonomy," and "right to equality and non-discrimination." Here, the interplay of these rights and their connection to ADM processes will be explored within the European context and their connection to the relevant EU body of secondary law from a feminist perspective. We will also briefly touch on the "freedom of expression" and "freedom of association and assembly," as exercising them online may reveal personal data on protected attributes and form a basis for discrimination when used by algorithms in job advertisement and human resources, while knowing that they may face discriminatory treatment as a result of exercising these rights may deter them from continuing to do so which may cause harm to individual and group interests, as well as to democracy and society in general.

2.1.1 Dearest Rights in a Democratic Society

The European Convention on Human Rights was born in the immediate post-

War era into the duty of preventing totalitarianism's rebirth. Therefore, a special attention is given to maintaining a democratic society which provides the best place for fundamental freedoms to flourish. In turn, ECHR permits fundamental freedoms to be limited only if it is necessary in a democratic society.

Reference to democracy reoccurs through the text of the ECHR and the jurisprudence of the ECtHR, in the CFREU and throughout various legislation of the EU and the case law of CJEU, especially when determining the limits of fundamental rights and freedoms. All the EU legislation and case law refers to ECHR as it is among the general principles of the EU law.[348] Thus it is important to first understand what does democracy mean under this framework,[349] as it is otherwise argued to be an essentially contested concept.[350]

ECtHR has given away hints about its understanding of democracy which is about striking a balance that ensures fair treatment of minorities and prevents the abuse of dominant positions,[351] and pluralism, tolerance and broadmindedness are essential for the existence of a democratic society.[352] However the ECtHR has also been criticized for not really evaluating what democracy is and taking assumptions of liberal political theory about democracy as granted.[353]

On this account, two major lines of criticism can be identified. First, liberal democratic practice does not really concern participation in the decision-making

---

[348] TEU Art. 6(3).

[349] Susan Marks, "The European Convention on Human Rights and Its 'Democratic Society'," British Yearbook of International Law 66, no. 1 (1996).

[350] Gallie, "Essentially Contested Concepts."

[351] Young, James and Webster v. The United Kingdom (Application no. 7601/76; 7806/77) 13 August 1981 https://hudoc.echr.coe.int/eng?i=001-57604

[352] Handyside v. The United Kingdom (Application no. 5493/72) 7 December 1976 https://hudoc.echr.coe.int/eng?i=001-57499

[353] Marks, "The European Convention on Human Rights and Its 'Democratic Society'."

beyond fair participation in free elections.[354] As it is stated by the ECtHR, democracy is not only about free elections and fair participation in them. If so, shortly after women won their suffrage rights, feminism would no longer be needed, women's participation in every office would be proportional to their numbers in the population, we would not be talking about gender pay gap, glass ceiling, gender mainstreaming or algorithmic discrimination against women today. Similarly, in the US, how can we explain the events of granting black people full citizenship rights including to vote and stand in elections, and election of Barack Hussein Obama as the first black president being 140 years apart? Why was there a need for civil rights movement of the 1960s or the Black Lives Matter (BLM) movement beginning in 2013 while having a black president in the White House? Second, social and economic inequalities have been long compromised promises of liberal democratic practice to treat citizens as free and equal. As systematic asymmetries of power and resources along the lines of gender, race, class and so on are continuously reproduced in the private spheres of employment, media, market, family, sexual and other personal relations; people are prevented from exercising their formal civil and political rights such as right to vote and stand for public office, freedom of expression, and freedom of association and assembly in the public sphere. Thus feminist critiques of democracy have called for the democratization of activities that fall in the so-called private sphere in the liberal theory.[355]

Thus it appears, within ECHR jurisprudence, democracy as a concept is tasked with (i) justifying, shaping and determining priorities among the guaranteed rights and freedoms of the ECHR; (ii) distinguishing legitimacy or illegitimacy of restriction of specific rights, particularly those between articles 8-11; (iii) grounding certain rights and institutions directly related to elections.[356]

---

[354] Ibid.

[355] Ibid.; ibid.

[356] Ibid., 231.

First, the interplay of group interests, public debate, media's role, the freedom of expression, association, and assembly are directly linked to democracy;[357] thus democracy is identified with "pluralism, tolerance and broadmindedness."[358]

According to the ECtHR, "democracy does not simply mean views of the majority must always prevail," instead democracy is about striking a balance that ensures fair treatment of minorities and prevents the abuse of dominant positions.[359] Alongside with ensuring fair treatment, freedom of political debate[360] is deemed essential to democracy by the ECtHR. In this regard, freedom of expression have been named closely attached to democracy, most openly in the landmark *Handyside* case. According to the ECtHR, freedom of expression lies at the core of a democratic society, serving as a prerequisite for its progress and personal development of the individuals living in it. Freedom of expression specifically "protects information or ideas that offend, shock or disturb" in order to maintain "pluralism, tolerance and broadmindedness without which there is no democratic society".[361] The ECtHR also regarded that people who are assembling in the streets or other public places exercise their freedom of expression and any restrictions against such acts must ensure individuals are not discouraged from communicating their opinions to the public.[362] Thus, ECtHR draw a strong connection between the Article 10 and 11 of the ECHR and maintaining of democracy. In *Klass v. Germany* regarding surveillance against espionage and terrorism, the ECtHR draw a direct line between right to privacy (Art. 8 ECHR)

---

[357] Ibid.

[358] Handyside v. The United Kingdom (Application no. 5493/72) 7 December 1976

[359] Young, James and Webster v. The United Kingdom (Application no. 7601/76; 7806/77) 13 August 1981 https://hudoc.echr.coe.int/eng?i=001-57604

[360] Lingens v. Austria (Application No . 9815/82) 8 July 1986 https://hudoc.echr.coe.int/eng?i=001-57523

[361] Handyside v. The United Kingdom (Application no. 5493/72) 7 December 1976

[362] Ezelin v. France (Application no. 11800/85) 26 April 1991 https://hudoc.echr.coe.int/eng?i=001-57675

and the maintenance of democracy which could be destroyed in the name of protecting it if adequate safeguards are not provided.[363]

However, democracy is primarily linked to government which is assumed to be properly done by being allocated to professionals. This reduces the citizen interest in democracy to be protected from the misuses of power by the State and accountability in public affairs.[364]

Secondly, mentioned in the second paragraph of rights between article 8-11, necessity in a democratic society becomes the defining ground of proportionality in reconciling the tension between community interests and individual rights or between the rights of different individuals. ECHR jurisprudence follows the public/private distinction that is prominent in liberal political thinking, for example by applying a wider margin of appreciation in cases concerning freedom of expression in commercial context as opposed to those in public domain.[365] It is often that right to privacy and freedom of expression end up in a stand-off in which ECtHR has been more likely to take sides with the freedom of expression as it is understood as more contributing to the community as opposed to the ECtHR perception of right to privacy as more individualistic. This is also related to the public/private distinction that relationships between private actors including economic relations are understood to be non-political so that not too essential for the democracy.[366]

Thirdly, ECHR jurisprudence have been also criticized for understanding Protocol 1 Article 3 as the state's duty to ensure all citizens are treated equally in

---

[363]Klass and Others v. Germany (Application no. 5029/71) 6 September 1978 https://hudoc.echr.coe.int/eng?i=001-57510

[364] Marks, "The European Convention on Human Rights and Its 'Democratic Society'," 232.

[365] Ibid., 233.

[366] Ibid.

exercising their right to vote while not considering the conditions that are necessary for such equality to be effective. ECtHR seems to be mostly blind to asymmetries of power in the realms of civil society which affect efficacy of formal freedom and equality in every aspect of life including participation in formal political institutions.[367] Indeed, ECHR jurisprudence does not really go beyond formal equality to the realm of social and economic rights. This is done in Europe by the Charter of Fundamental Rights of the European Union.[368]

2.1.2 Privacy and Equality in Collaboration

Sandra Wachter argued that privacy is a necessary precondition for the realisation of rights between Article 9-11 of the ECHR and Article 3 of the 1st Protocol.[369] Those are the rights that are acknowledged to be in a direct relation with democracy both in the text of the ECHR and the case law of ECtHR.[370]

According to Wachter, the right to privacy is related to these rights in three different layers. First of all, privacy is about freely developing a personality and choosing a good life for one's self, which is the basis for pluralism without which there is no democracy. Secondly, further development of ideas and life choices depends on being able to express them in interaction with others. Here privacy as self-determination comes in play, as individuals need to choose which part of their opinions or life choices they want to share with which audience in order to feel comfortable in sharing them and putting them in practice. It is important to know expressing one's ideas, interacting with like-minded people or seeking for

---

[367] Ibid.

[368] Under CFREU Title IV.

[369] Wachter, "Privacy: Primus Inter Pares — Privacy as a Precondition for Self-Development, Personal Fulfilment and the Free Enjoyment of Fundamental Human Rights".

[370] Handyside v. The United Kingdom (Application no. 5493/72) 7 December 1976 https://hudoc.echr.coe.int/eng?i=001-57499, Ezelin v. France (Application no. 11800/85) 26 April 1991 https://hudoc.echr.coe.int/eng?i=001-57675

information would not cause negative ramifications. Otherwise chilling effects may occur and this would limit the free market of ideas and therefore harm the continuous development of a democratic society.[371] Finally, privacy constitutes a prerequisite for those rights which may reveal sensitive information about the individual who exercises them such as freedom of thought, conscious and religion, the right to assembly and associations, and the right to free elections. All of these layers are connected with informational self-determination and the prohibition of discrimination (as in ECHR Art. 14 and Art. 1 12th P.).[372]

Wachter concludes by declaring privacy to have an elevated position among other fundamental rights and freedoms. She bases her arguments mostly on the

---

[371] On the chilling effects caused by interference with right to privacy, see Laurent Pech, The Concept of Chilling Effect: Its Untapped Potential to Better Protect Democracy, the Rule of Law, and Fundamental Rights in the Eu, (Open Society European Policy Institute, 2021), https://www.opensocietyfoundations.org/publications/the-concept-of-chilling-effect#publications_download; Jonathon Penney, "Understanding Chilling Effects," *Minnesota Law Review* 106 (2021); Karen Levy, "Chilling Effects and Unequal Subjects: A Response to Jonathon Penney's Understanding Chilling Effects," ibid. (2022); Jonathon W. Penney, "Chilling Effects: Online Surveillance and Wikipedia Use," *Berkeley Technology Law Journal* 31, no. 1 (2016); "Internet Surveillance, Regulation, and Chilling Effects Online: A Comparative Case Study," *Internet Policy Review* 6, no. 2 (2017), https://policyreview.info/articles/analysis/internet-surveillance-regulation-and-chilling-effects-online-comparative-case; Moritz Büchi, Noemi Festic, and Michael Latzer, "The Chilling Effects of Digital Dataveillance: A Theoretical Model and an Empirical Research Agenda," *Big Data & Society* 9, no. 1 (2022). Such chilling effects may occur even when personal data or any information is not processed, see Dara Hallinan, "Opinions · Data Protection without Data: Could Data Protection Law Apply without Personal Data Being Processed?," *European Data Protection Law Review* 5, no. 3 (2019). In *Digital Rights Ireland*, Villalón AG acknowledged the effect of mass surveillance on freedom of expression and assembly, however opined that "that effect would be merely a collateral consequence of interference with the right to privacy" (See, C-293/12 and C-594/12 Joined Cases [2014] *Digital Rights Ireland Ltd* Opinion of Villalón AG, ECLI:EU:C:2013:845, para 52.)

[372] Wachter, "Privacy: Primus Inter Pares — Privacy as a Precondition for Self-Development, Personal Fulfilment and the Free Enjoyment of Fundamental Human Rights".

value of privacy as informational self-determination, which acts as a guardian against discrimination and allows individuals to have control over their personal information that is crucial for having room for self-development. Again she continuously refers to people's fear of being discriminated against as their thoughts, beliefs or life choices would offend, shock or disturb the society. Therefore, it appears that privacy provides an *ex ante* protection for the democracy by protecting individuals in expressing themselves and in interacting with others without revealing information about themselves more than they wish, while non-discrimination provides *ex post* protection to democracy by protecting individuals against being treated unfairly because of what such expressions and associations have revealed. For now, this formulation makes a case for collaboration of right to privacy and non-discrimination in protecting exercise of other rights which are instrumental in building and maintaining a democratic society which is based on equal participation of the people in democratic governance. On the other hand privacy,[373] equality and even democracy are all essentially contested concepts.[374]

This is to say that both equality and privacy have different multi-faceted and open-textured meanings.[375] This makes them hard to pin down under one dimensional definition that would have been easier to automate or standardize, but also keeps them flexible in their broad relation with all the other fundamental rights and freedoms they protect. According to Berlin's concepts of positive and

---

[373] Mulligan, Koopman, and Doty, "Privacy Is an Essentially Contested Concept: A Multi-Dimensional Analytic for Mapping Privacy." As to the vagueness of privacy as a concept, Solove suggests Ludwig Wittgenstein's argument about some concepts could only be understood as family resemblances. Privacy is one of these concepts that is not one thing but a plurality of many distinct yet related things. See Daniel J. Solove, "The Meaning and Value of Privacy," in *Social Dimensions of Privacy: Interdisciplinary Perspectives*, ed. Beate Roessler and Dorota Mokrosinska (Cambridge: Cambridge University Press, 2015), 75.

[374] Gallie, "Essentially Contested Concepts."

[375] Mulligan, Koopman, and Doty, "Privacy Is an Essentially Contested Concept: A Multi-Dimensional Analytic for Mapping Privacy."

negative freedom,[376] substantive rights such as freedoms of expression, association and assembly reflect the positive notion of "freedom to" take actions. On the other hand, regulatory rights such as privacy and non-discrimination reflect the negative notion of "freedom from" others' actions that may infringe upon a person's positive freedoms, thus they regulate the actions of others to protect the substantive rights and freedoms of their subjects of protection.[377]

For example, Solove points at the fact that ECtHR has chosen to define privacy based on the reasonable expectation test.[378] However, reliance on the reasonable expectation of privacy fails to address the privacy paradox.[379] According to this paradox there is a mismatch of people's online information sharing behaviour and their intentions for protecting their privacy. Solove suggests two important reasons for such information sharing behaviour. First, many times there isn't much choice but to share information with businesses.[380] Second, everyone does not know how

[376] See Isaiah Sir Berlin, "Two Concepts of Liberty" (1958). https://jmaggio.typepad.com/no_call_me_jay/files/two_concepts_of_liberty.pdf

[377] Raphaël Gellert et al., "A Comparative Analysis of Anti-Discrimination and Data Protection Legislations," in *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases*, ed. Bart Custers, et al. (Berlin, Heidelberg: Springer Berlin Heidelberg, 2013).

[378] Solove, "The Meaning and Value of Privacy," 75.

[379] Susanne Barth and Menno D. T. de Jong, "The Privacy Paradox – Investigating Discrepancies between Expressed Privacy Concerns and Actual Online Behavior – a Systematic Literature Review," Telematics and Informatics 34, no. 7 (2017).

[380] Midas Nouwens et al., "Dark Patterns after the Gdpr: Scraping Consent Pop-Ups and Demonstrating Their Influence," in Proceedings of the 2020 Chi Conference on Human Factors in Computing Systems (Association for Computing Machinery, 2020); Luiza Jarovsky, "Dark Patterns in Personal Data Collection: Definition, Taxonomy and Lawfulness," (2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4048582; ibid.; Sara Suárez-Gonzalo, "Personal Data Are Political. A Feminist View on Privacy and Big Data," Recerca Revista de pensament i anàlisi 24 (2019).

this information can be and will be used.[381] Suárez-Gonzalo argues that both reasons are true as a result of the asymmetric power relation as data domination between individuals and private companies who exploit their data and control the data flows.[382] Another problem with defining privacy problems stem from associating privacy too much with individual sovereignty in the sense of an individual's right to shut out from collective life and focus on their personal development. It is in this sense gets undervalued in balancing tests of the courts against public interests.

According to Dewey,[383] there is no dichotomy between the individual rights and the society. The society shapes the individual as much as each individual shapes the society. Thus, contribution of an individual right to the welfare of the society must be acknowledged as the source of its value, otherwise most social interests would outweigh individual rights.[384] However, this depends on how one defines the society and the welfare of that society. Whose welfare we are talking about, how are the conflicted interests of groups of people to be balanced are among the questions to be answered. Privacy is contextual, thus what it means, how it can be protected, why it should be protected are essential questions that can change in time, according to the context and whose privacy we talk about. Privacy is an

---

[381] Solove, "The Meaning and Value of Privacy." As an example for data subjects being given inadequate information and choice see the Litigation Chamber decision on the merits 21/2022 of 2 February 2022 of Belgian Data Protection Authority (Case number: DOS-2019-01377) <https://edpb.europa.eu/system/files/2022-03/be_2022-02_decisionpublic_0.pdf> accessed 24 April 2022. On the other hand, even though data subjects are adequately informed and asked for their consent in compliance with the applicable data protection laws, data subjects may still lack the knowledge because of the complexity of the operation and the technology, as well as a result of lacking familiarity with the subject. Here, we can talk about an informational power asymmetry.

[382] Suárez-Gonzalo, "Personal Data Are Political. A Feminist View on Privacy and Big Data," 176.

[383] John Dewey, "Ethics," in The Collected Works of John Dewey V. 5; 1908, Ethics : The Middle Works, 1899-1924, ed. Jo Ann Boydston (Carbondale, United States: Southern Illinois University Press, 2008).

[384] Solove, "The Meaning and Value of Privacy."

essentially contested concept, that is why it is hard to pin down a definition of it.[385]

Demands of equality is similar to the demands of privacy in being contested in essence as a concept. As we described in section 1.1, privacy has been something that was debated over heavily by feminists. Privacy is desirable for women to have control over their bodies, to be able to dress as they wish, to be able to have an abortion, to be able to experience their sexuality freely. On the other hand, women are critical about the privacy in the sense of non-intervention of the State in the market and family relations where women's bodies are oppressed and labour is exploited.[386] However, how centred the theory of separate spheres of public and private has been in feminist theory, attracted criticism from black feminists as the whole theory was built on the experience of middle-class white women in the western world and did not correlate with the perceptions of black womanhood or manhood in the society, nor did with the experience of a black household, nor did how black women experienced state intervention and privacy of their home.[387]

Similarly, demands of justice and equality is hard to pin down once and for all under one definition and demand. If we follow the feminist route once more, women were first fighting against the exclusion so they first used the Aristotelian notion of procedural justice also known as formal equality. Here, the rule is to "treat like cases alike and unlike cases differently" in proportion to their differences. Women have instrumentalized this notion of equality to achieve access to the public sphere privileges traditionally enjoyed rather by men. However, this approach sets men as the norm. When men are taken as the norm to being human, then women start from an unequal position not only because they

---

[385] Mulligan, Koopman, and Doty, "Privacy Is an Essentially Contested Concept: A Multi-Dimensional Analytic for Mapping Privacy."

[386] MacKinnon, *Toward a Feminist Theory of the State*.

[387] Crenshaw, "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics."

are not men, but also due to past exclusion and an unfair setting for competing as women are traditionally burdened with primary responsibility for the young children. Thus, when men are taken as the norm, experiences that are solely part of women's lives such as pregnancy turn into a burden in equality claims. So feminist thought was trapped for a while in the dilemma of whether to demand equal rights to men or special rights that covers their different cases such as pregnancy and childcare which bears a risk of stigmatization of being different than the norm.[388] Also known as the merit principle, which seeks for individuals to be treated based on their merits rather than stereotypical assumptions derived from their particular attributes, it seems simple to achieve. However, deciding on which education, experience and other qualifications make one person identical to another, would require a value judgment to be that simple, rather than being an objective one. Another criticism is that it would not make any change in the condition of the WMG as it is satisfied when two individuals are treated equally badly or beneficially.[389] In other words, while demanding justice should be blind to the group memberships and treat people as individuals, formal equality not only ignores the reality of social power structures (e.g. usually problem starts from disadvantaged group members not having the equivalent access to means to acquire the similar merits) but also perpetuates and legitimizes them. Then comes the substantive distributive fairness, also known as "substantive equality" or "equality of outcome" which requires affirmative action. According to this thought, "equality does not mean treating people the same but requires treating them differently so as to achieve equal or equivalent effects."[390] In this notion of equality, it is acknowledged that blind justice is no justice, achieving fairness requires inquiring into identities, capacities and practices, thus group affiliations shall be acknowledged and individuals should be treated in proportion to the

[388] Fredman, *Women and the Law*.

[389] Ellis and Watson, Eu Anti-Discrimination Law.

[390] <http://www.ruleoflaw.ca/substantive-equality-some-people-are-more-equal-than-others/> Accessed 17 June 2022.

advantages/disadvantages or strengths/weaknesses of the group they are a member of.

As seen here, similar to ML literature in legal literature there are also different types of equality/fairness and such like it is the case for ML fairness metrics, formal equality and substantive equality exclude each other. Between the extremes of formal equality and substantive equality stands "equality of opportunity" which shares the idea of substantive equality that an unbending adherence to formal justice may perpetuate existing injustices and inequalities. However, it takes the individual approach of formal equality. It recognizes the fact that structural discrimination distorts the life chances of individuals because of their group memberships. Therefore, it aims at equalizing the starting point. Once the individuals enjoy equal opportunities in this sense, then they can be subjected to formal equality as they're now free from the burdens of structural injustices. Thus, it rejects the quotas and targets which aim at equal outcomes in the workforce.[391]

In relation to proxy discrimination, Prince and Schwarcz discussed how unexpected disclosure of protected traits in the form of proxy information as an intended or unintended violation of privacy is likely to undermine outcomes that are pursued with the right to non-discrimination. One of the goals of non-discrimination is to socialize individual risk, such as in the case of protection of pregnant women which involves distribution of employment-related costs of pregnancy in the society, so that they are not borne entirely by women. Second, limiting the effects of past discrimination is an important goal which targets the structural disadvantages resulting from historical subordination. Third goal is anti-stereotyping which protects non-conforming members of groups from being treated on the statistical basis of their group's characteristics. Non-discrimination principle ensures that past discrimination and stereotypes informed from it, not to block life chances for individuals. Due to both historical discrimination and

[391] Ellis and Watson, *Eu Anti-Discrimination Law*.

stereotyping creeping into ADM, WMG may have difficulty in beating hiring algorithms and landing a job, therefore ADM undermines these two goals of the non-discrimination principle as well. Finally, an important goal is preventing the chilling of valuable activities for a democratic society such as expressive or associational actions. News stories about how employers are able to predict someone's pregnancy, sexual orientation, or political opinion from their subtle social media behaviour such as inferred interests, likes, comments, groups and so on may lead to awareness and challenging these structures, but also it may lead members of disadvantaged groups to refrain from community building, seeking for help and activism that women and marginalized groups very much need. Without the preventive protection offered by non-discrimination laws, access to credit, insurance, healthcare can be negatively affected from not having a steady job due to algorithmic discrimination, this may curb access to university education, which may have been also the result of lack of access to others. Such vicious feedback loops in combination with the chilling effect that causes social isolation and silence would make these group interests to remain and reinforce the disadvantage loop.[392]

Hence, right to privacy and equal treatment are essential for democracy to prevail and they have both been under threat alongside with the rights and freedoms they are closely related to, as well as democracy itself, as informational technologies are becoming prominent in every aspect of individual and collective lives.[393] Custers et al. point out that privacy and non-discrimination are similar as their value is evaluated through their relation to other rights. Moreover, even though they may have derived from different legal sources, they share a similar

[392] Or "the vicious cycle of discrimination," see. Andrea Romei and Salvatore Ruggieri, "A Multidisciplinary Survey on Discrimination Analysis," *The Knowledge Engineering Review* 29, no. 5 (2014): 585.

[393] Gellert et al., "A Comparative Analysis of Anti-Discrimination and Data Protection Legislations."

remedy. While discrimination concerns focus on being distinguished based on particular attributes of individuals (such as gender, sexual orientation, racial or ethnic origin, disability, age, political opinion, religious belief etc), privacy concerns focus on processing of identifying attributes (such as name, phone number etc) in combination with sensitive attributes which mostly overlaps with protected attributes under antidiscrimination clauses. In both cases, individuals are advised not to share their personal data.[394]

2.1.3 A Separate Right to Data Protection

Digitization has resulted in private sphere growing to include personal data. This type of privacy is also called informational self-determination. According to Westin, "Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others."[395] Then such privacy aims at equipping the individual to have control over their personal information.[396] Article 8 CFREU and data protection laws reflect this understanding of privacy and thus they have been strongly focussed on control over one's personal data and personal autonomy. We will later discuss whether new tools and strategies introduced in the GDPR can be understood as a shift from this positioning.

Both among the scholarly work[397] and the case law of the ECtHR and the CJEU,[398] it is common to take privacy and data protection as compounds of each

---

[394] Bart Custers et al., Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases (Springer Publishing Company, Incorporated, 2013), 344.

[395] Alan Westin, Privacy and Freedom, (New York: Ig Publishing, 1967). 24.

[396] Custers et al., Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases, 343.

[397] Daniel J. Solove, "Conceptualizing Privacy," California Law Review 90, no. 4 (2002).

[398] Juliane Kokott and Christoph Sobotta, "The Distinction between Privacy and Data Protection in the Jurisprudence of the Cjeu and the Ecthr," International Data Privacy Law 3, no. 4 (2013).

other. However, they have different descents and they may serve different purposes that the other cannot. Mostly, the right to privacy applies to the matters of private life, personhood and identity of an individual even though there is no processing of personal data is involved. The right to data protection, on the other hand, applies whenever personal data is processed without being necessarily about private life at all.[399]

Gutwirth and De Hert identify privacy as an opacity tool, as opposed to data protection as a transparency tool. The opacity that right to privacy offers to the private sphere by guaranteeing the non-interference of the state has been a major critique of law by feminist legal scholars as an enabler of many forms of unequal power relations and women's oppression such as domestic violence. Transparency tools on the other hand tend to make the powerful accountable.[400] According to Nancy Fraser, in order to react to injustice, one has to be able to detect what is happening and have the means to identify it as unfair.[401] Domination becomes possible many times by disrupting the means of the oppressed to detect and identify injustices they experience.[402] Thus, being a black box is not always the "nature" of algorithmic systems, it is also intentional by those who dominate by controlling the data flows hidden behind the complexity of the technology, intellectual property rights and trade secrets.[403] While there has been a hot debate

[399] Plixavra Vogiatzoglou and Peggy Valcke, "Chapter 1: Two Decades of Article 8 Cfr: A Critical Exploration of the Fundamental Right to Personal Data Protection in Eu Law " in Research Handbook on Eu Data Protection Law, ed. Eleni Kosta, Ronald Leenes, and Irene Kamara (Edward Elgar Publishing, 2022). In *Digital Rights Ireland Ltd* The CJEU recognized that legal persons also enjoy CFREU Art. 7 right to privacy and Art.8 right to data protection (See. C-293/12 and C-594/12 Joined Cases [2014] *Digital Rights Ireland Ltd* ECLI:EU:C:2014:238). y

[400] Gutwirth and Hert, "Regulating Profiling in a Democratic Constitutional State."

[401] Nancy Fraser, "On Justice," New Left Review MAR/APR 2012, no. 74 (2012).

[402] Ibid.

[403] Pasquale, Black Box Society : The Secret Algorithms That Control Money and Information.

upon whether GDPR provides such transparency in relation to ADM,[404] transparency being a tool at all in exposing unfair uses of power has also been challenged.[405]

Max von Grafenstein argued that the informational power asymmetry is so central to the right to data protection, it actually defines the scope of this right.[406] Data protection took its place in the Art. 8 CFREU as a fundamental right serving next to the right to privacy (Art. 7) in protecting the autonomy of individuals in the sense of making autonomous decisions and taking autonomous actions in a

---

[404] Casey, Farhangi, and Vogl, "Rethinking Explainable Machines: The Gdpr's 'Right to Explanation' Debate and the Rise of Algorithmic Audits in Enterprise."; Goodman and Flaxman, "European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"."; Kaminski, "The Right to Explanation, Explained."; Gianclaudio Malgieri, "Automated Decision-Making in the Eu Member States: The Right to Explanation and Other 'Suitable Safeguards' for Algorithmic Decisions in the Eu National Legislations," Computer Law & Security Review, 2019 Forthcoming; "Right to Explanation and Algorithm Legibility in the Eu Member States Legislations."; Selbst and Powles, "Meaningful Information and the Right to Explanation."; Wachter, Mittelstadt, and Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation."; Wachter, Mittelstadt, and Russell, "Counterfactual Explanations without Opening the Black Box: Automated Decisions and the Gdpr."

[405] Edwards and Veale, "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For."; Ananny and Crawford, "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability."

[406] Maximilian von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part I," European Data Protection Law Review 6, no. 4 (2020). See also W. Steinmüller et al., "Grundfragen Des Datenschutzes: Gutachten Im Auftrag Des Bundesministeriums Des Innern [Bt-Drs. Vi/3816 Anlage 1]," (1971)., Jörg Pohle, "Datenschutz Und Technikgestaltung: Geschichte Und Theorie Des Datenschutzes Aus Informatischer Sicht Und Folgerungen Für Die Technikgestaltung" (Humboldt-Universität zu Berlin, 2018)., Gianclaudio Malgieri and Jędrzej Niklas, "Vulnerable Data Subjects," *Computer Law & Security Review* 37 (2020); O. Lynskey, *The Foundations of Eu Data Protection Law* (Oxford University Press, 2015), 211-26. For a critical respond to Lynskey's assessment on the essence and rationale of the right to data protection see F. Bieker, *The Right to Data Protection: Individual and Structural Dimensions of Data Protection in Eu Law* (T.M.C. Asser Press, 2022), 158-60.

free and democratic society which also serves as a basis for other aforementioned rights and freedoms to flourish.[407] In other words, Art. 8 CFREU protects the individuals against the risks to their other fundamental rights caused by processing of personal data.[408] This is because if the collection or processing of personal data reveals information about someone's private life, then right to private life in CFREU Art. 7 (and ECHR Art. 8) would be the first right to be affected and other rights may follow[409] as discussed above. Once such information is revealed to a person the damage is irreparable as the information cannot be deleted from the mind of this person, it cannot either be known whether this person passed the information onto others and whether they acted upon this information. For instance, in the case of discrimination against a job applicant based on information gained about them the right to equality would be affected. However, right to non-discrimination as well as other rights, which benefit from right to private life to be protected, may be at risk even when the collected data do not reveal any information about the private lives of any individuals.

In Nissenbaum's informational privacy as contextual integrity, protection of personal data is justified on appropriateness and the flow of data. First ground relates to that revealing information in one context may not be appropriate in another context, such as revealing an information at a doctor's office may not have

---

[407] Rouvroy and Poullet, "The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Inportance of Privacy for Democracy."

[408] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part I." According to von Grafenstein the scope of the interplay between the right to data protection and the other rights is rather vague and a referral to the concepts of risk regulation may help clarify as a risk based approach as suggested by the GDPR provides more effective protection regarding also the other rights.

[409] "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Ii," European Data Protection Law Review 7, no. 2 (2021): 17; Wachter, "Privacy: Primus Inter Pares — Privacy as a Precondition for Self-Development, Personal Fulfilment and the Free Enjoyment of Fundamental Human Rights".

the same effect as revealing it at a cocktail party, and the flow of information must respect what is appropriate to the context into which the information is transferred.[410] In this regard, similar to the Nissenbaum's moment when contextual integrity is violated,[411] von Grafenstein argues that the purpose specification aims at providing protection before it is too late.[412]

In this sense, the right to data protection as a separate right is indispensable to the protection of other fundamental rights as it protects them before their own scope applies.[413] However, von Grafenstein warns against the right to data protection becoming a super-fundamental right in this regard,[414] similar to the criticism of data protection law becoming the law of everything.[415] This would happen when the right to data protection overrides the diversity of all the other fundamental rights that would have applied in an analogous set-up by superseding all of them as more and more social interactions amongst persons take place in digitalised environments.[416]

In order to avoid such result, von Grafenstein turns to identifying the scope of the right to data protection. In this regard, German Constitutional Court gives weight to processing being automated, whereas the ECtHR considered an event

---

[410] Mulligan, Koopman, and Doty, "Privacy Is an Essentially Contested Concept: A Multi-Dimensional Analytic for Mapping Privacy."

[411] Helen Nissenbaum, "Privacy in Context - Technology, Policy, and the Integrity of Social Life" (2009).

[412] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Ii."

[413] Ibid., 20.

[414] Ibid.

[415] Purtova, "The Law of Everything. Broad Concept of Personal Data and Future of Eu Data Protection Law."

[416] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part I."

receding to the past[417] and personal information being systematically collected and stored for a long period of time[418] to make it part of the private life.[419] In *Jehovah's Witnesses*[420] the CJEU also opted for not considering the automated processing necessary for justifying legal protection and applied the DPD directly. According to von Grafenstein, even though this indicates to a lower threshold than the German criterion of automated processing, it still indicates to the informational power asymmetry[421] which is useful to be considered as the reason for data protection as explained by Jörg Pohle[422] by diving deep into the history of privacy and data protection.[423] On the other hand, when there is no informational power asymmetry stemming from the processing of personal data, then the protection of Art. 8 CFREU is irrelevant and the very right that is being intervened must take over.[424] According to von Grafenstein this was supposed to be the case in *Nowak,*[425] that instead of data protection law, the affected right to education should

[417] M.M. v. The United Kingdom (Application no. 24029/07) 13 November 2012

[418] Rotaru v. Romania (Application no. 28341/95) 4 May 2000

[419] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part I," 7; Kokott and Sobotta, "The Distinction between Privacy and Data Protection in the Jurisprudence of the Cjeu and the Ecthr."

[420] C-25/17 [2018] *Jehovan todistajat* ECLI:EU:C:2018:551

[421] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Ii," 8. argues that this is also the reason for the household exception in Art. 2(2)(c) GDPR. However, the related case law of the CJEU in C-101/01 [2003] *Lindqvist* ECLI:EU:C:2003:596; C-212/13 [2014] *Ryneš v Úřad pro ochranu osobních údajů* ECLI:EU:C:2014:2428; C-25/17 [2018] *Jehovan todistajat* ECLI:EU:C:2018:551 does not indicate to such consideration.

[422] Pohle, "Datenschutz Und Technikgestaltung: Geschichte Und Theorie Des Datenschutzes Aus Informatischer Sicht Und Folgerungen Für Die Technikgestaltung."

[423] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Ii."

[424] In this regard we will briefly discuss the shortcomings of anti-discrimination law in the face of algorithmic discrimination as so far identified in scholarly works.

[425] C-434/16 [2017] *Peter Nowak v Data Protection Commissioner* ECLI:EU:C:2017:994

have applied.[426]

Another question to be resolved is what Art.8 CFREU really protects. According to von Grafenstein, it protects individuals' autonomy against the risks caused by personal data processing, and the risk it may pose to the autonomous exercise of the data subject's all the other fundamental rights.[427] The right to data protection also protects from chilling effects to all the other rights that it protects. As the data subjects are like in a panoptical situation[428] in which they are not sure whether someone and if yes who is collecting, processing and sharing their personal data with whom, this constant uncertainty may cause modification of their behaviour[429] and refraining from exercising their rights which are the dearest to a democratic society. In this context, the right to data protection also protects even though there is no information processed through transparency[430] measures as knowing that the law requires disclosure of who processes their data[431] supposed to free them from such panoptic prison.

In this regard, the separate right to data protection plays a key role in

---

[426] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Iii."

[427] "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Ii."

[428] Michel Foucault, Discipline and Punish: The Birth of the Prison (London: Penguin Books, 1991). We agree with Haggerty and Ericson, that panopticism is not adequate to explain the surveillant assemblage experienced today. And even though data subjects must know they are being watched in everything they do online and the rhizomatic flow of information must be their real concern, in the face of the informational asymmetry that they also experience, we refer to panopticon to describe their angst. See Haggerty and Ericson, "The Surveillant Assemblage."

[429] CAHAI, "Ad Hoc Committee on Artificial Intelligence (Cahai): Feasibility Study".

[430] Hallinan also counts Art. 25 and Art. 35 GDPR among such measures which we will delve into later. See also Reuben Binns, "Data Protection Impact Assessments: A Meta-Regulatory Approach," Information Privacy Law eJournal (2016).

[431] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Ii."

maintaining a democratic society, however, as the data subjects do not have subjective rights to abstract constitutional principles such as democracy, it cannot be tasked with saving the democracy. Von Grafenstein proposes political microtargeting as an example, which may manipulate individual opinions[432] and undermine autonomous exercise of the right to vote (Art. 39 and 40 CFREU) which is protected by their right to data protection, however this would not amount to require the data controller to take measures in order to maintain a public political debate but to be taken into consideration as an exacerbating factor in the balancing exercise.[433]

2.1.4 Risks of ADM to the Rights of WMG

The right to privacy under Art. 8 ECHR covers a person's general privacy, physical psychological, and moral integrity, as well as identity and autonomy. Identification, authentication, assessing, predicting, influencing a person's behaviour, surveillance, profiling and categorising individuals through algorithmic systems affect all the aforementioned aspects of the right to privacy. Such interventions may cause adapting to certain norms and deepen the power imbalance between those who control these systems and individuals who use them. According to CAHAI, Art. 8 ECHR entails a private place free from such tracking technologies for the sake of personal development and democracy.[434]

As the public sphere is mediated now by algorithmic systems controlled by a few private platforms, constant tracking and surveillance would chill individuals and groups from exercising their political freedoms such as freedom of expression, assembly and association as guaranteed in the ECHR At. 10-11. Especially the

[432] EDPS Opinion 3/2018, "European Data Protection Supervisor's Opinion on Online Manipulation and Personal Data," (2018).

[433] "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Ii," 22.

[434] CAHAI, "Ad Hoc Committee on Artificial Intelligence (Cahai): Feasibility Study".

loss of group anonymity would be effective in discouraging from reading certain books and newspapers, engaging with certain social and political causes and groups online. When it comes down to the not having anything to hide rhetoric, such chilling effects are more likely to effect WMG as they become targets of discrimination just because of their very core identities[435] and by being openly struggling for their rights lead to vulgar reactions.[436] Content curation and moderation, search and recommender algorithms may reinforce stereotypes,[437] polarization and extremism by creating eco chambers and filter bubbles which are the opposite of broadmindedness, plurality and diversity that are necessary for a democratic society. When these values that make a democratic society are at risk, WMG are the first ones to be affected as their life styles and speech may be strange and shocking to majority groups, especially when they are trapped in a bubble without diversity. Platforms are increasingly turning to algorithmic systems to identify, flag, downrank and remove content that does not conform with their terms and services.[438] Nuances and the context related to WMG is easier to miss for algorithmic systems which may lead to false positives.

As biases, stereotypes and discrimination based on various grounds including on proxies and intersectional grounds are made easier to be perpetuated and amplified by algorithmic systems, the right to equality (Art. 14 ECHR and Protocol 12) is also under major risk. Most crucially, both organizational and technical opacity and complexity of these systems bear the risk of these risks and harms may remain undetected.[439] Even though ECHR, CFREU and the anti-discrimination framework under the secondary laws of the EU do apply to the

---

[435] Jessica L. Roberts, "Protecting Privacy to Prevent Discrimination," William and Mary Law Review2015.

[436] Dinar, "The State of Content Moderation for the Lgbtiqa+ Community and the Role of the Eu Digital Services Act"; Marshall, "Algorithmic Misogynoir in Content Moderation Practice".

[437] Noble, Algorithms of Oppression: How Search Engines Reinforce Racism.

[438] CAHAI, "Ad Hoc Committee on Artificial Intelligence (Cahai): Feasibility Study". 9.

[439] Ibid.

discrimination stemming from the use of algorithmic systems, it is also acknowledged that there are many limitations.

All the mentioned risks to these rights are amplified as a few private platforms have become the de facto public sphere where there is an ever growing power imbalance and limited regulation to ensure the rule of law principles including transparency, accountability, non-discrimination, and equality apply.[440]

ECtHR have not developed specific case law related to AI assisted algorithmic systems, however, Art 8, 10 and 14 have come up generally algorithm related cases. In *Big Brother Watch v. The United Kingdom*[441] and *Centrum för rättvisa v. Sweden,*[442] the ECtHR found that bulk interception of online communications was not in itself incompatible with the ECHR, provided that appropriate oversight and safeguards are at place. The Court did not consider algorithmic surveillance and the dangers of correlations in the given context. In *Delfi AS v. Estonia,*[443] a positive obligation was created on news and other content sites to police hate speech and an automated *ex ante* removal of content amounting to hate speech was not deemed contrary to Article 10 ECHR.[444] In *Magyar Kétfarkú Kutya Párt v. Hungary*[445] a mobile application made available by a political party and allowing voters to share anonymous photographs of their ballot papers found by the ECtHR

---

[440] Ibid.

[441] Big Brother Watch And Others v. The United Kingdom (Applications nos. 58170/13, 62322/14 and 24960/15) 25 May 2021 https://hudoc.echr.coe.int/eng?i=001-210077

[442] Centrum För Rättvisa v. Sweden (Application no. 35252/08) 25 May 2021 https://hudoc.echr.coe.int/eng?i=001-210078

[443] Delfi AS v. Estonia (Application no. 64569/09) 16 June 2015 https://hudoc.echr.coe.int/eng?i=001-155105

[444] A. Murray, *Information Technology Law: The Law and Society*, 3 ed. (Oxford University Press, 2016), 205-08.

[445] Magyar Kétfarkú Kutya Párt v. Hungary (Application no. 201/17) 20 January 2020 https://hudoc.echr.coe.int/eng?i=001-200657

not to violate voter secrecy and fairness of elections.[446]

Upon commissioning a study that looked into various ethical AI frameworks from the globe, CAHAI concludes that mandatory governance by governmental and intergovernmental authorities is required. It is very important that CAHAI refers to an unalignment of interests between the developers and deployers of these technologies and those who are likely to be bearing the negative effects as a reason for not taking the route to self-regulation. Another reason for the need for specific regulation addressing AI systems is that existing fundamental rights and freedoms are formulated as general principles, moreover they are value-laden and contested, in conjunction with information asymmetries and black box problem of these systems may lead to interpretation problems. For instance, despite criticism that we will discuss later, GDPR is an example of a specific law that turns such general principles into concrete rights for data subjects and obligations for controllers, alongside with the AI Act Proposal. We will later dig into the advantages and shortcomings of these rights and obligations in tackling algorithmic discrimination against the WMG. CAHAI suggests that in a legal framework addressing such issues, the right to a fair trial would turn into a right to contest and getting insight into evidence in algorithmic systems. Another example would be deriving a due diligence obligation from the right to non-discrimination for the developers and deployers of algorithmic systems to look for and prevent unjust bias throughout the life cycle of these systems.[447]

Most importantly CAHAI warns that without clear obligations the rights of individuals may lack adequate protection against negative effects of algorithmic systems. CAHAI gives weight into transparency and explainability of these systems alongside with mechanisms of responsibility, accountability and redress to ensure that they can be assessed for infringement of fundamental rights and

[446] CAHAI, "Ad Hoc Committee on Artificial Intelligence (Cahai): Feasibility Study".

[447] Ibid.

freedoms. Otherwise, the information asymmetry between the developers/deployers and not only negatively affected individuals/groups but also regulators and law enforcers would make it impossible to challenge ADM and assess the existence of rights infringements.[448]

From the very days it was emerging, data protection was understood to aim at protecting the parliamentary democracy by protecting the minorities as they can be easily discriminated against through the use of advancing data processing technologies.[449]

Modernised Convention n. 108 in Art. 6(2) links processing of special categories of personal data without appropriate safeguards to a notable risk of discrimination in addition to risks to other interests, rights and fundamental freedoms of the data subject. However, as we showed above, processing of non-special category personal data may also lead to discriminatory results for WMG such as proxy discrimination. Furthermore, gender is not a special category data but carries a notable risk of discriminatory outcomes. Moreover, group profiles and other processing of anonymous data may still result in discrimination against WMG such as in the case of statistical discrimination.[450]

GDPR Art. 1 sect. 2 clearly mentions that it protects fundamental rights and freedoms of natural persons in general and their right to the protection of personal data in particular. Recitals 2 and 4 GDPR, by using the word “respect” indicate to

---

[448] Ibid.

[449] Theilen et al., "Feminist Data Protection: An Introduction."; Steinmüller et al., "Grundfragen Des Datenschutzes: Gutachten Im Auftrag Des Bundesministeriums Des Innern [Bt-Drs. Vi/3816 Anlage 1]."

[450] Sandra Wachter, "Affinity Profiling and Discrimination by Association in Online Behavioural Advertising (May 15, 2019). , Vol. 35, No. 2, 2020, Forthcoming, Available at Ssrn: Https://Ssrn.Com/Abstract=3388639 or Http://Dx.Doi.Org/10.2139/Ssrn.3388639," *Berkeley Technology Law Journal* 35, no. 2 (2020).

the negative obligation of the data controller in the other fundamental rights of the data subjects. Recital 4 not only indicates to a need for balancing between the data protection and other fundamental rights and freedoms but also data protection as a prerequisite for the effective exercise of all these other rights.[451] Recital 4 provides a non-exhaustive list of rights that are recognized in CFREU and enshrined in the EU Treaties that are protected among which there is the right to privacy and freedom of expression, surprisingly though the wording left out the right not to be discriminated against.

As we have discussed in this section, rights to privacy and equality are both constitute a prerequisite for each other and they also ensure other rights, especially those which form the backbone of our democratic societies namely freedoms of expression, association and assembly. However, social injustices are woven in the very structures of our societies and this results in the system not living up to its promises to everyone equally. Thus balancing different rights, freedoms and interests of different groups of people and individuals in the society becomes an almost impossible quest. We have already disclosed that this study is interested in taking multiple disadvantages some historically marginalized groups experience into consideration in a manner of prioritizing their needs and interests in not being discriminated against by the use of algorithmic systems. This requires a broader understanding of historical social injustice and structural power imbalances and a strategy to address them to prevent algorithmic discrimination and stigmatization of WMG. In this direction, one of the prominent suggestions came from Malgieri

[451] Hielke Hijmans, "Article 1 Subject-Matter and Objectives," in The Eu General Data Protection Regulation (Gdpr): A Commentary, ed. Christopher Kuner, Lee A. Bygrave, and Christopher Docksey (New York: Oxford University Press, 2020); von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Ii."; ibid.

and Niklas[452] who draw from layered vulnerability approach developed by Luna.[453] This study is also remarkable in its direct references to feminist theory.

[452] Malgieri and Niklas, "Vulnerable Data Subjects."

[453] Florencia Luna, "Elucidating the Concept of Vulnerability: Layers Not Labels," International Journal of Feminist Approaches to Bioethics 2, no. 1 (2009); "Identifying and Evaluating Layers of Vulnerability - a Way Forward," Dev World Bioeth 19, no. 2 (2019).

2.2 EUROPEAN ANTIDISCRIMINATION LAW

Discriminating between things, concepts and other human beings is the way our brains function as they make sense of the world and navigate in it.[454] What becomes a subject of law is when such discrimination is based on a morally irrelevant attribute and causes a person to be treated less favourably on that ground. Then the first question of legal rules which prohibit discrimination, is a moral one that asks what constitutes an unacceptable ground. For example, properties that one cannot choose or change are by consensus fall in the morally unacceptable category such as sex,[455] race and disability.[456] However, religion, political and philosophical opinions, are seemed to be areas where one can have a choice. Also, which brings these areas into more controversial categories for the legal principle of non-discrimination. Again, although age is a category which one cannot choose, retirement age and not employing underaged children are widely accepted.[457]

Pluralism, tolerance and broadmindedness are essentials of a democratic society, which Europe considers itself to be one, as discussed above. This is closely linked to the acknowledgment of the need to be treated in a similar way with the other human beings as a fundamental human right[458] even though one is different than the majority in such a society from some aspects.[459] In this regard, we discussed the interplay of fundamental rights that are closely related to our lead examples.

---

[454] Hildebrandt, "Profiling and the Rule of Law," 57-58.

[455] We shall note here that regarding sex as something immutable is exclusive of transgender individuals, see. Os Keyes, "The Misgendering Machines: Trans/Hci Implications of Automatic Gender Recognition" (paper presented at the The ACM on Human-Computer Interaction, 2018).

[456] Ellis and Watson, Eu Anti-Discrimination Law.

[457] Ibid.

[458] Ibid.

[459] And sometimes to be treated not the same way because of being in a different situation.

ECHR and CFREU have an important role in interpreting antidiscrimination law of the EU. The sources of EU antidiscrimination law are main primary sources of the EU law (the Treaty on European Union-TEU and the Treaty on the Functioning of the European Union-TFEU); secondary legislation of the EU (this area is regulated by directives, mainly the Recast Directive,[460] the Race Directive,[461] the Framework Directive,[462] the Goods and Services Directive[463] that we will refer to as Equality Directives when we address all off them together); decisions of The Court of Justice of the European Union (CJEU) and the General Court; and instruments for the protection of fundamental human rights (The Charter of Fundamental Rights of the European Union-CFEU and The European Convention on Human Rights-ECHR). Also general principles of law as recognized by the CJEU has a similar status to the Treaties.[464]

EU Law forbids discrimination on grounds of nationality, sex, part-time and temporary employment, racial or ethnic origin, religion or belief, disability, age, and sexual orientation. Article 21 CFREU adds on to these grounds a possibility of inexhaustive list of protected grounds, however, currently CJEU continues to apply fragmented and limited scope of protection in Equality Directives.[465]

In this part, we will delve into the antidiscrimination law that is applicable to our lead examples from human resources field. As explained in the introduction the main reason of focusing in the employment related algorithmic discrimination is because EU antidiscrimination law covers all the protected grounds only in this

[460] Directive 2006/54/EC

[461] 2000/43/EC

[462] 2000/78/EC

[463] 2004/113/EC

[464] Ellis and Watson, *Eu Anti-Discrimination Law*, 99.

[465] Angela Ward, "The Impact of the Eu Charter of Fundamental Rights on Anti-Discrimination Law: More a Whimper Than a Bang?," *Cambridge Yearbook of European Legal Studies* 20 (2018).

sphere.[466] Here we will explore applicability and the shortcomings of the EU non-discrimination law even if when it is applicable to all the protected grounds.

2.2.1 Applicability of the European Antidiscrimination Law to Algorithmic Discrimination

The EU Antidiscrimination law is designed to cope with human discrimination which is different from our subject matter of discrimination by algorithmic processes.[467] A special report by the European Commission on Algorithmic discrimination in Europe identified six challenges for the European Antidiscrimination Law in this regard as (i) human factor, stereotyping and cognitive bias, (ii) data, (iii) correlations and proxies, (iv) transparency and explainability, (v) scale and speed, (vi) responsibility, liability and accountability.[468]

There seems to be a consensus among scholars that indirect discrimination is a better fit in capturing algorithmic discrimination.[469] It is also suggested that when applied to algorithmic discrimination, the distinction between direct and indirect discrimination is blurred.[470] Ellis and Watson are in the opinion that whichever form the discrimination takes, discrimination would consist of two elements: (i) adverse treatment that causes harm, (ii) such treatment is based upon a protected

---

[466] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 29.

[467] Ibid., 40.

[468] Ibid., 51.

[469] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."; Barocas and Selbst, "Big Data's Disparate Impact."; Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai ".

[470] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

ground, or in other words, the causality is linked to such ground.[471] Therefore, a correct understanding of the concept of discrimination shall transcend the merely technical distinctions among its direct and indirect manifestations. We will now have a closer look into these two types of discrimination and how they could apply to algorithmic discrimination.

2.2.1.1 Direct Discrimination[472]

It reflects formal equality[473] and requires three conditions to be met to occur. First, an individual must be treated less favourably than others. Second, these others (comparators) must be in a similar situation. Third, there must be a causal link between the unfavourable treatment and a 'protected ground' which covers a particular characteristic they hold.[474]

All Equality Directives phrase the definition of direct discrimination including not only being less favourably treated by comparison to those in a similar situation "have been treated," but also extend it to hypothetical comparisons that less favourable treatment by comparison to those in a similar situation "would be treated." Direct discrimination can be both overt and covert.[475] Moreover, the intention of the discriminator is irrelevant no matter which type of discrimination is triggered unlike the US anti-discrimination law which led to scholars[476] debating that implicit bias in algorithmic systems would fall under disparate impact (indirect discrimination). Thus, in the case of machine learned algorithms European notion of direct discrimination is likely to capture implicit bias as well

---

[471] Ellis and Watson, *Eu Anti-Discrimination Law*, 173.

[472] It is referred as "disparate treatment" under US equality law.

[473] Ellis and Watson, *Eu Anti-Discrimination Law*, 142.

[474] *Handbook on European Non-Discrimination Law*, 2018 Edition (Luxembourg: Publications Office of the European Union, 2018).

[475] Ellis and Watson, *Eu Anti-Discrimination Law*, 145.

[476] Barocas and Selbst, "Big Data's Disparate Impact."

as explicit bias of the decision-maker if the protected ground was directly used as a variable (e.g. via labelling)[477] that is causally linked to the output with a discriminatory effect on some individuals compared to others.

However, the need of a comparator in order to establish direct discrimination, is by itself an important problem of the anti-discrimination law, especially in the case of algorithmic discrimination because of the black box problem, as well as protection of IP rights and trade secrets. These may make gathering evidence[478] that shows another group which has been or would have been treated more favourably and does not share characteristics that fall under the protected ground on which the plaintiff is allegedly treated unfavourably, really gruelling. On the other hand, the notion of direct discrimination have some advantages. According to the CJEU's case law, when the discrimination on the protected ground is obvious, a comparator group is not necessary[479] as in the case of discrimination based on pregnancy because only female sex is capable of getting pregnant.[480] The CJEU also has found that an abstract victim would be sufficient in some direct discrimination cases[481] such as when an employer openly expresses not to employ

---

[477] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 10.

[478] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 69; Wachter, "Affinity Profiling and Discrimination by Association in Online Behavioural Advertising (May 15, 2019). , Vol. 35, No. 2, 2020, Forthcoming, Available at Ssrn: Https://Ssrn.Com/Abstract=3388639 or Http://Dx.Doi.Org/10.2139/Ssrn.3388639," 46.

[479] C-177/88 [1990] *Dekker v Stichting Vormingscentrum voor Jong Volwassenen* ECLI:EU:C:1990:383; C-32/93 [1994] *Webb v EMO Air Cargo* ECLI:EU:C:1994:300

[480] See Ellis and Watson, *Eu Anti-Discrimination Law*, 333-38. for a broader understanding of the role of the comparator in discrimination on grounds of pregnancy and maternity.

[481] Wachter, "Affinity Profiling and Discrimination by Association in Online Behavioural Advertising (May 15, 2019). , Vol. 35, No. 2, 2020, Forthcoming, Available at Ssrn: Https://Ssrn.Com/Abstract=3388639 or Http://Dx.Doi.Org/10.2139/Ssrn.3388639," 36.

gay football players[482] or immigrants.[483]

Anyhow, the scope of the notion of direct discrimination especially in its ability to capture algorithmic discrimination is yet uncertain.[484]

In *Coleman*,[485] the CJEU applied direct discrimination to the case when a woman was treated unfavourably not on a protected ground which covers a particular characteristic she held, but on the basis that she had to take care of her disabled son. It is now known as "discrimination by association." In the light of this, some authors found a potential for direct discrimination to extend towards proxy discrimination and miscategorization by algorithms where they directly cause unfavourable treatment based on a protected ground.[486]

Furthermore, the European Commission considered that the Race Directive and the Framework Directive were also prohibitive for direct discrimination "on the basis of a wrong perception or assumption of protected characteristics"[487] which would capture one of the prominent risks in algorithmic discrimination. However, it is pointed out that the CJEU has not yet established its case law in this

---

[482] C-81/12 [2013] *Asociaţia Accept* ECLI:EU:C:2013:275

[483] C-54/07 [2008] *Feryn* ECLI:EU:C:2008:397

[484] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

[485] C-303/06 [2008] *S. Coleman v Attridge Law and Steve Law* ECLI:EU:C:2008:415

[486] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

[487] "European Commission (2014) ", *Joint Report on the application of Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin ('Racial Equality Directive') and of Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation ('Employment Equality Directive')* COM(2014) 2 final (Brussels), https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52014DC0002&from=EN.

direction.[488] In *Kaltoft*,[489] Jääskinen AG mentioned the question whether discrimination resulting from the false presumption of the employer for the employee's disability would be covered by the Framework Directive, however he concluded that the national court has to decide whether obesity amounts to disability and if it does the discrimination Mr. Kaltoft experienced would be on the basis of disability under the Directive but not on the basis of presumed disability for his obesity. Citing the case law (*Chacón Navas*[490] and *Coleman*[491]), the CJEU decided that the scope of the Framework Directive "should not be extended by analogy beyond the discrimination based on the grounds listed exhaustively in Article 1 thereof"[492] which lists disability but not obesity and left it to the national court to decide whether obesity is to be considered as disability according to the Court's definition in *HK Danmark*.[493] *Kaltoft* was interpreted to be problematic as regards to the applicability of direct discrimination to cases where algorithmic profiling results in an individual being assigned to a profile over a falsely inferred identity or group membership.[494]

Wachter argued, on the other hand, that the job ads that are delivered or not delivered based on other users behaviours on the platform, such as in the case of

---

[488] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 68.

[489] C-354/13 [2014] *Fag og Arbejde (FOA) v Kommunernes Landsforening (KL)* ECLI:EU:C:2014:2463

[490] C-13/05 [2006] *Sonia Chacón Navas v Eurest Colectividades SA*, ECLI:EU:C:2006:456

[491] C-303/06 [2008] *S. Coleman v Attridge Law and Steve Law* ECLI:EU:C:2008:415

[492] We will discuss later that such a static categorical approach (see. Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 70.) to protected grounds is not a good fit for capturing dynamic categories algorithmic discrimination may be based upon.

[493] C-335/11 and C-337/11 [2013], *HK Danmark v Dansk almennyttigt Boligselskab (C-335/11), and HK Danmark v Dansk Arbejdsgiverforening (C-337/11)* ECLI:EU:C:2013:222 paragraph 47.

[494] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 69.

Facebook's look-a-likes feature would be captured by this newly emerging doctrine of discrimination by association.[495] As we will discuss later though, discrimination by algorithmic processes bring by serious enforcement and redress problems, one of which is becoming aware of the possibility of having been treated differently on a characteristic one holds that under a protected ground by an algorithm, as in the case of not being delivered job ads based on information provided by the user and inferred from the user's behaviour on social media which had put them in a profile and now their future behaviour is being predicted based on the behaviour of those who look like their perceived, ascribed profile on this channel.

Even if those who manage to become aware of being subject to such algorithmic discrimination, in both direct and indirect discrimination cases, first of all, the person alleging discrimination must be able to establish a presumption of discrimination (prima facie discrimination). Once the prima facie discrimination is established though, the burden of proof shifts to the defendant. In direct discrimination cases, in order to rebut the plaintiff's claim, defendant must either prove that the claimant is not in a similar or comparable situation or the unfavourable treatment is based on objective differences not the protected ground. If the discrimination claim cannot be rebutted, then the plaintiff may resort to objective justification.[496] One of the strengths of direct discrimination is that objective justification only applies to indirect discrimination[497] unless it falls under a few exceptions accepted in direct discrimination cases.[498]

---

[495] Wachter, "Affinity Profiling and Discrimination by Association in Online Behavioural Advertising (May 15, 2019). , Vol. 35, No. 2, 2020, Forthcoming, Available at Ssrn: Https://Ssrn.Com/Abstract=3388639 or Http://Dx.Doi.Org/10.2139/Ssrn.3388639."

[496] *Handbook on European Non-Discrimination Law*.

[497] C-356/09 [2010] *Pensionsversicherungsanstalt v Kleist* ECLI:EU:C:2010:703 at para 41-43.

[498] Direct discrimination can be justified with genuine occupational requirements; exceptions in relation to religious institutions; exceptions particular to age discrimination. For more details thereon, see *Handbook on European Non-Discrimination Law*, 94-108.

In the employment context, a genuine and determining occupational requirement may justify direct discrimination as set out in the Race (Art. 4(1)), the Framework (Art. 4(1)) and the Recast Directives (Art. 14(2)). For this to apply the established case law of the CJEU is clear that "it is not the ground on which the difference of treatment is based but a characteristic related to that ground which must constitute a genuine and determining occupational requirement."[499] For example, CJEU has accepted age to be "objectively linked to the reduction of physical capabilities."[500] In *Bougnaoui*, the CJEU decided that "willingness of the employer to take account of the particular wishes of the customer" in dismissing the employee based on wearing headscarf was a subjective consideration and did not meet the objectivity criterion.[501]

According to Hacker, an algorithmic system may use a protected ground as a direct input or those who share a characteristic pertaining to a protected ground may receive a worse score from the algorithm, in either case differential treatment can only be justified if the input or scoring practice capture something essential for the job to be performed.[502] Moreover, we understand from the relevant case law that a correlation would not be enough and the employer would need to prove causality between the protected ground and characteristic that is objectively essential to the performance of the task.[503] Such justification does not only mean a negative effect on the WMG. Membership in a protected group can be essential for a particular role or performance in artistic professions; women to work in women-only fitness clubs, or sexual orientation or ethnicity playing a crucial role

[499] C-229/08 [2010] *Wolf v Stadt Frankfurt am Main* ECLI:EU:C:2010:3 para 35.

[500] C-447/09 [2011] *Prigge and Others* ECLI:EU:C:2011:573 para 28.

[501] C-188/15 [2017] *Bougnaoui and ADDH* ECLI:EU:C:2017:204

[502] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 21.

[503] C-258/15 [2016] *Salaberria Sorondo v Academia Vasca de Policía y Emergencias* ECLI:EU:C:2016:873 para 48.

in counselling in related areas can be examples[504] to preference of WMG not to be blocked by formal equality rules.

This of course brings to mind the possibility of positive action, that is the differential treatment for the benefit of WMG occurs not because of a genuine and determining occupational requirement but in order to remedy historical discrimination.[505] Thus far, EU primary law and equality directives allow positive action as an exception[506] to the individual fairness principle. Especially in the case that fairness metrics are applied to algorithmic processes to achieve statistical parity,[507] it may fall under the definition of positive action.[508] Even though positive action justifies both direct and indirect discrimination, not every attempt to achieve fairness will be acceptable according to the case law of the CJEU.

In *Kalanke*, the Tesauro AG interpreted positive action that seeks to remedy historical discrimination which is in a compensatory nature in favouring the disadvantaged groups in the form of quotas and goals to be amounting to reverse discrimination.[509] Ellis and Watson finds it unfortunate that the Advocate General did not distinguish between quotas and goals, the latter having a temporary nature

---

[504] *Handbook on European Non-Discrimination Law*.

[505] Ellis and Watson, *Eu Anti-Discrimination Law*, 425.

[506] Article 157(4) TFEU, Article 23(2) CFEU, Article 6 of the Race Directive (2000/43/EC), Article 7 of the Framework Directive (2000/78/EC), Article 6 of the Goods and Services Directive (2004/113/EC), and Article 3 of the Recast Directive (2006/54/EC).

[507] "*Statistical parity* is the property that the demographics of those receiving positive (or negative) classifications are identical to the demographics of the population as a whole. Statistical parity speaks to *group fairness" Cynthia Dwork et al., "Fairness through Awareness," ArXiv abs/1104.3913 (2012).*

[508] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 32.

[509] Tesauro AG, C-450/93 [1995] *Kalanke v Freie Hansestadt Bremen* ECLI:EU:C:1995:105 at 3065.

and apply until the effects of past discrimination is remedied.[510] *Kalanke* was concerned with a tie-break situation, meaning that women would be given priority in appointment or promotion to official posts over male applicants provided that they had equal qualifications and number of women in the sector is less than fifty percent.[511] The CJEU decided that a rule which guarantees women "absolute and unconditional priority for appointment or promotion" was disproportionate to the aim of promoting equal opportunities and was not covered by the positive action exception in question.[512] On this account, Hacker identifies three conditions that may amount to reverse discrimination during efforts for obtaining algorithmic group fairness. First, a corrective measure applied to the algorithm would result in a member of a protected group (a female employee) to be treated more favourably than a member of another group (a male employee) with similar proximity to the algorithmic output before the correction such as a performance score that determines the promotion. Second, the original output was not based on biased input and training data and had predictive accuracy, in other words was correct in reflecting the reality. For instance, the score given to female employees is slightly less because height is one of the factors weighed, and sex correlates with height. Third, if a decision was based on the original output (performance score) it would have been justifiable, such as a certain height being a genuine requirement of the job. Thus, correcting a factually correct indirect proxy discrimination through algorithmic fairness methods may be captured by reverse discrimination.[513]

Algorithmic fairness measures seem to be more applicable at stages before selection, especially during the training stage of the algorithm before it applies to

---

[510] Ellis and Watson, *Eu Anti-Discrimination Law*, 425.

[511] Ibid.

[512] C-450/93 [1995] *Kalanke v Freie Hansestadt Bremen* ECLI:EU:C:1995:322 para 22. See also *Handbook on European Non-Discrimination Law*, 74.

[513] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 32.

real life, especially by correcting biased training data.[514] In *Marschall*,[515] the rule at question had a saving clause which required in order to prefer women, reasons specific to another candidate not to outweigh. The CJEU decided that such a clause which guarantees that male candidates with equal qualification to women candidates will be subject to objective assessment with a possibility of discovery of their specific qualities that may override the rule giving female candidates priority, saved this rule from being absolute and unconditional and deemed it proportionate to the equal opportunities objective.[516] In *Badeck*, the CJEU was concerned with the hiring funnel and has come to a distinction where provisions such as a call for interview which do not attempt to achieve a final result such as appointment or promotion to a position, may apply a quota provided that a preliminary examination ensures only qualified candidates to be called to interview.[517] Once screened candidates come down to the selection stage though, further conditions in *Marschall* will apply, which means that the algorithmic correction applied will have to take into consideration factors that, as Hacker argues, are not easy to be mathematically expressed and translated into code. Thus Hacker suggests a human-in-the-loop would be required to ensure fairness applied by the model had taken all the factors into consideration as a final review[518] otherwise such efforts may be considered direct discrimination. Thus, especially for diversity and inclusion tools in HR, it is to be put into consideration.

---

[514] Ibid., 33.

[515] C-409/95 [1997] *Marschall v Land Nordrhein-Westfalen* ECLI:EU:C:1997:533 para 33 and 35.

[516] *Handbook on European Non-Discrimination Law*, 75; Ellis and Watson, *Eu Anti-Discrimination Law*.

[517] C-158/97 [2000] *Badeck and others v Landesanwalt beim Staatsgerichtshof des Landes Hessen* ECLI:EU:C:2000:163 para 60-63.

[518] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

2.2.1.2 Indirect Discrimination[519]

Many scholars agree that indirect discrimination is a better fit in capturing algorithmic discrimination.[520] Indirect discrimination occurs when (i) a neutral rule, criterion or practice (ii) disadvantages a person or a group; (iii) defined by a protected ground, (iv) in a significantly more negative way (v) in comparison to others in a similar situation, (vi) unless that provision, criterion or practice is objectively justified by a legitimate aim and the means of achieving that aim are appropriate and necessary.[521]

The concept of indirect discrimination reflects equality of opportunity.[522] It is the result of acknowledgement that different situations should be treated differently as treating them the same way may put certain people, especially those burdened with historical and structural exclusion and discrimination such as the WMG, at a particular disadvantage.[523] However, it shall be noted that indirect discrimination does very little to actually change the disadvantaged position and stereotyping experienced by these groups, such as in the case that CJEU has acknowledged throughout its case law that part-time work is predominantly performed by women due to their care responsibilities for their families and homes, thus discrimination on the basis of part-time work is captured by indirect

---

[519] It is referred as "disparate impact" under US equality law.

[520] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."; Barocas and Selbst, "Big Data's Disparate Impact."; Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai ".

[521] *Handbook on European Non-Discrimination Law*, 54. See also Article 2(2) of the Burden of Proof Directive (Directive 97/80, OJ [1998] L 14/6), Article 2(2)(b) of the Race Directive (2000/43/EC), Article 2(2)(b) of the Framework Directive (2000/78/EC), Article 2(b) of the Goods and Services Directive (2004/113/EC), and Article 2(1)(b) of the Recast Directive (2006/54/EC).

[522] Ellis and Watson, *Eu Anti-Discrimination Law*, 176.

[523] *Handbook on European Non-Discrimination Law*, 53.

discrimination based on sex. However, according to Ellis and Watson, this have cemented women's position as primary caregivers while doing nothing to challenge family unfriendly way work is organized.[524] On the other hand, notion of indirect discrimination suits better in capturing algorithmic discrimination as it applies when a facially neutral policy and practice disproportionately causes a negative effect on the members of a protected group, which captures covert direct discrimination that would be the case in masking discriminatory intent on the basis of membership in a protected group by using a facially neutral proxy instead, as well as implicit bias that causes a similar result.[525] Furthermore, sampling bias and incorrect labelling resulting from lack of care rather than implicit or explicit bias on the side of developers would also be captured by indirect discrimination.[526] Also the focus shifting from disparate treatment of individuals to adverse impact on groups makes it a better fit to deal with the very logic of algorithms.[527] It is argued that the focus on the effects rather than the inner operation of the algorithm may be liberating for the claimants to circumvent some challenges posed by algorithmic discrimination such as access to the datasets and the code, difficulty in understanding the inner workings of algorithms for non-experts, and being able to bring a case directly against the end-user rather than being lost in finding whose action in the chain of actors caused the disparate treatment.[528]

As we already introduced above, in order to bring a case alleging direct or indirect discrimination under EU antidiscrimination law "a claimant must establish prima facie discrimination." This is done by demonstrating that (i)

---

[524] Ellis and Watson, *Eu Anti-Discrimination Law*, 176.

[525] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

[526] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

[527] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 70.

[528] Ibid.

existence or likelihood of a harm; (ii) which effects or likely to affect people with characteristics that fall under a protected ground; and (iii) which is significantly disproportionate compared with the comparator group.[529]

If these requirements are met the burden of proof shifts to the alleged defendant who either can demonstrate that the victim is not in a similar situation to their 'comparator'; or that the difference in treatment is based on some objective factor, unconnected to the protected ground. In other words, indirect discrimination can be justified if a legitimate aim is pursued and the measure is necessary and proportionate. Predictive accuracy of machine learning models may serve as a good justification for this.[530]

Statistical evidence plays an important role in both bringing out presumption of discrimination[531] from the side of the plaintiff and in justifying the differential treatment from the side of the defendant in indirect discrimination cases. We will focus on the side of the defendant below among the difficulties to seek redress as statistical nature of the ADM makes justifying easy. Here we will focus on the victim's side first and also on the possibility of building systems that would refrain from indirect discrimination as an ex-ante measure on the developers' side.

One of the important nuances in the anti-discrimination doctrine is that once a rule is known, there is no need to wait for adverse effect to occur and the likelihood of discrimination can be brought before the court. Race, Framework and Recast Directives all define indirect discrimination inclusive of contingent harm.[532]

---

[529] Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai " 15.

[530] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

[531] Ibid.; *Handbook on European Non-Discrimination Law*; Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai ".

[532] Ellis and Watson, *Eu Anti-Discrimination Law*.

First of all, it must be noted that in some areas where statistical evidence is hardly to be found due to the sensitivity of the protected ground such as in the case of sexual orientation,[533] it may bring disadvantage on the side of the victims. The important point is to provide evidence that the facially neutral rule, criterion or practice is likely to effect the protected group more than the comparator group. If it causes adverse effect on 100% of the members of the protected group without any exceptions, then it will be considered direct discrimination.[534]

Here, how the comparator group is chosen and how is more adverse effect or advantageous treatment is measured become important problems in establishing prima facie discrimination in indirect discrimination cases.

Both in choosing the comparator group and statistical evidence, the reach of the contested rule, criterion or practice also plays a crucial role. In finding a suitable comparator for equal pay for equal work (between women and men) principle, CJEU clarified in *Allonby* that the principle was "not limited to situations in which men and women work for the same employer." It would rather be determined based on the source of the contested rule, criterion or practice; be it a legislative provision or collective labour agreement, work carried out in the same establishment or service, whether private or public. When there is no such single source then it would be out of the scope of the equal pay for equal work principle set out in the Treaty simply because then who is responsible for the inequality and could restore equal treatment could not be identified.[535] In such, if a national

---

[533] Ibid.

[534] C-267/06 [2008] *Tadao Maruko v Versorgungsanstalt der deutschen Bühnen* ECLI:EU:C:2008:179; C-267/12 [2013] *Frédéric Hay v. Crédit agricole mutuel de Charente-Maritime et des Deux-Sèvres* ECLI:EU:C:2013:823; *Handbook on European Non-Discrimination Law*, 58.

[535] C-256/01 [2004] *Allonby v Accrington & Rossendale College* ECLI:EU:C:2004:18 para 45 and 46.

legislation is the source then regional[536] or global[537] statistical evidence is not admissible, however when a regional legislation is the source then regional statistics can be used.[538] On this account, Wachter et. al pointed out the problem in determining reach may cause in addressing indirect discrimination occurring on a platform[539], such as in the case of one of our lead examples of job ad delivery.

Courts have discretion on choosing the comparator group for the case at hand which is determining of the result thereof and sometimes they may lack the perspective of the disadvantaged group and be rigid in their legal analyses.[540]

The ex-ante intervention mechanism from the victims' side based on contingent harm can only work in cases where the rules of an algorithm is made transparent by those who hold rights to it. Most of the time intellectual property rights and trade secrets merged with competition concerns will impede this from happening. Thus, it leaves ex-ante prevention to the good will of the developers and deployers of ADM systems. Then, the question arises whether it would be possible for developers and users of ADM systems to design and monitor their algorithms in a way it would not cause indirect discrimination under EU antidiscrimination law.

Under the US antidiscrimination law[541] there is a rule of thumb of 4/5th for

---

[536] C-300/06 [2007] *Ursula Voß v Land Berlin* ECLI:EU:C:2007:757

[537] C-527/13 [2015] *Lourdes Cachaldora Fernández v Instituto Nacional de la Seguridad Social (INSS), Tesorería General de la Seguridad Social (TGSS)* ECLI:EU:C:2015:215

[538] C-1/95 [1997] *Hellen Gerster v Freistaat Bayern* ECLI:EU:C:1997:452

[539] Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai " 19.

[540] C-249/97 [1999] *Gruber v Silhouette International Schmied GmbH & Co. KG.* ECLI:EU:C:1999:405

[541] Equality laws of the US are relevant not only because they inspired the indirect discrimination doctrine of the EU antidiscrimination law but also because many of the vendors who provide diversity and inclusion solutions as B2B services in the field of human resources are from the US.

disparate impact that was introduced by The U.S. Equal Employment Opportunity Commission (EEOC). However, in the EU antidiscrimination law there isn't any statistical threshold published as in the US law,[542] disparate impact doctrine of which was an inspiration of the CJEU in developing the indirect discrimination doctrine in the 1980s when it was not introduced yet in the EU law,[543] nor do any of the fairness metrics developed in the technical community yet capture a European conceptualisation of discrimination which is highly contextual.[544]

The CJEU has relied on various wording to determine whether the protected group was affected by the seemingly neutral rule, criterion or practice in a significantly more negative way to establish prima facie discrimination. Examples from the Court's wording are "a far greater number of women than men,"[545] "a considerably lower percentage of men than women,"[546] "far more women than

---

In the US, in an indirect discrimination (disparate impact) case (i) the plaintiff must establish that the employer's selection procedure generates disparate impact; (ii) if established, then employer must defend itself by justifying the disparate impact by reference to some business need; (iii) if justified, then the plaintiff may challenge the justification to be faulty or demonstrate that an alternative practice exists that would serve the employer's business need equally while reducing the disparate impact provided that such practices are not too costly.

[542] For a comparison of the US and EU antidiscrimination laws, see. A Comparative Study on Algorithmic Discrimination between Europe and Northamerica, (Italian Equality Network, 2022), https://www.italianequalitynetwork.it/a-comparative-study-on-algorithmic-discrimination-between-europe-and-north-america?action=genpdf&id=1564.

[543] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

[544] Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai ".

[545] C-171/88 [1989] *Rinner-Kühn v FWW Spezial-Gebäudereinigung* ECLI:EU:C:1989:328

[546] C-33/89 [1990] *Maria Kowalska v Freie und Hansestadt Hamburg* ECLI:EU:C:1990:265; C-184/89 [1991] *Nimz v Freie und Hansestadt Hamburg* ECLI:EU:C:1991:50

men,"[547] "significantly high proportion of non-nationals, compared to nationals,"[548] "significantly greater proportion,"[549] "considerably lower percentage,"[550] and so on.[551] In *Rinner-Kühn*, 89 per cent of part-time workers being women in Germany was interpreted as a significant disproportion.[552] Here the examination of the Court was focussed on the disadvantaged group.

The effect of the contested rule has been examined by the CJEU on only the disadvantaged group, only the comparator group (advantaged group), or on both. In *R v Secretary of State for Employment, ex parte Nicole Seymour-Smith,* the CJEU had to formulate a legal test which is interpreted by Wachter et. al as setting a "gold standard." According to the test, the Court accepted that:

> "…the best approach to the comparison of statistics is to consider, on the one hand, the respective proportions of men in the workforce able to satisfy the requirement of two years' employment under the disputed rule and of those unable to do so, and, on the other, to compare those proportions as regards women in the workforce. It is not sufficient to consider the number of persons affected, since that depends on the number of working people in the Member State as a whole as well as the

---

[547] C-343/92 [1994] *De Weerd, Roks and Others v Bestuur van de Bedrijfsvereniging voor de Gezondheid, Geestelijke en Maatschappelijke Belangen and Others* ECLI:EU:C:1994:71

[548] C-703/17 [2019] *Krah v Universität Wien*, Opinion of Leger AG, ECLI:EU:C:2019:450

[549] C-161/18 [2019] *Villar Láiz* ECLI:EU:C:2019:382; C-527/13 [2015] *Cachaldora Fernández* ECLI:EU:C:2015:215

[550] C-33/89 [1990] *Maria Kowalska v Freie und Hansestadt Hamburg* ECLI:EU:C:1990:265; C-167/97 [1999] *Regina v Secretary of State for Employment, ex parte Nicole Seymour-Smith and Laura Perez* ECLI:EU:C:1999:60

[551] See. Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai " 29-30; Ellis and Watson, *Eu Anti-Discrimination Law*, 153; *Handbook on European Non-Discrimination Law*, 57.

[552] C-171/88 [1989] *Rinner-Kühn v FWW Spezial-Gebäudereinigung* Opinion of Darmon AG ECLI:EU:C:1989:158 para 31

percentages of men and women employed in that State."[553]

However, as criticised by Wachter et. al, the CJEU continued to not applying this ideal legal test and comparing the disadvantaged or advantaged only, despite the availability of statistics to go further. Even in *R v Secretary of State for Employment, ex parte Nicole Seymour-Smith,* it concluded that there was no indirect discrimination as statistics had shown 77.4% of men and 68.9% of women fulfilled that condition, by only looking at the advantaged group. According to Wachter et. al, while comparing only the advantaged or the disadvantaged groups can still be useful in capturing disadvantage experienced by groups that are evenly distributed in the society such as the percentage of women and men, it falls short to capture cases when the disadvantaged group is a minority or a heterogenous group[554] especially in cases of discrimination based on ethnicity and sexual orientation, as well as intersectional discrimination. They give the hypothetical example, if in a company's hiring practice, one only looks at who is not hired (the disadvantaged group) and it seems that no protected group was disproportionately disadvantaged the practice would prove non-discriminatory. However, if one looks into the advantaged group as well and finds out that those who are hired are all white men, the story would change entirely.[555] Similarly, in *Maniero,* the CJEU did not look into the composition of the advantaged group and found no indirect discrimination as the disadvantaged group did not seem to have any disproportionate effect on a given ethnicity.

Wachter et. al[556] derived from the case law of the CJEU that the Court applied two legal tests, in other words fairness methods in indirect discrimination cases.

---

[553] C-167/97 [1999] *Regina v Secretary of State for Employment, ex parte Nicole Seymour-Smith and Laura Perez* ECLI:EU:C:1999:60 para 59.

[554] In *Maniero* CJEU did not look into the composition of ethnicity.

[555] Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai " 41.

[556] Ibid., 51-52.

First one is demographic parity, while the second one is called by the authors negative dominance. Demographic parity is satisfied when the proportion of a protected group is the same both in the advantaged group and the disadvantaged group in consistency with the proportions in the general population. Negative dominance is derived from the CJEU's general definition of indirect discrimination used in various indirect discrimination cases that it occurs "when a seemingly neutral rule puts considerably more workers of one sex at a disadvantage than the other." Thus in negative dominance a twofold assessment is carried out. The majority of the disadvantaged group has to belong to a protected class;[557] and a minority of the protected group has to be part of the advantaged group.[558] According to Wachter et. al, negative dominance (i) provides less protection to minority groups such as in the cases of discrimination based on ethnicity and sexual orientation; (ii) makes a divide and conquer strategy possible for justification, if the disadvantaged group is divided into smaller subgroups to test the fairness of the system, (iii) turns the legal debate into a battle of numbers which may grant an easy win for the algorithmic systems. The writers suggest conditional demographic disparity as a method to capture algorithmic fairness compatible with the EU antidiscrimination law that does not fall short on where negative dominance does.[559]

In *R v Secretary of State for Employment, ex parte Nicole Seymour-Smith,*

---

[557] C-171/88 [1989] *Rinner-Kühn v FWW Spezial-Gebäudereinigung GmbH & Co. KG.* ECLI:EU:C:1989:328, in Report for the Hearing p.2746. %89 of the disadvantaged group was women. In this case the second part of the assessment is evident.

[558] C-167/97 [1999] *Regina v Secretary of State for Employment, ex parte Nicole Seymour-Smith.* Women's proportion (68.9%) in the advantaged group was very close to men's (77.4%), not showing only a minority of women was being advantaged which led the Court to conclude that indirect discrimination did not occur.

[559] Amazon started to implement this method in their bias and explainability software. <https://www.amazon.science/latest-news/how-a-paper-by-three-oxford-academics-influenced-aws-bias-and-explainability-software> Accessed on 30 September 2022.

CJEU also draw attention to the role of the national courts in assessing whether the statistical evidence submitted is relevant to reflect the population at the time of the case. In *Rinner-Kühn*, it was considered that the numbers can change in accordance with the time and place data is collected. Thus, the applicability of the data has to be considered when testing algorithmic systems for their statistical fairness.

Definition of proxy discrimination and indirect discrimination almost overlap at parts. As mentioned in Chapter 1, proxy discrimination occurs when a facially neutral classifier replaces a suspect classifier which falls under a protected ground and predicts the target variable. Indirect discrimination captures such situations where a neutral rule, criterion or practice disadvantages a person or a group defined by a protected ground. In *CHEZ*, electricity meters were placed higher from the ground compared to other service areas of an electricity company in a neighbourhood where the residents were predominantly of Roma origin. The Court found that when the neutral practice of placing meters higher happens in an area where predominantly Roma people resided, therefore affecting Roma people significantly more than others, it would be captured by indirect discrimination.[560] Here, the neighbourhood acts like a facially neutral proxy for ethnicity, and it is important that the notion of indirect discrimination captures its discriminatory effect.

As to the proxies that may predict ethnicity and race, in *Jyske Finans,*[561] as in *CHEZ*[562] the CJEU acknowledged that the concept of ethnicity has "its origin in the idea of societal groups marked in particular by common nationality, religious faith, language, cultural and traditional origins and backgrounds" and that this was

[560] C-83/14 [2015] *CHEZ Razpredelenie Bulgaria AD v Komisia za zashtita ot diskriminatsia* ECLI:EU:C:2015:480.

[561] At para 17.

[562] At para 46.

not an exhaustive list. Thus, a person's country of birth could be one of the factors to be taken into account in concluding on someone's ethnicity, however it was not by itself decisive. So that, while the CJEU accepted the link, refused that birth place as a proxy for ethnicity. In *Maniero,* similarly an equivalent foreign diploma would have been a proxy for foreign ethnicity in indirect discrimination analysis however, the Court did not find an adverse effect on a particular ethnicity. However, these cases are interpreted by some authors to signal the strength of indirect discrimination doctrine in responding to proxy challenge in dealing with algorithmic discrimination,[563] especially in complex protected grounds such as ethnicity, political opinions and sexual orientation. To be noted though, in *Jyske Finans*, decision of the CJEU was also interpreted to be problematic for not considering indirect discrimination even though the law against money laundering and terrorist financing with the specific identification, which was the contested rule, was likely to adversely impact persons who are not ethnically Danish significantly more than persons who are.[564] The Court could have found this out if it had compared the composition of the advantaged group.

Moreover in *CHEZ*, the applicant was not of Roma descent but she felt discriminated on this basis as she had a business in the same area. The Court followed its judgement in *Coleman,* that principle of equal treatment set out in Article 1 of the Race Directive, aims at protecting all persons as mentioned in Recital 16 of the Directive and does not apply to a particular category of person but by reference to the ground of racial and ethnic origin. On this account, the Court held that it applied to persons who are "not themselves a member of the race or ethnic group concerned, nevertheless suffer less favourable treatment or a

[563] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 73.

[564] Ward, "The Impact of the Eu Charter of Fundamental Rights on Anti-Discrimination Law: More a Whimper Than a Bang?," 50.

particular disadvantage on one of those grounds."[565] Thus, the CJEU established indirect discrimination by association,[566] which is one of the concerns in the way algorithmic discrimination operates. This allows persons, who feel discriminated based on a characteristic an algorithmic system incorrectly inferred that they hold, which also works as a proxy to a protected group membership, to bring an indirect discrimination case, provided that they can establish prima facie discrimination on the said protected ground.

2.2.2 Limitations of the European Antidiscrimination Law

As we introduced above, although it captures more algorithmic discrimination cases, indirect discrimination is even harder to detect and prove when the bias arises in machine learning models. There are several reasons for this. First of all, it is hard to establish a prima facie case of discrimination. Regulators and judges might not have an intuitive feeling with some proxies that cause discrimination against protected groups in the same way they had in the case of part-time work. When such automated discrimination is also not directly experienced or 'felt' by potential claimants,[567] and when access to (information about) the system is limited, knowing where to look and obtaining relevant evidence that could reveal prima facie discrimination will also be difficult. Secondly, even though prima facie discrimination could be detected, it is very challenging to point out the source of algorithmic bias so that it cannot be justified on the basis of predictive accuracy as the benefits and mathematical objectivity of decision making processes may

---

[565] C-83/14 [2015] *CHEZ* Para 56.

[566] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*; Wachter, "Affinity Profiling and Discrimination by Association in Online Behavioural Advertising (May 15, 2019). , Vol. 35, No. 2, 2020, Forthcoming, Available at Ssrn: Https://Ssrn.Com/Abstract=3388639 or Http://Dx.Doi.Org/10.2139/Ssrn.3388639."

[567] Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai ".

appear to be objective. Third, algorithms remain to be black boxes, especially to those who have not developed or have access to the code of the allegedly discriminating algorithm. According to the ruling of CJEU in Meister, antidiscrimination law does not grant access rights.[568] Thus, even if there is a suspicion of discriminatory practice, bringing up a case would be challenging. We will have a closer look into these difficulties below.

2.2.2.1 Limitation to Protected Classes

As mentioned earlier, as long as it does not form a proxy for a historically disadvantaged group, we do not think discrimination based on groups that are outside of the protected classes such as people being discriminated based on their device type is too much of a priority. In this case android users cannot claim to be discriminated based on this ground, however the treatment they receive can still be unfair.[569] Moreover, it is also important to know, as pointed out by Xenidis, which future decisions these emergent forms of discriminatory grounds inform, which chances they deny to those who associate with them, into what kind of classifications and predictions do they turn in time, how does the knowledge they produce shape the future? Are their predictions based on nonintuitive[570] correlations treated like causal links that inform future behaviours, get caught up in feedback loops and eventually turn into self-fulfilling prophecies? When the societies they operate in cannot have a democratic oversight on their workings to evaluate whether they are socially acceptable or desirable or should they be allowed, especially if they are becoming the foundation of important decisions, we

---

[568] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

[569] Janneke H. Gerards and Frederik J. Zuiderveen Borgesius, "Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence," *EngRN: Electronic* (2020): 8.

[570] We use the term as explained in Selbst and Barocas, "The Intuitive Appeal of Explainable Machines," 1091. at footnote 30.

agree that antidiscrimination law should be able to provide protection to those who are going to be adversely affected.[571] As Custers et al have noted, data mining will unintentionally use proxies that are socially unacceptable, as hard it is to be discovered, new groups that are facing stereotyping and seclusion, may not even know the adverse effects they are experiencing and even if they do, they may be so scattered as a population they may lack the political power to challenge these.[572]

On defence of unprotected groups may still face unfair treatment, Gerards and Borgesius give three examples. In their first example, they suggest the unfairly treated group might be composed of low-income individuals, therefore reinforce structural inequalities and punish the poor. We are in the opinion that this example does not really reflect a new algorithmic group problem but it is worth questioning why discrimination on economic-status is underexplored under EU antidiscrimination law.[573] Later in their article authors point out current discussions on whether to add economic-status as a protected ground and discuss in a semi-closed list of grounds would allow it to be considered by the courts in time as more discrimination by AI on the basis of economic status is brought to courts. The second example is seemingly irrelevant criteria, as in a real example from Netherlands, when an insurance company charged people whose house number included a letter next to a number such as 1A, 21C and so on to be more likely in a car accident, based on a correlation captured in the historical data rather than any causal link. Third example is about incorrect predictions. This may occur when a number of attributes which do not fall under a protected ground are found together in a client, the model of a bank predicts a high probability of default based on the

---

[571] Raphaële Xenidis, "Tuning Eu Equality Law to Algorithmic Discrimination: Three Pathways to Resilience," *Maastricht Journal of European and Comparative Law* 27, no. 6 (2020): 754.

[572] Bart Custers et al., "The Way Forward," in *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases*, ed. Bart Custers, et al. (Springer Publishing Company, Incorporated, 2013), 352-53.

[573] Sarah Ganty, "Poverty as Misrecognition: What Role for Antidiscrimination Law in Europe?," *Human Rights Law Review* 21, no. 4 (2021).

correlations it found in the historical data that 95 percent of clients who carry all these attributes default. A person who is in the 5 percent and never misses a payment would feel unfairly treated.[574]

After a detailed analysis of different systems that deal with non-discrimination, Gerards and Borgesius find out best fit to answer the challenges brought by algorithmic discrimination is a hybrid system with a semi-closed list of grounds and open possibility of exemptions, which would offer legal clarity for the companies who would align their systems in order to refrain from legal consequences while the suspect groups would be more aware of their rights. Also the source of protection would come from the democratic basis, rather than leaving it to the scrutiny of unelected judges. The exemptions shall be left open as the authors conclude that it would not be possible to scratch out every possibility these new technologies may bring. Article 21 CFREU and Article 14 ECHR are found to be in this category by the authors, with their limitations of the first being only applicable in areas covered by the EU law, while latter is limited to the exercise of other rights under the ECHR and cannot always be invoked based on the national system where the discrimination arises.[575]

However, currently EU antidiscrimination law provides protection on an exhaustive list of categories, that are not fit to capture the dynamism of classifications carried out by machine learning algorithms.[576]

First of all, CJEU's current case law raises more questions than providing

---

[574] Gerards and Zuiderveen Borgesius, "Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence."

[575] Ibid., 67.

[576] Xenidis, "Tuning Eu Equality Law to Algorithmic Discrimination: Three Pathways to Resilience."

clarity on tackling proxy discrimination. In *Jyske Finans*,[577] the Court did not find birth place decisive enough to predict ethnic origin by itself, even though it may be among the factors that may presume it, despite Article 21(1) CFREU explicitly prohibits discrimination on the basis of birth and is higher in the hierarchy of norms than the Framework Directive. This approach is clearly at odds with how proxy discrimination functions. When a model produces a discriminatory output based on a facially neutral input variable that correlates with a protected ground, it is yet unclear how that link can be measured to be decisive enough.

Furthermore, the CJEU have rejected to extend the scope of the Framework Directive by analogy beyond the discrimination based on the exhaustively listed grounds in the directive and ruled that this was so despite the inexhaustive list given in CFREU Article 21[578] which is a primary source of EU law since 2009. If the CJEU followed the approach of the ECtHR in applying the "any other status" in Article 14 ECHR,[579] after an inexhaustive list of grounds were introduced in CFREU Article 21, this would have armed the EU antidiscrimination law much better against proxy discrimination.

In its current system indirect discrimination is the best fitting tool that the EU antidiscrimination law has, however it is still bound with the material scope of the equality directives, as well as the protected grounds, while being open to much more chances of justification than direct discrimination for the challenged algorithmic systems. We will explore the justification problem in more detail below.

---

[577] C-668/15 [2017] *Jyske Finans A/S v Ligebehandlingsnævnet, acting on behalf of Ismar Huskic* ECLI:EU:C:2017:278 para 17-23.

[578] See C-13/05 *Chacón Navas* para.56; C-354/13 *Kaltoft* para. 36.

[579] Prisoners in *Aydin v. Turkey* App no 23178/94 (25 September 1997); asylum seekers in *M.S.S. v. Belgium and Greece* App no 30696/09 (21 January 2011); HIV survivors in *Kiyutin v. Russia* App no 2700/10 (10 March 2011).

2.2.2.2 Domain Based Limitations

Under the EU antidiscrimination law not every ground as protected equally, which has been called out to create a hierarchy of protection.[580]

Protection against discrimination based on ***sex*** is provided in Recast Directive, in the area of employment and occupation; and in Goods and Services Directive in access to and supply of goods and services.[581]

Recast Directive covers equal treatment of men and women in relation to access to employment, including promotion, and to vocational training; working conditions, including pay; and occupational social security schemes.[582] It also covers any less favourable treatment of a woman related to pregnancy or maternity leave.[583] There are additional gender equality directives in this area,[584] alongside with the direct protection provided by the Treaties.

The Goods and Services Directive applies to goods and services, which are made available to the public by both the public and private sectors, including public bodies.[585] However it does not apply to the content of media and advertising nor to education,[586] it also leaves out matters of employment and occupation,[587] as

---

[580] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

[581] 2004/113/EC

[582] 2006/54/EC, Article 1.

[583] 2006/54/EC, Article 2(2).

[584] Pregnancy Directive (92/85/EEC), the Parental Leave Directive (2010/18/EU), the Part-time Work Directive (97/81/EC).

[585] 2004/113/EC, Article 3(1)

[586] 2004/113/EC, Article 3(2)

[587] 2004/113/EC, Article 3(3)

well as self-employment.[588] It is already too limited that protection in access to goods and services is limited to the ground of sex, but also many discrimination cases resulting from automated decision-making will arise from services provided online.[589]

One important loophole comes from the definition of services which are normally provided for remuneration under EU Law according to TFEU Article 57. Many online services are provided for free if one doesn't include personal data in the definition of renumeration[590] as the infamous saying "if something is free, you are the product" is regularly used to explain advertisement revenue based services provided for free to the users, however the real service they provide is delivering people to the advertisers who are their real customers. According to Zuboff it is an incorrect analogy as it goes beyond providing eyeballs for advertisement, and is more of a "raw-material-extraction operation" for future behaviour to be traded in the markets of surveillance capitalism.[591] As payment by third parties is included in renumeration,[592] Hacker concludes such services would be covered by the Directive.[593]

---

[588] 2004/113/EC, Article 3(4)

[589] Wulf, *Automated Decision-Making Systems and Discrimination: Understanding Causes, Recognizing Cases, Supporting Those Affected*

[590] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 12. See von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Iii," 381. for an opposing opinion. According to von Grafenstein, data subjects exchange a risk to their fundamental rights against a data-driven service or product.

[591] Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power,* 10.

[592] Case 352/85 [1988] *Bond van Adverteerders and others v The Netherlands* State ECLI:EU:C:1988:196 at para 16.

[593] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 13.

Another problem may arise from the clash of microtargeted advertisement, prices, and contractual provisions with the public availability requirement for the application of the Directive, and leave highly personalized services outside its scope. According to Hacker, these personalized services are beyond the reason behind the limitation to publicly available good and services, such as someone selling their old phone to a friend, should not be brought to court on the basis of discrimination. Thus, irrespective of the personalized content, offer or contract, as far as providers of highly personalized services are generally in the intention of providing their services to the public, then it is just the content of the goods and services that may diverge, thus they ought to be covered by the Directive.[594]

As the content of media and advertising is outside the scope of the Directive, suppliers of goods and services may differentiate in targeting men or women in the content of their advertisement, however the goods and services themselves should be offered under same terms to both men and women.[595] However such gender versioning may also open way to justification of differential treatment under Article 4(5)[596] if there is a legitimate aim and it is proportionate. Even though the scope of this study is limited to job ads delivery in terms of advertisement, these practices may have stereotyping results which may affect all areas including employment. We will discuss job ads delivery more in detail below.

Recast Directive, Race Directive and Framework Directive, all cover (i) conditions for access to employment, to self-employment or to occupation, including selection criteria and recruitment conditions, including promotion; (ii) access to all types and to all levels of vocational guidance, vocational training, including practical work experience; (iii) employment and working conditions,

---

[594] Ibid.

[595] Ellis and Watson, *Eu Anti-Discrimination Law*, 368.

[596] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 62.

including dismissals, as well as pay; (iv) membership of, and involvement in, an organisation of workers or employers, or any organisation whose members carry on a particular profession.[597]

Protection against discrimination based on ***race and ethnicity*** is covered under the Race Directive. It has the widest scope which covers in addition to abovementioned areas (i) social protection, including social security and healthcare; (ii) social advantages; (iii) education; (iv) access to and supply of goods and services, including housing.[598] It does not cover difference of treatment based on nationality as it is covered elsewhere.[599] As it does not mention any exlusion regarding the content of media and advertising, it is interpreted that this area is also covered with respect to the ground of race and ethnic origin.[600]

Discrimination based on the grounds of ***religion or belief, disability, age or sexual orientation*** are prohibited under the Framework Directive which lays down a general framework for combating discrimination on these grounds as regards employment and occupation[601] both in public and private sectors.[602] It does not apply to payments of any kind,[603] and leaves out nationality as it is covered elsewhere.[604]

---

[597] 2006/54/EC, Article 14; 2000/43/EC, Article 3(1); 2000/78/EC, Article 3(1).

[598] 2000/43/EC, Article 3(1).

[599] 2000/43/EC, Article 3(2). Protection on the basis of nationality is a very complex issue under the EU law that is too wide to cover which is the reason that we will not delve into any detail regarding this ground under this study.

[600] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

[601] 2000/78/EC, Article 1.

[602] 2000/78/EC, Article 3(1).

[603] 2000/78/EC, Article 3(3).

[604] 2000/78/EC, Article 3(2).

As mentioned above, employment is an area where the most protection is offered. All of the abovementioned directives cover conditions for access to employment including selection criteria and recruitment conditions. As job advertising is a part of the recruitment process[605] it is included within the scope of all the above-mentioned directives and discrimination in this area on grounds of sex, race and ethnic origin, religion or belief, disability, age and sexual orientation is prohibited. Case law of the CJEU is supportive of this conclusion as even in the absence of an actual recruitment process it has ruled that deterring job applicants from protected grounds to be considered discrimination.[606]

Another area of employment, where solely automated decision-making is common, is the so-called gig economy. Platform workers' pay is often determined by algorithms based on various variables including the feedback in terms of likes or stars or reviews given by customers, workers' availability and flexibility and so on as these algorithms also remain black boxes, details of their workings have not been discovered. However, studies indicate to a digital gender pay gap[607] to be in the forming, unsurprisingly as structural disadvantages of women do not fit the alleged criteria of these algorithms.

As for the ever growing gig economy, work conditions, promotions, dismissals and other employment related rights and benefits might also be subject to ADM. Many of these areas would be captured by the equality directives, once an employment relationship is established. First of all, in *Allonby*, the CJEU ruled

---

[605] Post and Holtmaat, "A False Start: Discrimination in Job Advertisements."

[606] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*. See C-54/07 [2008] *Feryn* ECLI:EU:C:2008:397; C-81/12 [2013] *Asociaţia Accept* ECLI:EU:C:2013:275; C-507/18 [2020] *NH v Associazione Avvocatura per i diritti LGBTI — Rete Lenford* ECLI:EU:C:2020:289.

[607] Arianne renan barzilay and Anat Ben-David, "Platform Inequality: Gender in the Gig-Economy," *Seton Hall Law Review* 47 (2017); Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

that it will apply the definition of a worker under EU law even if it is defined differently under the national law, and that a person classified as self-employed under the national law may still be classified as an employee under the EU law if "independence is merely notional, thereby disguising an employment relationship."[608]

In *Asociación Profesional Élite Taxi,* the CJEU decided that UBER was not providing 'an information society service' as an intermediary within the meaning given in E-Commerce Directive (2000/31/EC) but a transport service as it "…exercises decisive influence over the conditions under which that service is provided by those drivers. On the latter point, it appears, inter alia, that Uber determines at least the maximum fare by means of the eponymous application, that the company receives that amount from the client before paying part of it to the non-professional driver of the vehicle, and that it exercises a certain control over the quality of the vehicles, the drivers and their conduct, which can, in some circumstances, result in their exclusion."[609] Thus, Hacker concludes, platforms with an "intermediation plus control" business model are covered by antidiscrimination law in relation to the substantive service they intermediate, such as online recruitment agencies.[610] In *Bougnaoui,* the CJEU ruled that preferences of customers cannot serve as a justification in employment related discrimination cases.[611] While these developments in case law indicate to most ADM processes used in HR management in gig economy would fall under the scope of EU Antidiscrimination Law, legislative efforts of the Commission seems like

---

[608] C-256/01 [2004] *Allonby v Accrington & Rossendale College* ECLI:EU:C:2004:18 para 71.

[609] C-434/15 [2017] *Asociación Profesional Élite Taxi v Uber Systems Spain SL* ECLI:EU:C:2017:981 para 39-40.

[610] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 16.

[611] C-188/15 [2017] *Bougnaoui and Association de défense des droits de l'homme (ADDH) v Micropole SA* ECLI:EU:C:2017:204

clarifying where this most pervasive use of HR management sits in EU law soon.[612]

However, discrimination based on the grounds of religion or belief, disability, age or sexual orientation is not protected in access to goods and services including in advertisement. In addition, grounds that are not listed in the closed-list system of Equality Directives that can be very determining for the unfavourable treatment such as economic status remain problematic. For the least, CFREU Article 21 could be used by the CJEU in remedying this hierarchy of protection problem, in order to address discrimination arises in one of the domains falling under an equality directive but the ground on which the discrimination is based is not protected under that directive.[613]

2.2.2.3 Failure to acknowledge intersectionality and the comparator problem

Intersectional discrimination occurs when more than one protected ground intersect to form a different type of discrimination which is not well-addressed by the CJEU.[614] This kind of discrimination is very likely to occur in machine-led profiling.

The term intersectionality was coined by Crenshaw in order to explain the

---

[612] European Union: European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Better Working Conditions for a Stronger Social Europe: Harnessing the Full Benefits of Digitalisation for the Future of Work, 9 December 2021, Com(2021) 761 Final* (European Union: European Commission, 2021); *Proposal for a Directive of the European Parliament and of the Council on Improving Working Conditions in Platform Work, 9 December 2021, Com(2021) 762 Final* (European Union: European Commission, 2021); ibid.

[613] Xenidis, "Tuning Eu Equality Law to Algorithmic Discrimination: Three Pathways to Resilience."

[614] Dagmar Schiek, "On Uses, Mis-Uses and Non-Uses of Intersectionality before the Court of Justice (Eu)," *International Journal of Discrimination and the Law* 18, no. 2-3 (2018).

discrimination experienced on multiple grounds to be a completely different and an amplified experience than one dimensional discrimination. She criticized the blindness of the anti-discrimination law and the courts to such discrimination suggesting that black women experienced a different kind of discrimination than black men and white women.[615] Similar results are shown by computer scientists with respect to machine learning algorithms, such as the Gender Shades study that showed how facial recognition algorithms were significantly less accurate for darker skin women than lighter skin women and most accurate for light skinned man,[616] it also reflects the matrix of domination.[617] As in real life interaction of womanhood and blackness may give rise to a completely new type of experience, interactions between labels in a system may also give rise to new categories.[618]

The concept of discrimination based on more than one ground does not have a settled terminology in law yet. It has been interchangeably termed as multiple discrimination, cumulative discrimination, compound discrimination, combined discrimination and intersectional discrimination, although their meaning slightly differs.[619] Fredman identified three main ways discrimination may occur on more than one ground. First, in an employment context, an employee may be subject to racist treatment and unfavourable treatment on their disability in separate occasions which led to their dismissal. In such cases, both grounds are separately taken into consideration by courts. Second, a lesbian woman may claim that on an

---

[615] Crenshaw, "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics."

[616] Buolamwini and Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification."

[617] Patricia Hill Collins, *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*, Second ed. (New York and London: Routledge, 2000).

[618] Anna Lauren Hoffmann, "Where Fairness Fails: Data, Algorithms, and the Limits of Antidiscrimination Discourse," *Information, Communication & Society* 22 (2019): 906.

[619] Commission European et al., *Intersectional Discrimination in Eu Gender Equality and Non-Discrimination Law* (Publications Office, 2016).

occasion she was harassed because of her sexual orientation and her gender. In these cases, two grounds are additive to each other but a court will still apply two grounds separately, that discrimination on each ground has to be proved individually. Finally, more than one ground come together and create a completely different experience of discrimination that cannot be captured by applying two merging grounds separately, nor can it be proved on each ground separately. For example, in a dismissal process black women were made redundant first, this is a different experience than black men and white women experience, even if the whole percentage of those who were made redundant are black women, it cannot be proven in comparison to women and in comparison to black people because the number of dismissals are neither disproportionate among women nor among black employees.[620] This last type of discrimination is what we mean with intersectional discrimination.

In the European context, ECtHR has been considering intersectionality without calling it by its name.[621] In *N.B. v. Slovakia*,[622] ECtHR did not find it necessary to decide on both grounds of sex and ethnicity under Article 14 ECHR after finding out there was discrimination on grounds of ethnicity in conjunction with a violation of Article 8, despite the claim of the plaintiff was discrimination on both grounds. In *B.S. v. Spain*,[623] two NGOs acting as the third party interveners brought up intersectional discrimination as discrimination on multiple grounds. ECtHR found violation of Article 14 in conjunction with Article 3 as "the domestic courts failed to take account of the applicant's particular vulnerability inherent in her position as an African woman working as a prostitute." Even though the Court gave signals of considering intersectional discrimination, it did not mention a

[620] Ibid.

[621] *Handbook on European Non-Discrimination Law*, 60.

[622] N.B. v. Slovakia (Application no. 29518/10) 12 June 2012 https://hudoc.echr.coe.int/fre?i=001-111427

[623] B.S. v. Spain (Application no. 47159/08) 24 July 2012 https://hudoc.echr.coe.int/eng?i=001-112459, see para 62.

violation of Article 14 both on grounds of race and sex. In *S.A.S. v. France*, the ECtHR was made aware of the intersectional effect again by intervening NGOs, thus even though the Court considered gender and religious belief grounds together as "the ban prevents certain women from expressing their personality and their beliefs," it found no violation of the Article 14 as the ban was justified.[624] Finally, in *Carvalho Pinto De Sousa Morais v. Portugal*, the ECtHR found violation of article 14 in conjunction with Article 8 as "the applicant's age and sex appear to have been decisive factors in the final decision, introducing a difference of treatment based on those grounds."[625] In this case intersectional discrimination within the true meaning of it as explained above was found without mentioning multiple discrimination and intersectional discrimination again, as the national court decision under scrutiny was exactly discriminatory only for women above a certain age, having combined effect on these two grounds.

In the EU, Recital 14 of the Race Directive[626] and Recital 3 of the Framework Directive,[627] state that "women are often the victims of multiple discrimination." Also the new gender equality strategy for 2020-2025 takes intersectionality in its centre.[628] However, CJEU has been restrictive in applying combined effect. In recent two cases *Achbita*[629] and *Bougnaoui*,[630] the Court considered Muslim

---

[624] S.A.S. v. France (Application no. 43835/11) 1 July 2014 https://hudoc.echr.coe.int/eng?i=001-145466 , see para 153.

[625] Carvalho Pinto De Sousa Morais v. Portugal (Application no. 17484/15) 25 July 2017 https://hudoc.echr.coe.int/eng?i=001-175659, See para 53.

[626] Directive 2000/43/EC

[627] Directive 2000/78/EC

[628] European Union: European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic Tnd Social Committee and the Committee of the Regions - a Union of Equality: Gender Equality Strategy 2020-2025, 5 March 2020, Com(2020) 152 Final* (European Union: European Commission, 2020).

[629] C-157/15 [2017] *Achbita, Centrum voor gelijkheid van kansen en voor racismebestrijding v G4S Secure Solutions NV* ECLI:EU:C:2017:203

[630] C-188/15 [2017] *Bougnaoui and ADDH* ECLI:EU:C:2017:204

women's discrimination based on wearing headscarf as discrimination against religion even though the plaintiffs brought a combined effect case as it is indeed a special kind of discrimination only being experienced by Muslim women who wear headscarves, as Muslim men do not.[631]

This brings first of all the comparator problem that we have already introduced above in 2.2.1. In the original example of Crenshaw the comparator problem was examined as the reason for the black women's failure in winning discrimination cases, as they had to either be compared to black men regarding discrimination based on race or to white women regarding discrimination based on gender. Imagine, a recruitment algorithm for the selection stage discovers women aged between 30 to 40 from a specific neighbourhood is likely to have multiple breaks in their careers and weeds them out by using an algorithm that gathered data on online behaviour for the purposes of online advertising. This neighbourhood may be proxy for an ethnic group and women who have careers of this group may be giving birth more than once during this period. Age, sex and ethnicity becomes intertwined to create a single experience that is then marginalized by an algorithm.

One of the main problems related to this example in the EU antidiscrimination law is what we discussed above as the domain-based limitations and the hierarchy of protected grounds that stem from that. We will take an example given by Fredman. Let's say older Turkish immigrants living in Berlin are experiencing a disproportionately negative effect as a result of a housing policy that is led by gentrification. Discrimination based on ethnicity is covered by the Race Directive which provides protection regarding housing, while age is only protected under the Framework Directive which only covers employment and vocational training. If the young residents of Turkish decent in the area are not similarly affected from the same policy, however old Germans were also being adversely affected, because they cannot rely on age and cannot prove discrimination on ethnicity, old

[631] Schiek, "On Uses, Mis-Uses and Non-Uses of Intersectionality before the Court of Justice (Eu)."

Turkish residents cannot redress the detriment they suffered.[632]

If the different dimensions of the unfavourable treatment is covered by the equality directives and could be brought before the Court, then prima facie discrimination has to be established. The lack of disaggregated data by sex and race, ethnicity and disability, alongside with the comparator problem makes it harder for intersectional discrimination claims.[633] It is best described by the basement example of Crenshaw, that the single-axis understanding of discrimination and these single categories being developed on the experiences of the most advantaged of these disadvantaged categories, marginalize the experiences of those who are suffering from multiply disadvantages. They are only protected when their experiences come closer to those who are more privileged under that protected ground such as white women or black men.[634] And as it was pointed out by Wachter et. al as an act of divide and conquer,[635] the smaller of a group they form, the more peculiar their experience appear, therefore it may lead to conclusions of no discrimination. So that empowering ways of collecting and making sense of disaggregated data reflecting these experiences should be the one of the steps to be taken.[636] Data is the new oil sort of approach to data extraction and leaving making sense of such data to correlations, certainly is not the empowering way.[637]

---

[632] European et al., *Intersectional Discrimination in Eu Gender Equality and Non-Discrimination Law*, 63.

[633] Ibid., 28.

[634] Crenshaw, "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics," 151.

[635] Wachter, Mittelstadt, and Russell, "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai " 69.

[636] For examples and suggestions thereto, see D'Ignazio and Klein, *Data Feminism*.

[637] Hoffmann, "Where Fairness Fails: Data, Algorithms, and the Limits of Antidiscrimination Discourse," 909.

In *Z.*, a woman who was born without an uterus decided to become mother via surrogacy and denied maternity leave by her employer as she had not been pregnant. Having looked into them separately, the CJEU found discrimination neither based on sex (because "a commissioning father who has had a baby through a surrogacy arrangement is treated in the same way") [638], nor disability (because lacking an uterus and having a baby through surrogacy as a result do "not hinder full and effective participation in professional life on an equal basis with other workers").[639] She was also denied protection under Directive 92/85 because, she could not, "by definition, be subject to less favourable treatment related to her pregnancy, given that she has not been pregnant with that baby."[640]

In *Parris*, the intersectional discrimination claim was straightforward and based on age and sexual orientation. [641] Mr. Parris would have been included in a survivor pension if he had married before turning 60, which he couldn't as by that time same sex partnership was not legalized. CJEU's reasoning in this case in rejecting discrimination both based on sexual orientation and age separately was problematic already. However, the Court's rejection on the basis of intersectional discrimination was much more unfortunate. The Court first accepted that, discrimination could be based on multiple grounds under the Framework Directive. Then it ruled, that a new combined category of discrimination could not arise if there was no discrimination found on the basis neither of the grounds in isolation.[642] However, Kokott AG's interpretation in *Parris* was on point with intersectionality literature and such judicial analysis alongside with acknowledgements as in recitals of Race and Framework Directives, as well as

---

[638] C-363/12 [2014] *Z. v A Government department and The Board of management of a community school* ECLI:EU:C:2014:159 para. 50

[639] Ibid. Para 80.

[640] Ibid para. 57.

[641] C-443/15 [2016] *David L. Parris v Trinity College Dublin and Others* ECLI:EU:C:2016:897 para 80-81.

[642] Ibid para 79-82

endorsements such as in the Gender Equality Strategy for the 2020-2025 are interpreted as hope-giving developments.[643]

As discussed above, it is once more suggested that making use of Article 21 CFREU's inexhaustive list of protected grounds would make it possible to create new protected grounds that is more fitting to the context of the case at hand, especially in discrimination claims regarding intersectionality. This would then help to assign a more fitting comparator group and thus to be captured by the antidiscrimination law.[644]

Finally, Dagmar Schiek's proposal of reorganizing EU antidiscrimination law in a heterarchical way, instead of its current hierarchical way, around the nodes of race, gender and disability, can be a good way out from the limitations regarding static single axis grounds of protection, answering to intersectional discrimination and the proxy challenge.[645]

2.2.2.4 Difficulties to Seek Redress

In this part we will identify difficulties to seek redress against algorithmic discrimination under EU antidiscrimination law.

---

[643] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*, 142.

[644] Ibid., 141.

[645] Dagmar Schiek, "Organising Eu Equality Law around the Nodes of 'Race', Gender and Disability," in *European Union Non-Discrimination Law and Intersectionality : Investigating the Triangle of Racial, Gender and Disability Discrimination*, ed. Dagmar Schiek and Anna Lawson (London and New York: Routledge, 2011); Xenidis, "Tuning Eu Equality Law to Algorithmic Discrimination: Three Pathways to Resilience."

2.2.2.4.1 Individual Redress Mechanism

The fundamental right to equality and not to be discriminated are designed as individual rights. This brings down the existence of anti-discrimination law to protect individual victims of discrimination if they are eager to take active measures to challenge their personally experienced case of injustice.[646] This is not to be taken as an entirely invaluable opportunity to challenge inequalities in the society, however it may prove quite inadequate against algorithmic discrimination for a couple of important reasons we will briefly lay down below.

As discussed above, EU antidiscrimination law has developed some strong tools and concepts that may capture algorithmic discrimination.

EU Law differs from classical international law, that it actually grants rights to individuals that certain provisions of the EU law can be enforced in the national Courts.[647] According to the CJEU, "the vigilance of individuals concerned to protect their rights amounts to an effective supervision."[648] This is called direct effect, and Article 157 (1) of the TFEU on equal pay for equal value having direct horizontal effect[649] (meaning that it was not only invokable against the state by individuals but also against other individuals)[650] has played a crucial role in the

---

[646] Kristina Koldinská, "Eu Non-Discrimination Law and Policies in Reaction to Intersectional Discrimination against Roma Women in Central and Eastern Europe," in *European Union Non-Discrimination Law and Intersectionality : Investigating the Triangle of Racial, Gender and Disability Discrimination*, ed. Anna Lawson and Dagmar Schiek (Burlington, 2011), 253.

[647] Ellis and Watson, *Eu Anti-Discrimination Law*, 52.

[648] Case 26-62 [1963] *Van Gend & Loos v Netherlands Inland Revenue Administration* ECLI:EU:C:1963:1, p.13

[649]Case 43-75 [1976] *Defrenne v Sabena* ECLI:EU:C:1976:56

[650] See <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:l14547> Accessed on 30 September 2022.

development of EU antidiscrimination law.[651]

Direct effect of Treaty provisions over conflicting national legislation is a result of the supremacy of the EU law.[652] Unlike directly applicable provisions of the primary law (which require to be precise, unconditional and to not need implementation)[653] of the EU and regulations; directives, which constitute pretty much the entirety of EU law addressing antidiscrimination, require to be transposed to national law first. Furthermore, as a result of being a compromise they may leave discretion to Member States, as well as not being precise. On the other hand, legal integration and effectiveness of EU law requires that directives which regulate very important areas of the EU law to be also properly implemented and Member States frequently failed in this. Thus the CJEU ruled directives to be also capable of direct effect,[654] reasoning that (i) if individuals can rely on them, directives would be more effectively enforced, (ii) according to Article 267 TFEU, national courts can refer questions regarding directives to the CJEU even in cases between two individuals,[655] (iii) a member state who failed to implement a directive within its implementation period cannot deny the binding effect of the directive to individuals based on its own wrongdoing as otherwise the individual would have been able to rely on the national law.[656]

While vertical direct effect of equality directives is an advantage for the victims

---

[651] Xenidis and Senden, "Eu Non-Discrimination Law in the Era of Artificial Intelligence: Mapping the Challenges of Algorithmic Discrimination," 23.

[652] C-6/64 [1964] Flaminio Costa v E.N.E.L. ECLI:EU:C:1964:66 See. Ellis and Watson, *Eu Anti-Discrimination Law*, 45-52; Paul Craig and Gráinne de Búrca, *Eu Law : Text, Cases, and Materials*, 6. edition ed. (Oxford: Oxford University Press, 2017), 266-314. for a more detailed view on the supremacy doctrine of the EU law.

[653] *Eu Law : Text, Cases, and Materials*, 201.

[654] Ibid.

[655] Case 41-74 [1979] *Van Duyn v Home Office* ECLI:EU:C:1974:133

[656] Case 148/78 [1979] *Ratti* ECLI:EU:C:1979:110

discriminated by algorithms used by public bodies, many algorithmic discrimination cases will involve private companies especially in our lead examples of human resources and job ads delivery. Unfortunately in these cases, individuals cannot directly invoke equality directives as directives lack horizontal direct effect.[657] As regards the supremacy of the EU law, it is established that a Member State shall not adopt contradictory laws after a directive comes into effect, during the implementation period of a directive contradictory domestic law remains applicable, however new contradictory laws shall not be enacted. Once the implementation period is over, contradictory domestic law becomes inapplicable.[658]

Importantly, the CJEU granted horizontal direct effect to general principles of law,[659] that take their source from international instruments to which Member States signed up (most prominently ECHR, and European Social Charter), constitutional traditions and laws of the Member States, The Community Charter of Fundamental Social Rights for Workers, and the CFREU.[660] In *Mangold*, transposition period of the Framework Directive was not yet exhausted, however, the CJEU decided that non-discrimination on grounds of age was a general principle of law which was expressed in the Framework Directive. In *Küçükdeveci*, the CJEU ruled that if a case before the national court falls under the scope of the EU law, the national court must disapply the national law that

---

[657] C-152/84 [1986] *Marshall v Southampton and South-West Hampshire Area Health Authority (Teaching)* ECLI:EU:C:1986:84; C-91/92 [1994] *Paola Faccini Dori v Recreb Sri* ECLI:EU:C:1994:292; C-201/02 [2004] *Wells v Secretary of State for Transport, Local Government and the Regions* ECLI:EU:C:2004:12.

[658] C-129/96 [1997] *Inter-Environnement Wallonie ASBL v Région wallonne* ECLI:EU:C:1997:628; C-144/04 [2005] *Werner Mangold v Rüdiger Helm.* ECLI:EU:C:2005:709; C-555/07 [2010] *Seda Kücükdeveci v Swedex GmbH & Co. KG.* ECLI:EU:C:2010:21 Ellis and Watson, *Eu Anti-Discrimination Law*, 66.

[659] Craig and Búrca, *Eu Law : Text, Cases, and Materials*, 194.

[660] Ellis and Watson, *Eu Anti-Discrimination Law*, 103.

contradicts with the general principle of non-discrimination. Here, the discrimination on grounds of age in employment context fell under the scope of the EU law as it was regulated in the Framework Directive. In *Dansk Industri*,[661] the CJEU confirmed that general principle of non-discrimination on age had horizontal direct effect, and then in *Egenberger* on grounds of religion or belief by interpreting the Framework Directive together with the CFREU as it lays down the general principle.[662]

From the abovementioned case law of the CJEU, two crucial points remain vague while another crucial point becomes clear. First of all, it remains unclear whether individuals who felt discriminated by an algorithmic system in a private relation may invoke the entirety of the antidiscrimination law which is mostly regulated by directives. Second, it is not yet clearly established whether general principles of law may extend the protected grounds set out in equality directives as CFREU Article 21 and the ECHR has a more unexhaustive approach. Third, it is clear that the reliance on individuals to bring cases of algorithmic discrimination without any clarity on the chances of success is not an ideal strategy, especially as they are expected to discriminate in never before seen speed and scale. Moreover, currently the public, supervisory level institutions and tools such as Article 258 TFEU, and Union level equality and non-discrimination agencies European Institute of Gender Equality (EIGE) and the Fundamental Rights Agency (FRA) do not have enforcement power.[663]

---

[661] C-441/14 [2016] *Dansk Industri (DI) v Estate of Karsten Eigil Rasmussen* ECLI:EU:C:2016:278 para 35-36

[662] C-414/16 [2018] *Vera Egenberger v Evangelisches Werk für Diakonie und Entwicklung eV* ECLI:EU:C:2018:257 at para 76 which reads as "The prohibition of all discrimination on grounds of religion or belief is mandatory as a general principle of EU law. That prohibition, which is laid down in Article 21(1) of the Charter, is sufficient in itself to confer on individuals a right which they may rely on as such in disputes between them in a field covered by EU law."

[663] Xenidis and Senden, "Eu Non-Discrimination Law in the Era of Artificial Intelligence: Mapping the Challenges of Algorithmic Discrimination," 26-27.

Asking who questions such as who benefits or who cannot is a good feminist method to reveal power disparities.[664] On this account, individual redress is not likely to benefit the most disadvantaged because of the costs of legal proceedings, access to professionals and time needed.[665] Cathy O'Neil gives the example of Kyle Behm, who was rejected by an hiring algorithm used at screening stage to explain that decisions about masses (or about the poor as Eubanks[666] puts it) are being automated, the privileged will continue to enjoy individualised treatment provided by humans. In Kyle Behm's case if his friend hadn't heard the reason he was rejected and his father hadn't been a lawyer, he couldn't have known the discrimination he experienced and even if he did he might have not been able to bring it to the court.[667] In this regard, introduction of collective redress and representative action mechanisms to the EU antidiscrimination law may help overcoming the shortcomings of the reliance on individual redress especially against pressing algorithmic discrimination issues.[668]

2.2.2.4.2 Black Box Effect

Awareness of discrimination has two sides. On the one side those being subject to automated profiling and decision-making have little means to observe whether the job ad, content, price presented to them reflects a potential bias or not.

---

[664] D'Ignazio and Klein, *Data Feminism*, 58.

[665] See Xenidis and Senden, "Eu Non-Discrimination Law in the Era of Artificial Intelligence: Mapping the Challenges of Algorithmic Discrimination," 25. for the reasons individuals deter from seeking redress for their right to equality.

[666] Eubanks, "Automating Inequality : How High-Tech Tools Profile, Police, and Punish the Poor."

[667] O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, 105-12.

[668] Sara Benedi Lahuerta, "Enforcing Eu Equality Law through Collective Redress: Lagging Behind?," *Common Market Law Review* (2018); Federica Ceccaroni, "Collective Discrimination without an Identifiable Victim in Eu
Law. Discrimination by Public Speech," *Forum di Quaderni Costituzionali* (2020).

Moreover, many times both the technology used and also the organisational structures create black box effect and such lack of transparency further hinders awareness of discrimination. On the other side of the coin is to make others aware of the prima facie discrimination that is experienced so that it may be disputed before the court. However, black box effect in the form of complexity of the systems, trade secret and IP law protection and hardships in comparing the treatment received with the others, such as in the case of lacking the chance of comparing a job offer with the offers received by other job seekers, makes it an equally hard task.

2.2.2.4.2.1 Awareness of Discrimination

In order to have a case under Antidiscrimination Law, the complainant must establish evidence that indicates to discrimination based on a protected ground. If successful the burden of proof shifts to the defendant. However, before starting to collect evidence, one must first become aware of or feel being discriminated.

Normally at the workplace, a hardworking female employee keeps seeing male colleagues who are not as qualified as herself being promoted, it would not get too long for her to put the pieces of the puzzle together. On the other hand, a qualified female job seeker, uploads her CV on an online job seeking platform and receives advertisements of some positions, would just apply to positions she was advertised and might initially think it must be a bad time for seeking jobs or she is maybe not qualified enough for better positions. She would not be able to pin down that she is being treated unfavourably and even if she feels so, she might not come up with sex discrimination immediately as she also doesn't see what kind of job advertisements other people get. Moreover, maybe the hostility of the job market towards women is too commonplace and it might cross her mind. What if an algorithm discovered a thirty something urban millennial who owns a cat and listens to hip-hop music is more likely to change jobs? Or maybe another algorithm prefers people whose first name is Jarred and played high school

lacrosse?

Let's assume a non-profit organization devoted to algorithmic justice discovered that company or platform weeds out job seekers who have a cat and listens to hip-hop music, and they speculated this may affect disproportionately women of a particular ethnic origin based on some studies with publicly available statistical data. People start sharing their experience about this service provider online, she becomes convinced that despite not being from that ethnic origin, she was assumed to be based on her music genre preference. She contacts the organisation and accepts to become their representative case in strategic litigation pleading indirect discrimination based on sex and ethnicity (by association or ascription) regarding access to employment under Framework and Race Directives.[669] Such consciousness raising through sharing of personal experience is a feminist method[670] and we find it crucial for bringing algorithmic discrimination before the courts. Most of the time, non-expert victims will not become aware of being discriminated unless the output is too suspicious or causes some public scandal, and even if they do, bringing enough evidence to establish prima facie discrimination would not be possible.[671]

2.2.2.4.2.2 Access to Evidence

Now we will leave our hypothetical example above and explore the other side of the coin. Access to evidence within these complex technical and organisational structures is both practically hard and legally not so well if at all supported.

---

[669] Wulf, *Automated Decision-Making Systems and Discrimination: Understanding Causes, Recognizing Cases, Supporting Those Affected* 20.

[670] Bartlett, "Feminist Legal Methods," 863-67.

[671] See Wulf, *Automated Decision-Making Systems and Discrimination: Understanding Causes, Recognizing Cases, Supporting Those Affected* 16-17. For some tips on recognizing algorithmic systems.

Burden of Proof Directive, requires complainants to establish "facts from which it may be presumed that there has been direct or indirect discrimination."[672]

First of all, finding evidence in order to establish prima facie discrimination for a discriminating ADM process would be time consuming and require allocation of serious resources. For an ADM process that pretty much follows always the same logic, a testing procedure can be used. A test and a control subject who are similar in all their characteristics but for the suspect ground. They apply for the same house offer, credit or job posting; and the fact that they received different results can be used to establish prima facie discrimination[673] if the national court finds it admissible.[674]

This practice is reminiscent of situation testing.[675] In *Kratzer*, the CJEU ruled making an application to a job post solely to gain the formal status of applicant rather than really obtaining that post, cannot be considered as 'access to employment, to self-employment or to occupation' under Recast and Framework Directives and could even be considered abuse of rights.[676] However, in this case the sole purpose of the applicant was determined to be seeking compensation.

Another way of collecting evidence can be acting as a user or a publisher to

---

[672] Directive 97/80/EC, 15 December 1997, Article 4(1).

[673] Wulf, *Automated Decision-Making Systems and Discrimination: Understanding Causes, Recognizing Cases, Supporting Those Affected* 19.

[674] C-553/11 [2012] *Rintisch v Klaus Eder* ECLI:EU:C:2012:671 para 15. Finding and assessing the facts in the case in the main proceedings are under the jurisdiction of the national court alone.

[675] *Handbook on European Non-Discrimination Law*, 241.

[676] C-423/15 [2016] *Nils-Johannes Kratzer v R+V Allgemeine Versicherung AG* ECLI:EU:C:2016:604 para 44.

invoke a victimless discrimination case.[677] AlgorithmWatch published two job opportunities on Facebook in 2020, one for a truck driver and one for an educator, by purposely leaving out the targeting option, but the truck driver ad was delivered disproportionately to men, while educator ad was delivered disproportionately to women.[678] Xenidis, suggested that CJEU's approach in *Feryn*, *Rete Lenford* and *Accept*, that it did not require an identifiable complainant to claim to have been the victim of discrimination,[679] may allow EU Antidiscrimination Law to capture dissemination of this kind of harmful stereotyping, as well as collective harms arising from being systemically excluded from opportunities in terms of access to employment, credit, housing and so on, due to biased ADM processes.[680]

Once the prima facie discrimination is established, that from the facts presented the unfavourable treatment can be presumed, the burden of proof shifts to the defendant. This redeems for the claimants lack of access to the whole process at place to prove precisely the facts that are causally linked to the discriminatory result. Especially in algorithmic discrimination, no one is better placed than the developers and deployers of the model, and even them sometimes can have a hard time discovering or illustrating what causes the discriminatory result such as in the case of deep learning algorithms as some models could be preferred to be

---

[677] *Handbook on European Non-Discrimination Law*, 241. C-54/07 [2008] *Feryn* ECLI:EU:C:2008:397; C-81/12 [2013] *Asociaţia Accept* ECLI:EU:C:2013:275; C-507/18 [2020] *NH v Associazione Avvocatura per i diritti LGBTI — Rete Lenford* ECLI:EU:C:2020:289.

[678] Wulf, *Automated Decision-Making Systems and Discrimination: Understanding Causes, Recognizing Cases, Supporting Those Affected* 9. Similarly Sweeney, "Discrimination in Online Ad Delivery."; Datta, Tschantz, and Datta, "Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination."; Angwin and Jr., "Facebook Lets Advertisers Exclude Users by Race"; Adam Gabbatt, "Facebook Charged with Housing Discrimination in Targeted Ads," https://www.theguardian.com/technology/2019/mar/28/facebook-ads-housing-discrimination-charges-us-government-hud.

[679] *Handbook on European Non-Discrimination Law*, 114.

[680] Xenidis, "Tuning Eu Equality Law to Algorithmic Discrimination: Three Pathways to Resilience," 750.

uninterpretable as it increases the predictive power,[681] also complexity of the system may make it inscrutable and nonintuitive.[682] ADM processes usually encompass an assemblage of complex company procedures, datasets and models.[683] In *Danfoss*, a similar problem arose from the way system of individual supplements applied to basic pay was implemented, that a woman employee could not comprehend why she was receiving a different pay check than a male colleague performing the same work.[684] The CJEU ruled that "where an undertaking applies a system of pay which is totally lacking in transparency, it is for the employer to prove that his practice in the matter of wages is not discriminatory, if a female worker establishes, in relation to a relatively large number of employees, that the average pay for women is less than that for men."[685]

Danfoss brings two very important conclusions that (i) users of ADM systems cannot hide behind the corporate or technological complexity, they have to find a way of explaining their system in a way people who are affected by it can evaluate whether they are being treated unfavourably compared to others, (ii) this turn in a case can only be achieved if the output of the algorithm and group affiliations of the other persons (other applicants for a screening algorithm or people who received the job ad) are known.[686]

---

[681] Asghari et al., "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making.". 16; Jenna Burrell, "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms," *Big Data & Society* 3, no. 1 (2016).

[682] Selbst and Barocas, "The Intuitive Appeal of Explainable Machines."

[683] Ananny, "Toward an Ethics of Algorithms:Convening, Observation, Probability, and Timeliness."

[684] Case 109/88 [1989] *Handels- og Kontorfunktionærernes Forbund I Danmark v Dansk Arbejdsgiverforening, acting on behalf of Danfoss* ECLI:EU:C:1989:383 Para 10.

[685]Ibid Para 16.

[686] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 23.

Meister deals with one of the main questions that arise in establishing prima facie discrimination in claims regarding access to employment. Even if an ADM systems is not in place, it is difficult for an applicant who did not pass the screening stage to know the qualifications and group affiliations of other applicants to compare herself with. Thus, as confident in her qualification for the job she applied, Ms Meister asked for the production of the file for the person who was engaged, so that she could prove that she was more qualified.[687] As she claimed discrimination based on sex, ethnicity and age, the Court interpreted Recast, Race and Framework Directives together and found that none of them entitles a rejected applicant who meets the requirements in the job advertisement to have access to information on employer's engagement with another applicant.[688] Yet, the CJEU continued to decide that the fact that employer refuses to grant any information may be taken into account by the national court in establishing prima facie discrimination.[689] This raises the question whether an employer who uses an ADM process for screening and refuses to grant access to their model, is presumed to have discriminated, so that the burden of proof shifts. According to Hacker, nondisclosure of a company may have many legitimate reasons for a national court to assess such as IP rights, trade secrets, costs of algorithmic auditing.[690]

Thus, as beneficial as it is that a claimant only have to establish a presumption towards the likelihood of direct or indirect discrimination and is not burdened with proving that a complex technology disadvantaged them, three problems remain: (i) it is still difficult for algorithmic discrimination to get this far in legal proceedings if at all, (ii) antidiscrimination law does not seem to provide a key to unlock black boxes of ADM systems, (iii) even if the prima facie discrimination

---

[687] C-415/10 [2012] *Meister v Speech Design Carrier Systems GmbH* ECLI:EU:C:2012:217 para 29.

[688] Ibid para 46

[689] Ibid para 47.

[690] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."; ibid.

is established, the plaintiff may still disprove the claim that the complainant and the comparator are not in a similar situation or that there is an objective reason behind the differential treatment, and if disproving fails, defendant still have an option to justify the difference in treatment.

2.2.2.4.3 Justification

As we already mentioned above, there are limited exceptions in Equality Directives to direct discrimination and justification is possible with respect to indirect discrimination formulated within the very definition of indirect discrimination. Thus indirect discrimination does not occur if the contested provision, criterion or practice that is neutral on its face, can be justified with a (I) legitimate aim and (II) overall proportionate to the aim pursued; meaning it is (i) necessary in the sense that the same result could not be achieved through a similarly effective but less discriminatory method, and (ii) is appropriate for reaching this aim.[691]

Most of the case law of the CJEU is regarding Member State legislation regarding the legitimate aim. As regards actions by employers may be objectively justified if they correspond to a real need on the part of the undertaking and differential treatment is based on factors unrelated to any discrimination on a protected ground.[692]

Aims pursued by employers who make use of automation and predictive technologies, may vary from increasing objectivity, inclusion and diversity, reducing financial and human resources allocated to hiring processes, making

---

[691] Ibid.

[692] Case 170/84 [1986] *Bilka-Kaufhaus GmbH v Karin Weber von Hartz* ECLI:EU:C:1986:204 para 39; C-196/02 [2005] *Vasiliki Nikoloudi v Organismos Tilepikinonion Ellados AE* ECLI:EU:C:2005:141 para 47.

better hiring decisions, composing better fitting teams, preventing theft, increasing employee tenure,[693] reaching to a wider range of candidates and so on. Hacker argues, as long as an algorithm is not built to mask discriminatory intent, all of these aims would be accepted to be legitimate.[694] However, according to the established case law of the CJEU "objectives, which are related essentially to personnel management and budget considerations cannot be considered objective reasons."[695] The CJEU explicitly mentioned discrimination cannot be justified by purposes such as rigorous personnel management[696] and saving money.[697] In this regard we find it considerable that automating human resources in a way it causes indirect discrimination might not be justified as a legitimate aim so easily as it is many times about better personnel management and reducing costs. Moreover, while the CJEU accepted that Member States enjoy a broad margin of discretion on social policy, the same is not true for the social or employment policies of employers on which they cannot rely as a legitimate aim.[698]

Proportionality is defined by being suitable to the aim pursued and necessary to achieve this aim. According to Hacker, suitability should be understood as

---

[693] Bogen and Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias."

[694] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

[695] C 644/19 [2020] *FT v Universitatea 'Lucian Blaga' Sibiu, GS and Others, HS, Ministerul Educaţiei Naţionale* ECLI:EU:C:2020:810 para 52

[696] C-486/08 [2010] *Zentralbetriebsrat der Landeskrankenhäuser Tirols v Land Tirol* ECLI:EU:C:2010:215 para 46; C-443/16 [2017] *Francisco Rodrigo Sanz v Universidad Politécnica de Madrid* ECLI:EU:C:2017:109 para 52.

[697] C-16/19 [2021] *VL v Szpital Kliniczny im. dra J. Babińskiego Samodzielny Publiczny Zakład Opieki Zdrowotnej w Krakowie* ECLI:EU:C:2021:64 para 59

[698] C-281/97 [1999] *Andrea Krüger v Kreiskrankenhaus Ebersberg* ECLI:EU:C:1999:396 para. 28-29.Christa Tobler, *Limits and Potential of the Concept of Indirect Discrimination* (Luxembourg: : Office for Official Publications of the European Communities, 2008), 32.

effectiveness.[699] The CJEU requires evidence to be brought rather than mere generalisations to show that the means chosen were suitable for achieving the aim.[700] Thus Hacker argues that the fact that machine learning models are based on statistics and predictive accuracy is measured, it would be no problem to show evidence as regards the suitability to the pursued legitimate aim for an ADM system. This is more true in cases of indirect proxy as the model will produce indirect discrimination but if the aim was legitimate and the predictive accuracy was high then it would be justifiable. In cases of biased training and test data, because the model will not perform high predictive accuracy in the real world, justification might be harder. However in any case, for the claimants it would be difficult to refute predictive accuracy of these models, as a result of lack of accessible proper statistics[701] on minority groups, and especially in the case of intersectional discrimination because of lack of disaggregated data.[702]

In assessing necessity, it must be shown that another reasonable alternative measure that produces less or no disparate effect would not be as effective.[703] Thus first of all, it would be argued that whether human decision-making instead of ADM could have produced same effectiveness with less disparate effect. Hacker, is in the opinion that if the predictive power of the ADM is high than its

---

[699] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

[700] C-167/97 [1999] *Regina v Secretary of State for Employment, ex parte Nicole Seymour-Smith and Laura Perez* ECLI:EU:C:1999:60 para 59.

[701] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

[702] Commission et al., *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*.

[703] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 18; *Handbook on European Non-Discrimination Law*, 94; Tobler, *Limits and Potential of the Concept of Indirect Discrimination*, 35.

effectiveness would always trump decision making solely by humans.[704] However, we are in the opinion that effectiveness would still be determined on the basis of the aim pursued and depending on the aim, especially if it is not about producing as many decisions as fast as possible, human decision-making may become preferable for the courts. Other comparison for necessity criterion could be whether there was a less biased training data or another model which could minimize bias while keeping the same level of accuracy. Here, Hacker also notes that, in evaluating whether adoption of such alternative system would be reasonable, as we discussed above, it is more likely the Court will not accept less biased alternative being more costly. An overall proportionality test would take into consideration what is at stake in terms of fundamental rights and freedoms. In cases where Netflix is suggesting a new series to be watched disproportionately to a protected group, it might be found unreasonable if the alternative costs much more than the data and model used. However, in a case of predictive policing it might be required that the user should have opted for the more costly option.[705]

As to our knowledge, there has not been a case brought before the CJEU on indirect discrimination caused by an ADM process yet, there is no clarity on what criteria would be used to assess objective justification. However, the fact that functioning of machine learning models and the limitation of the EU antidiscrimination law on a closed list of protected grounds, makes indirect discrimination more applicable to algorithmic discrimination, risks algorithmic discrimination to get away with an objective justification.

---

[704] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

[705] Ibid.

CHAPTER 3

3 EUROPEAN DATA PROTECTION LAW AND FORTHCOMING LEGISLATION ON ADM

3.1 EUROPEAN DATA PROTECTION LAW

In this part of our study, we will look into the European Data Protection Law and its rules that are regulating algorithms and applicable to algorithmic discrimination, while attempting to understand how GDPR's protection regime applies to algorithms, we will also juxtapose it with the limitations of European anti-discrimination law in order to see how it interplays, enhances and propels the protection offered by it against discriminatory harms.

Here, by the European Data Protection Law, we mainly mean Charter of Fundamental Rights of the European Union (CFREU) Article 8, the Regulation (EU) 2016/679 (GDPR), its predecessor Directive 95/46/EC, Directive 2002/58/EC (E-privacy Directive), Proposal for a Regulation on Privacy and Electronic Communications (E-privacy Regulation) as they are the most related to the topic of this study, as well as the guidance of Article 29 Working Party and its successor European Data Protection Board (EDPB), as well as opinions of the European Data Protection Supervisor (EDPS). However, we acknowledge that the data protection law instruments of the European Union are not limited to those and will refer to them when it fits our purposes.

3.1.1 Personal Data and Data Subjects

Here, we will first start with exploring the boundaries of the data protection law which is drawn around personal data related to an identifiable natural person.

3.1.1.1 Personal data

The material scope of the data protection law is limited to personal data. If the data at hand falls out of the legal definition and interpretation of personal data, the protection and toolset that data protection law offers will not be applicable. This binary nature[706] of data protection law causes difficulties in determining when the rules of personal data protection apply when it is concerned with more complex data processing operations such as creating profiles which result from processing of large datasets by use of machine learning techniques.

In Art. 4(1) of the GDPR personal data is defined as *"any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person."*

Similar to the OECD Guidelines and 95/46/EC Directive, GDPR also defines personal data with four constituent elements: (i) any information, (ii) relating to, (iii) identified or identifiable, (iv) natural person.[707]

How broad the scope of information should be understood from the wording *"any information"* might be problematic as discussed by Purtova (2018), especially as a result of broad understanding of the CJEU and Article 29 Working

[706] Bart Schermer, "Risks of Profiling and the Limits of Data Protection Law," in Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases, ed. Bart Custers, et al. (Springer Publishing Company, Incorporated, 2013).

[707] Christopher Kuner, Lee A. Bygrave, and Christopher Docksey, eds., The Eu General Data Protection Regulation (Gdpr): A Commentary (New York: Oxford University Press, 2020).

Party, former EU advisory authority on the matters of data protection (WP29), and lack of providing a definition for what constitutes information.

WP29 distinguishes between data types as data provided data (which are provided by the data subjects themselves), observed data (which are observed by the controller), derived data and inferred data.[708] According to WP29, a profile of the individual that has already been created, constitutes inferred data about that individual. Examples to inferred data include credit scores, algorithmic results by a service provider such as results of recommendation process. Thus, whether and to what extent GDPR applies within the chosen contexts of this study depends also on the types of data processed. Automated decisions, in which we are interested, are based on profiles or scores reflecting an evaluation process by the data controller. These processes are not always based on objective and verified information but often on subjective, non-verifiable information, opinions or assessments which can also be called inferences.[709] For instance, in the employment context a good employee, as we have discussed previously, does include this kind of assessments. Here then we talk about derived or inferred data. According to the WP29, derived and inferred data in the sense of inferences are personal data. Also, both objective information and subjective information such as opinions or assessments are included within the concept of personal data. WP29 refers to personal data processing in sectors such as banking, for the assessment of the reliability of borrowers ("X is a reliable borrower"), in insurance ("Y is not expected to die soon") or in employment ("Z is a good worker and merits promotion") as examples of subjective information.

---

[708] Article 29 Working Party, "Guidelines on the Right to Data Portability," WP 242 (2017); "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679"; Sandra Wachter and Brent Mittelstadt, "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and Ai," *Columbia Business Law Review* 2019, no. 2 (2019).

[709] "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and Ai."

In the narrow sense, information is defined as "data + meaning or significance."[710] However, Purtova argues that this understanding of information may lead to one of the reasons that data protection law may become the law of everything, as machine learning models are capable of attaching meaning or significance to any data.[711] On the other hand, Wachter and Mittelstadt argued that the ECJ is not in the same opinion with the WP29 and inclined to a more restricted interpretation of the scope of personal data, as well as the applicable rights through their analyses on *YS and M. and S*.[712] In this case, the CJEU denied access for the data subjects to the legal analyses of the immigration office on the ground that opinions, reasoning and assessments that underlie decisions are not considered personal data.

Actually, the CJEU engaged with the "any information" constituent of personal data for the first time in *Nowak*[713] and stayed in line with WP136 by stating that subjective kinds of information in the forms of opinions and assessments are also included in "any information."

However, Wachter and Mittelstadt come to the conclusion that even if such "subjective information" is considered personal data, CJEU excluded the decisions both in *YS and M. and S* and in *Nowak* cases.[714]

---

[710] Purtova, "The Law of Everything. Broad Concept of Personal Data and Future of Eu Data Protection Law," 53.

[711] Ibid.

[712] Wachter and Mittelstadt, "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and Ai." See. C-141/12 and C-372/12 YS Joined Cases [2014] *YS (C-141/12) v Minister voor Immigratie, Integratie en Asiel, and Minister voor Immigratie, Integratie en Asiel (C-372/12) v M, S.* ECLI:EU:C:2014:2081

[713] C-434/16 [2017] *Peter Nowak v Data Protection Commissioner* ECLI:EU:C:2017:994

[714] Wachter and Mittelstadt, "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and Ai."

Content of information is also not limited to private or family life.

(ii) For the second element of personal data, *"relating to"* again a broader understanding is adopted by WP29 in WP136 and the CJEU in *Nowak*.[715] According to the WP29, three not cumulative but alternative conditions, namely "content," "purpose" and "result," should be taken into consideration in determining whether data is relating to an individual. From the point of content element, "information "relates" to a person when it is "about" that person."[716] Data relate to an individual in purpose when they are used or are likely to be used with the purpose to evaluate, treat in a certain way or influence the status or behaviour of an individual.[717] Finally for the result element to be present, there must be a possible impact on a certain person's rights and interests. Although a major impact is not required, the possibility that the individual may be treated differently from other persons as a result of the processing of such data is sufficient.[718]

According to Purtova, the reference by WP29 to likelihood for both purpose and result elements, makes the status of personal data dynamic by being broad and context-dependent.[719] This leads to a concern about any data to be presumably relating to an individual as it is not really known[720] how the self-learning machines draw meaning from vast amounts of data collected from various interconnected environments in real time which increasingly includes offline life such as smart

---

[715] C-434/16 [2017] *Peter Nowak v Data Protection Commissioner* ECLI:EU:C:2017:994

[716] Article 29 Working Party, "Opinion 4/2007 on the Concept of Personal Data " WP 136 (2007).

[717] WP136ibid.

[718] WP136ibid.

[719] Purtova, "The Law of Everything. Broad Concept of Personal Data and Future of Eu Data Protection Law," 54. Purtova interpretes the lack of word "reasonably" with respect to the likelihood of "related to" element different than the language used for the "identifiablity" element as a lower threshold of likelihood.

[720] Asghari et al., "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making.".

cities, smart homes and IoT devices as well.[721] Another reason for all information to be likely to have an impact on individuals is the way all available data is being processed without purpose limitation in order to find out patterns and insights which would not be possible without making use of Big Data analytics powered with machine learning algorithms.[722]

(iii) WP29 understands an identified natural person as "distinguished from all the members of the group" and identifiable when identification is possible even though she is not identified yet.[723] According to WP29, while being directly identified is mostly related to the name to be known, being indirectly identified requires the possibility to be singled out in a crowd depending on the circumstances of the cases. As exemplified by the WP29, while a very common surname may not single out someone within the population, wearing a black suit may be enough to single someone out among passers-by standing at a traffic light. Similar to the Recital 26 of the DPD, in its Recital 26, GDPR sets the standard for the possibility of identification as follows:

"to determine whether a natural person is identifiable, account should be taken of *all the means reasonably likely to be used*, such as *singling out, either by the controller or by another person to identify the natural person directly or indirectly*."

In line with the opinion of WP29 in WP136, Recital 26 of the GDPR "To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and

---

[721] Purtova, "The Law of Everything. Broad Concept of Personal Data and Future of Eu Data Protection Law." Purtova provocates that even weather and waste water can be interpreted as personal data if the broad approach of WP29 is taken, which would lead GDPR to try regulating everything and the protection it offers to become obscure.

[722] Ibid.

[723] Article 29 Working Party, "Opinion 4/2007 on the Concept of Personal Data ".

the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments."

In the *Breyer* case the CJEU adopted a broad approach to identifiability by stating that all the information that enables identification doesn't have to be in the hands of one person.[724] On the other hand, the Court deviated from WP 136 in its assessment that identification is not considered reasonably likely to be used if it is legally banned.[725]

Another question in assessing applicability of the GDPR to ADM or profiles which result in causing discriminatory harms is whether the use of anonymized data in making the decision or producing the profile would make them fall out of the scope of GDPR. It is clarified in the Recital 26[726] that the GDPR does not concern the processing of anonymous information, "namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable." According to the opinion of WP29, data is still identifiable unless anonymisation is irreversible.[727] In this regard, Hacker argues that due to the current state-of-the-art deidentification techniques, most likely training data used in most supervised and reinforcement learning models will remain outside the scope and protection of the GDPR.[728] When the model produces decisions that

---

[724] C-582/14 [2016] *Patrick Breyer V. Bundesrepublic Deutschland* ECLI:EU:C:2016:779

[725] Purtova, "The Law of Everything. Broad Concept of Personal Data and Future of Eu Data Protection Law."

[726] Recital 26 reads as follows: "Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person."

[727] Article 29 Working Party, "Opinion 05/2014 on Anoymisation Techniques," WP 216 (2014).

[728] Philipp Hacker, "A Legal Framework for Ai Training Data—from First Principles to the Artificial Intelligence Act," *Law, Innovation and Technology* 13, no. 2 (2021).

apply to identifiable persons then GDPR comes to play. However, it is problematic when a model is developed on irreversibly anonymized data, this may also mean the data controller cannot go back and check the model to find out from where the discriminatory result emerged.

3.1.1.2 Data Subject

No matter (un)identified/(un)identifiable, data protection law is designed to be applied to a natural person's data called the data subject. Then, it is also important to ask who this subject is and how their subjecthood is constituted. If the data subject is a neutral term that involves every identified or identifiable natural person and provides the same protection to all of them, then is it possible that this neutrality veils the needs and experiences of actual data subjects who are gendered, racialized, and classed? Put differently, if the main aim of data protection law is to protect the individuals from the power imbalances that may arise from data processing, does data protection law take this person's vulnerability to such affect into consideration? If the gender, race, class and other markings of the data subject is not taken into consideration specifically, then how could the power imbalance based on them addressed?

In this regard, Malgieri and Niklas identified two dichotomies regarding the vulnerability of the data subject. First dichotomy is about particularistic and universalistic approaches. In particularistic approach, particular individuals and groups are understood to be vulnerable based on their specific situations or socio-economic context such as women, LGBTI+, racial minorities, asylum seekers, people with disabilities as it is seen in diversity and equality policies as well as in antidiscrimination laws.[729] According to the universalistic approach data protection law protects all individuals equally against the power asymmetries which stem from the use of data processing technologies. The second dichotomy

[729] Malgieri and Niklas, "Vulnerable Data Subjects," 3.

is related to vulnerability risks which may rise either from the data processing itself in the form of decisional vulnerability -such as in the cases of age and mental illness, as well as socio-economic status (literacy, poverty, minority status) which may inhibit the comprehension of information about the purposes and consequences of the data processing, provision of free consent and exercise of data subject rights- or the outcomes of such processing -such as discriminatory ADM. It is then suggested that Luna's layered vulnerability[730] would reconcile the criticisms of vulnerability approaches, assist risk-based approach and facilitate incorporating intersectionality.[731] First criticism of vulnerability approach is that it causes stigmatization on particular groups. Thus it is suggested that all humans are vulnerable as a universal human condition[732] which is relational and contextual as an individual maybe vulnerable in one situation but not in another, vulnerability may change in time as a child would grow out of vulnerability while another individual becomes vulnerable due to aging or loss of wealth, vulnerability may change in relation to the space as different countries may have different laws or culture, public and private spheres, urban and rural areas may also differ from their dangers to applicable law thereto. Nevertheless, there are two important reasons to oppose the universality of vulnerability. First, it may lead to arguments to the extent if everyone is vulnerable as a natural state of being human there is no reason to provide protection from it. Second, such an argument would disregard historical and structural subordination and oppression of some populations and complicity of some populations and institutions therein. On this account, Luna suggests thinking about vulnerabilities as layers that individuals can put on or take off according to time, space, situation, and context.[733] Moreover, she calls out fixed

[730] Luna, "Elucidating the Concept of Vulnerability: Layers Not Labels."; Luna, "Identifying and Evaluating Layers of Vulnerability - a Way Forward."

[731] Gianclaudio Malgieri and Gloria González Fuster, "The Vulnerable Data Subject: A Gendered Data Subject?," SSRN Electronic Journal (2021); Malgieri and Niklas, "Vulnerable Data Subjects."

[732] Martha Fineman, "The Vulnerable Subject: Anchoring Equality in the Human Condition," Yale J. of Law & Feminism 20 (2008).

[733] Luna, "Elucidating the Concept of Vulnerability: Layers Not Labels."

labels to be too simplistic of an approach to complicated problems. Layers may overlap and intensify vulnerabilities, but they may also be addressed and minimized layer by layer. Most importantly they do not label a subpopulation[734] as vulnerable, however each individual in it may have several layers of vulnerabilities which allows for addressing different situations of individuals within the groups that might be increasing their vulnerability. Malgieri and González Fuster find this solution compatible with the layered risk based approach in the GDPR and suggest that it may also allow accounting for intersectional discrimination.[735]

Another important point to take into account while accepting the fact that some individuals and groups might be more or less vulnerable, it also brings up the notion of a non-vulnerable or at least a standard data subject -which does not mark a black and white separation between those who are vulnerable and those who are not, rather represents the other pole of a wider spectrum of different levels of vulnerability in between.[736] According to Malgieri and González Fuster, GDPR does not determine the main characteristics of the (standard) data subject.[737] One characteristic that can be derived from the text of the GDPR is that the data subject suffers from information asymmetry. However, if they are properly informed then they are capable of overcoming it by weighing the risks and advantages of the intended processing, and upon being informed about their existence they can use their rights as data subjects. Thus we can say that the standard data subject is imagined to be a rational individual, as the reasonably informed, observant and circumspect consumer of the consumer law. Such definitions of standard or

---

[734] A term used in research but also in data science in order to refer to a fraction or part of the overall pool of the population from which the data were collected. A subpopulation or subgroup can be defined by gender, race/ethnicity, age (e.g. <40, 40-64, 65+), location and so on.

735 Malgieri and González Fuster, "The Vulnerable Data Subject: A Gendered Data Subject?." ibid., 6.

[736] Ibid.

[737] Ibid.

average individual is criticised as it may reinforce attributes of dominant groups.[738] In this study, we have challenged that the good employee is not a neutral definition and showed that they are being packed with gender, race, class and other attributes which may cause discrimination when they deviate from what is dominant. Nevertheless, despite the fact that Spanish and Polish DPAs have considered women to be a vulnerable category, GDPR and the data protection law remains to be gender-blind to the extent it is not listed among the special categories of data in Art. 9 GDPR which are subject to stricter protection as processing of such data may cause discrimination.[739] This is at great odds with how combatting sex discrimination has been the locomotive of the EU antidiscrimination law. We will discuss later whether this can be an advantage or disadvantage in addressing gender discrimination resulting from ADM. As we mentioned in the introduction by drawing on Crenshaw,[740] data protection should be concerned with finding out who the most vulnerable data subjects are, because if it manages to protect them, it manages to protect us all.

This is not an easy task in the face of unique challenges that algorithmic systems are bringing in sorting and classifying people as Buolamwini and Gebru showed in their gender shades study.[741] These identified or identifiable people, data related to whom are being processed to make decisions about them, are not represented well by the individualist and neutral notion of data subject as the real data subject is gendered, racialised, and classed. Moreover, it is collective in the sense that it represents the collective trauma and historical injustice that marks the group identity and memory, as well as the will to collectively heal and to be emancipated from it. According to Kadiri, only an intersectional data subject that

---

[738] Ibid.

[739] Ibid., 10-11.

[740] Crenshaw, "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics."

[741] Buolamwini and Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification."

accounts for interconnectedness, unsettle predefined categories, and acknowledges the structural aspect of discrimination, would be able to address these challenges[742] which have power struggles in the centre and are the digital continuation of historical objectification of women and Black people.[743]

3.1.1.3 Special categories of personal data

As Modernised Convention n. 108 in Art. 6(2),[744] Recital 51 of the GDPR also refers to the significant risk that processing of special categories of personal data pose on fundamental rights and freedoms. It is knowingly a lesson learned from the history of Europe that information on sensitive attributes such as racial or ethnic origin, health status, political opinions, religious or philosophical beliefs may be used in such a way that causes the greatest horrors. On the other hand, this approach of singling out a sub-category of personal data was criticized for being contradictory to the general understanding of sensitivity in data protection law as context-dependent.[745]

Even if the higher protection provided to these categories of personal data reasons from them to enable discrimination, special categories of data in Article 9 of GDPR does not overlap with the protected attributes under antidiscrimination law. Wachter & Mittelstadt (2019) note that, gender and age, despite causing significant discrimination, are not considered special category data. Same goes for

---

[742] Kadiri, "Data and Afrofuturism: An Emancipated Subject?." ibid., 9.

[743] Ibid., 16.

[744] Specifically refers to the risk of discrimination.

[745] Kuner, Bygrave, and Docksey, The Eu General Data Protection Regulation (Gdpr): A Commentary. Also see. Maximilian von Grafenstein, *The Principle of Purpose Limitation in Data Protection Laws: The Risk-Based Approach, Principles, and Private Standards as Elements for Regulating Innovation*, 1 ed., vol. 12, Schriften Zur Rechtswissenschaftlichen Innovationsforschung (Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG, 2018). for a criticism of contextuality by Nissenbaum and proposal of purpose specification being enough for it.

financial situation and personal profiles unless they contain data categories listed in Art. 9 GDPR or Art. 9 data related to an individual can be concluded from them. According to WP29, "profiling can create special category data by inference from data which is not special category data in its own right but becomes so when combined with other data".[746] On this account, proxy data may end up being treated as special category data.[747]

It has been suggested that even though processing of special categories of data is regulated with stricter rules such as requiring explicit consent of the data subject and is prohibited unless there is a legitimate ground as specified in Article 9 and 10, preventing discrimination in ADM systems might require processing of that very data.[748] Veale and Binns suggested three approaches which may be useful to ensure fairness without processing special category or sensitive data either because it is legally restricted to use such data or because it is not desirable due to information security or ethical concerns, however they are usually required for fairness metrics and other methods with respect to discovery and correction of algorithmic discrimination. Their alternative suggestion involves trusted third parties holding protected characteristics, knowledge bases about fairness in data and models, and exploratory fairness analysis. On the other hand Ivanova argued, there was no legal restriction keeping special categories of data to be processed for bias mitigation, both in the initial phase of training and testing, as well as in application stage. In the first case, as to the broad definition given to scientific research in Recital 159, it would be possible to rely on Article 89(1) provided that

---

[746] Article 29 Working Party, "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679". 15.

[747] Wachter and Mittelstadt, "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and Ai."

[748] Indrė Žliobaitė and Bart Custers, "Using Sensitive Personal Data May Be Necessary for Avoiding Discrimination in Data-Driven Decision Models," Artificial Intelligence and Law 24, no. 2 (2016); Michael Veale and Reuben Binns, "Fairer Machine Learning in the Real World: Mitigating Discrimination without Collecting Sensitive Data," Big Data and Society 4, no. 2 (2017).

the data was collected under a compatible purpose. In the latter case, Ivanova argues that decision of the CJEU in *GC and others*[749] would allow processing special category data with the sole purpose of bias discovery and correction by relying on substantial public interest (Article 9(2)(g)) in order to safeguard right to non-discrimination, provided that Article 22(4) does not apply and sufficient safeguards are ensured.[750]

3.1.1.4 Group Privacy[751]

"Sometimes the only way to protect the individual is to protect the group to which the individual belongs. Preferably before any disaster happens."[752] In anti-discrimination law, the subject of protection is not the group, but the individuals who share characteristic with a protected ground. Special categories of data are protected in relation to individual data subjects about whom they may reveal sensitive information that may cause discriminatory harms. For instance, in right to self-determination, the right holder is the whole nation instead of individual citizens.[753]

However, citizens of a country is a well-defined group. Group privacy or collective data protection is usually discussed in relation to privacy concerns regarding *ad hoc* groups that are the output of algorithmic profiling. These are specifically groups which are not already defined, as they do not even exist before

---

[749] C-136/17 [2019] *GC,AF, BH, ED v Commission nationale de l'informatique et des libertés* (CNIL) ECLI:EU:C:2019:773 para 68.

[750] Yordanka Ivanova, "The Data Protection Impact Assessment as a Tool to Enforce Non-Discriminatory Ai," *SSRN Electronic Journal* (2020): 5.

[751] See. Linnet Taylor, Luciano Floridi, and Bart van der Sloot, eds., *Group Privacy: New Challenges of Data Technologies* (Springer, 2017).

[752] Luciano Floridi, "Open Data, Data Protection, and Group Privacy," *Philosophy & Technology* 27, no. 1 (2014): 3.

[753] Ibid.

being classified to form a group with allegedly shared characteristics. They are formed by the discovery of a shared treats inferred from their online behaviour and its correlation to known treats or classes, which are not necessarily overlapping with a protected group. In time, these groups and individuals that are classified to belong to these groups may change based on new information or change of behaviour or any other factor. Thus it is not possible to identify them before they are formed, their group definitions and members are not static, they do not know they are a part of a group. Yet, they are subject to algorithmic decisions about themselves, they serve as the ground for decisions about other individuals or groups, and these decisions may violate their rights to privacy and equality, and all the other rights that are initially protected by these rights in a democratic society. Moreover, they remain outside the scope of individual protection provided by antidiscrimination and data protection laws, as they may not relate to any identified or identifiable person who is affiliated with a protected group (in a protected domain if we will reduce antidiscrimination protection to EU's secondary law).

It was discussed in the previous chapter that not offering any protection to such *ad hoc* groups, even though they might not be the priority today, may both be unfair still and may lead to different kind of structural problems for the future. On the other hand, how can law draft a protection and enforce it for groups, members of which are not aware of their affiliation with this group, do not know other members of the group, thus who cannot deliberate on their shared interests and take action upon them?

According to Mittelstadt, there are examples that in ethics and in law that despite such characteristics the need for group level protection has been acknowledged and put into action. First of all, in philosophy, it is acknowledged that once someone or something lacks or loses agency, this shall not mean they cannot be morally considered. These are called moral patients such as babies and toddlers, persons who suffer from loss of mental capabilities or conscious, and

they are indeed the objects of moral responsibility of moral agents. Thus, being aware of group membership and interests is not necessary to be acknowledged as a group worthy of protection. In US law, class actions are similarly are being brought on behalf of individuals who may not share any interest before the reason for the action occurred and who may at the time of the class action still not be aware of being part of a group, but their interests would be defended and at the end they may benefit from the compensation. Another example given by Mittelstadt is disparate impact doctrine in the US law but as it is very similar to the notion of indirect discrimination under the EU law we will just refer to it instead. Here, the difference to *ad hoc* groups is obviously, protection is granted to historically and structurally disadvantaged groups on the basis of a treat they have which is not morally acceptable to be the ground of unfavourable treatment. These groups are well defined, there are statistics available so that the courts can decide whether a group of persons that is disproportionately affected by a facially neutral rule, criterion or practice is disproportionately correlated with one of these protected groups. The similarity here to the *ad hoc* groups is that the disadvantaged group, say employees over the age of 50 in company x, did not exist as a group before they were subject to the contested rule, criterion or practice.

According to Mittelstadt, "group privacy based on the right to inviolate personality specifically aims to protect against (1) unwarranted third party manipulation of identity and (2) harms from automated decision-making based upon profiling identities assembled by third parties."[754] Currently such privacy protection is not offered to *ad hoc* groups in GDPR and they also remain outside the scope of antidiscrimination protection. One important step in the direction of such protection could be holding third parties who assemble and make decisions on ad hoc groups responsible for unjustified classifications and decisions, similar to objective justification in indirect discrimination doctrine. Such suggestion is elaborated in a later article by Wachter and Mittelstadt in the form of an individual

[754] Brent Mittelstadt, "From Individual to Group Privacy in Big Data Analytics," ibid.30 (2017): 15.

right to reasonable inferences, in which high risk inferences about individuals would require *ex ante* justification by the controllers and the right would grant the individuals an *ex post* right to challenge these inferences.[755]

3.1.2 General Principles in the ADM context

General principles of data protection law were first established in the 1980 OECD Guidelines and Convention n. 108 of 1981 and then incorporated also in 95/46/EC. Since then within the practice of data protection, they have been applied in various scenarios and proven to be sound. Therefore, with some adjustments and additions they appeared almost identically in the GDPR Art. 5 as well, this time covering a wider scope.[756]

3.1.2.1 Lawfulness, Fairness and Transparency

The core conditions for a lawful processing are listed under Art. 6 GDPR. In order to be lawfully process personal data, at least one of the conditions in Art. 6.1. should apply in addition to respecting all applicable legal requirements. Also conditions for lawful limitations in Art. 52.1 CFREU for Articles 7 and 8 of the CFREU must be considered when assessing the lawfulness of personal data processing under EU law. Thus, lawfulness also entails being in accordance with the law, pursuing a legitimate purpose and being proportionate in a democratic society to achieve this purpose.

Fairness of the processing is commonly understood as data to be obtained and processed through fair means which means not by deception or without the

---

[755] Wachter and Mittelstadt, "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and Ai," 79.

[756] Kuner, Bygrave, and Docksey, The Eu General Data Protection Regulation (Gdpr): A Commentary.

knowledge of the data subject.

As to the transparency of the processing, Recital 39 provides guidance. On this account, it is not only that the data subject should know that their personal data are collected, used, consulted or otherwise processed, to what extent the personal data are or will be processed, but also such information should be accessible and easy to understand meaning that a clear and plain language is used, as well as the audience of the information is considered such as in the case of data subjects being children. Asgharin et. al specifically stress the importance of contextuality, that an explanation regarding an ADM should consider the target audience and the purpose of such an explanation in order to provide meaningful transparency.[757]

Importance of transparency regarding ADM is also to enable data subjects to challenge lawfulness, fairness and accuracy of the decisions made about themselves. According to Wachter & Mittelstadt, while WP29's interpretation of the data protection law is in the direction that it aims to ensure accurate decision-making; CJEU's stance in *YS. And M. and S.,*[758] *Nowak*[759] and *Bavarian Lager*[760] shows that the Court is not in the opinion that the remit of data protection law is to ensure accuracy of decisions, as well as the reasoning and assessments behind them. Thus, even in the case where inferred data is considered to be objectively wrong, the data subject can only contest this situation under another procedure outside of data protection law.[761]

---

[757] Asghari et al., "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making.".

[758] C-141/12 and C-372/12 Joined Cases [2014] *YS (C-141/12) v Minister voor Immigratie, Integratie en Asiel, and Minister voor Immigratie, Integratie en Asiel (C-372/12) v M, S.* ECLI:EU:C:2014:2081

[759] C-434/16 [2017] *Peter Nowak v Data Protection Commissioner* ECLI:EU:C:2017:994

[760] C-28/08 [2010] *European Commission v The Bavarian Lager Co. Ltd.* ECLI:EU:C:2010:378

[761] Wachter and Mittelstadt, "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and Ai."

3.1.2.2 Purpose limitation vs. Data mining and ML

The purpose limitation principle (CFREU Art. 8[762], DPD Art. 6, GDPR Art. 5) is one of the cornerstones of European data protection law. It consists of two building blocks, namely purpose specification and compatible use. The first means that personal data should only be collected for 'specified, explicit and legitimate' purposes and it constitutes a prerequisite to all other data processing principles.[763] The latter building block of this principle requires that further processing must not be incompatible with the purposes for which personal data were collected.[764]

Article 6 (4) of the GDPR clarifies how a controller can ascertain whether the processing is compatible with the initial purpose by specifying that the controller shall take into account any link between the initial and further purposes, the context, the nature of the personal data (meaning whether special categories of data or data related to criminal convictions are being processed), the possible consequences for the data subjects and whether appropriate safeguards such as encryption or pseudonymisation are at place. On the other hand, as it draws on the right to private life under Art. 8 ECHR, ECtHR assesses whether the subsequent processing of personal data meets the reasonable expectations of the concerned individual.[765]

---

[762] As a result of being explicitly noted in CFREU, it had to be incorporated to GDPR and provides limited flexibility in defining its limits for the European legislators and regulators. See. Zarsky, "Incompatible: The Gdpr in the Age of Big Data.".

[763] According to von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Ii," 11.(Part II, p.11) it requires the controller to discover specific risks caused by its data processing against the data subject's autonomous exercise of fundamental rights.

[764] Article 29 Working Party, "Opinion 03/2013 on Purpose Limitation," WP 203, no. 00569/13/EN (2013), https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf.(pp.11-12)

[765] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part I."

It is argued that innovation in general and the Big Data analytics and data mining powered by machine learning clashes with the purpose limitation principle. First of all, the very core promise of Big Data analytics is to discover new meanings and knowledges from the vast amounts of data that is available to an entity. Such clash also occurs as a result of "relating to" constituent of personal data by reason of impact.[766]

GDPR Art.(5)(1)(b) allows for further processing for scientific research and statistical purposes provided that further processing is compatible with the initial purposes and subject to appropriate safeguards set forth in GDPR Art. 89(1). Accordingly, technical and organisational measures including pseudonymisation shall be taken in order to ensure data minimisation principle is respected.

In Recital 159, scientific research is broadly defined in a way that it includes technological development and demonstration, applied research and privately funded research. Thus it has been argued that some model development activities in private companies may be caught by this exemption.[767] Regarding the statistical purposes exemption, Recital 162 clarifies that the result of processing for statistical purposes are implied to be aggregate data and is not to be used in support of measures or decisions regarding any particular natural person. As it was explored in Chapter 1, even if the result of the development phase may be in the form of aggregate data, when the model is used in automated decision making it will fall out of the scope of processing for statistical purposes.[768] For example, if a company develops a model to predict overall employee churn, statistical purposes may apply. However, going further and predicting likely to leave employees and

---

[766] Purtova, "The Law of Everything. Broad Concept of Personal Data and Future of Eu Data Protection Law."

[767] Michèle Finck and Asia J. Biega, "Reviving Purpose Limitation and Data Minimisation in Personalisation, Profiling and Decision-Making Systems," *ArXiv* abs/2101.06203 (2021): 17.

[768] Ibid.

taking action to entice them such as giving an automated pay rise or promotion would make this operation fall out of statistical purpose.

GDPR 6(4) provides some examples to controllers, who wish to process personal data for further purposes, to take into account in order to ascertain the compatibility of initial and further processing. Zarsky argues that, safeguards provided in Art 6(4) are complex, difficult to execute and likely to undermine the whole process, and concludes that purpose specification principle is at odds with Big Data analysis.[769]

Maximillian von Grafenstein[770] argues, on the contrary that purpose limitation principle supports innovation by helping clarify for businesses early in the project a road map of what they want to do and to discover the risks thereof. According to von Grafenstein purpose specification indicates to determining the risk of the specific data processing to the fundamental rights and freedoms of data subjects. In this regard, it actually is grounded on an objective legal scale which provides predictability.[771]

Another criticism of purpose specification as a lead principle for privacy and data protection, is that currently there is no clarity into how detailed a purpose should be designed by the controllers. EDPB also doesn't seem to be clear on this. According to some scholars, processing or business operations of the controller may help specifying the purposes. However, many argue against it as it would mean the very core of data protection would be entirely determined by the controller as pretty much the whole legal system of data protection law depends

[769] Zarsky, "Incompatible: The Gdpr in the Age of Big Data."(p.1009)

[770] von Grafenstein, *The Principle of Purpose Limitation in Data Protection Laws: The Risk-Based Approach, Principles, and Private Standards as Elements for Regulating Innovation*, 12.

[771] Ibid.

on correctly specifying the purposes.[772] Nissenbaum argued business needs determined by the data controllers themselves does not provide a solid ground for privacy's moral imperative.[773] Nissenbaum rather proposes that the appropriateness of processing should be assessed its violation or conformity with social norms.[774] However, Kadiri points out that even though it is true that what is emancipatory and what is oppressive depends on the context, and such contextual understanding of data protection opens up space for acknowledging power relations within data processing activities and relations, we still have to be careful that dominant norms, which reflect structural racism, sexism and so on, would inform the boundaries of Nissenbaum's contextual integrity.[775]

Thus, it is suggested by some scholars that the data controllers should specify their purposes from the perspective of the data subject by taking the consequence for them in to account. Some argue gaining insight to this perspective by the data controller is not possible.[776]

According to von Grafenstein, the unmatching part of purpose principle has been the data determination aspect of it which refers to the understanding that the data subject decides how their data will be used later.[777] German Constitutional

---

[772] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part I," 513; Helen Nissenbaum, "Respect for Context as a Benchmark for Privacy Online: What It Is and Isn't," in *Social Dimensions of Privacy: Interdisciplinary Perspectives*, ed. Beate Roessler and Dorota Mokrosinska (Cambridge: Cambridge University Press, 2015).

[773] "Respect for Context as a Benchmark for Privacy Online: What It Is and Isn't."

[774] "Privacy in Context - Technology, Policy, and the Integrity of Social Life."

[775] Kadiri, "Data and Afrofuturism: An Emancipated Subject?."; ibid., 14-15.

[776] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part I," 512.

[777] Ibid.

Court in a recent decision[778] dropped granting this right to the data subjects by deeming it interference with the other private parties rights as they are principally free to collect process and use personal data of others and data subjects cannot unilaterally decide on the use of their data but can only influence it. This is because the personal data constitutes the basis of social interactions in the digitalised world.[779] We find such approach practical from a business and innovation, as well as application of the data protection law point of view, while finding it problematic from a point that power asymmetries between private parties as data controllers and data subjects would lead to the oppression of the weaker parties. Such approach is similar to the public/private dichotomy criticism in feminist legal theory according to which the state's negative duty of not to intervene in the private realm, where the market and private family life occurs, leads to women being oppressed both by capitalism in the free markets and by patriarchy in their private lives.

3.1.2.3 Data Minimization Principle vs. Big Data

The ongoing technological developments in data processing power promise that more data would lead to the discovery of more valuable information, therefore, from a futuristic point of view data controllers have an incentive to hold onto more data than they need at the moment. On the other hand, data minimization principle ensures that data controllers have minimum sources to undermine data protection rights of data subjects, minimizes the risk of cybersecurity threads, and supports data subjects' autonomy by reducing the anxiety of misuse.[780] The scope of protection the data minimization principle

---

[778] "Recht auf Vergessen I"Bverfg, (6th November 2019). 1 BvR 16/13 https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/DE/2019/11/rs20191106_1bvr001613.html

[779] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part I," 515-16.

[780] Zarsky, "Incompatible: The Gdpr in the Age of Big Data."

intends to provide have expanded in the GDPR compared to the DPD by change of wording from “not excessive in relation to the purposes” to “limited to what is necessary.”

Data minimization consists of three components, that data must be adequate, relevant, and limited to what is necessary which is determined by the purpose of the processing. New machine learning technologies are pushing the limits of relevance as they are able to find connections between distinct data points that can be meaningful for the same purpose. Adequacy component on the other hand can be at odds with the relevance component as in the case of more data being required in order to overcome unrepresented groups in datasets. Necessity component, on the other hand, would not let excessive amount of that in relation to the purpose to be used.[781]

Zarsky argues that data minimization requirements should have been loosened instead of tightened, and the righteous concerns addressed by this principle should be rather addressed with *ex post* regulation directed to unacceptable uses of data.[782] We do not agree with this position, especially because *ex post* intervention would be too late for the harms to general society to occur and most of the time undetectable and unchallenged as a result of individual rights based approach. We instead agree with Finck and Biega, that data minimization principle still holds a strong ground in the face of today’s technologies. While with relevance and necessity requirements it limits the amount of data that can be used, with adequacy it ensures the right kind of data to be put in use.[783]

---

[781] Finck and Biega, "Reviving Purpose Limitation and Data Minimisation in Personalisation, Profiling and Decision-Making Systems."

[782] Zarsky, "Incompatible: The Gdpr in the Age of Big Data."

[783] Finck and Biega, "Reviving Purpose Limitation and Data Minimisation in Personalisation, Profiling and Decision-Making Systems," 28.

3.1.2.4 Data Accuracy Principle

In ADM affecting WMG, the data gap which can be gendered, racialised, classed and so on, for example in training data may also affect the accuracy of the model's decisions specifically for these groups. Datafication of WMG then appears as a possible solution. Bu as we discussed before filling the data gap would be taking a technical short-cut to complex problems which is unlikely to solve them. It is not whether to close the data gap with more data collection on missing populations, or leaving the data gap as it is because it would clash with some data protection rules especially with respect to processing of special categories of data, but how and by whom to collect which data. Data protection can be turned into protection through data depending on the context which shall be determined by taking into consideration the historical power relations at the heart of the matter and to answer the questions of how and by whom to collect which data through this lens.[784]

Zarsky points out that pseudonymisation as a safeguard may clash with accuracy principle by undermining the quality of data.[785]

Data subjects are entitled to obtain rectification of inaccurate data concerning themselves or they may ask their data to be completed and provide information in that regard under Art. 16. The CJEU links this right to effective legal protection under CFREU Art. 47.[786] It is essentially linked to right to access under Art. 15[787] and will be examined together with Art. 22 below.

---

[784] Kadiri, "Data and Afrofuturism: An Emancipated Subject?."

[785] Zarsky, "Incompatible: The Gdpr in the Age of Big Data."

[786] C-362/14 [2015] *Maximillian Schrems v Data Protection Commissioner* ECLI:EU:C:2015:650 para 95

[787] Kuner, Bygrave, and Docksey, *The Eu General Data Protection Regulation (Gdpr): A Commentary*.

3.1.3 GDPR Article 22 (Automated Decision-Making Including Profiling)

Mendoza & Bygrave call Article 22 of the GDPR a bizarrely marginal right, given its relevance to the times we live in with increasing automation in almost every apect of life, as opposed to how it has rarely been enforced since its previous version as Article 15 in DPD[788], how it is poorly understood also thanks to its very complicated formulation and how easily it is circumvented.[789] It also differs from the other data subject rights under the GDPR as it targets a decision instead of data processing.[790]

3.1.3.1 Scope

Despite having almost identical wording, unlike its predecessor DPD Article 15, Article 22 GDPR caught a great public and academic attention, most possibly because of the broad reach of the GDPR combined with much higher fines compared to DPD alongside with the time it was enacted and the scope it applies that is capable of capturing some of the high risk technologies which are embedded in daily life and came to public attention with interconnected events of Cambridge Analytica scandal, Brexit campaign and 2016 US presidential elections leading to Donald J. Trump's win.[791]

---

[788] See https://gdprhub.eu/index.php?title=Category:Article_22_GDPR for the recent cases related to GDPR Art. 22.

[789] Isak Mendoza and Lee A. Bygrave, "The Right Not to Be Subject to Automated Decisions Based on Profiling," in Eu Internet Law: Regulation and Enforcement, ed. Tatiana-Eleni Synodinou, et al. (Springer, 2017), 78.

[790] Lee A. Bygrave, "Automated Profiling: Minding the Machine: Article 15 of the Ec Data Protection Directive and Automated Profiling," Computer Law & Security Review 17, no. 1 (2001): 17.

[791] Jamie Doward and Alice Gibbs, "Did Cambridge Analytica Influence the Brexit Vote and the Us Election?," *The Guardian* 2017. <https://www.theguardian.com/politics/2017/mar/04/nigel-oakes-cambridge-analytica-what-role-brexit-trump> Accessed on 20 September 2022

3.1.3.2 Profiling

Profiling is defined in Art. 4(4) GDPR as follows:

"any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements"

Hildebrandt suggests a simple definition of profiling to be "The process of 'discovering' correlations between data in databases that can be used to identify and represent a human or nonhuman subject (individual or group) and/or the application of profiles (set of correlated data) to individuate and represent a subject or to identify a subject as a member of a group or category."[792]

Most simply, profiling consists of two data processing steps, (i) generating a profile by inferring a set of characteristics about and individual or a collective entity, (ii) treating that individual/entity or others in a certain way by being informed by these characteristics.[793] The definition in Art. 4(4) adds two conditions to this understanding that is (i) being limited to automated processing of personal data (ii) which evaluates a natural person. In order to fall under Art. 22 GDPR, profiling has to be solely automated. The first criterion does not exclude human input or analogue data to be used in forming a profile, and the latter involves both the generation of a profile and application of it. According to Article 29 Working Party, a business which classifies their clients into their age or gender to acquire aggregated overview of their clientele does not amount to profiling unless it does

[792] Hildebrandt, "Defining Profiling: A New Type of Knowledge?."; ibid., 19.

[793] Bygrave, "Automated Profiling: Minding the Machine: Article 15 of the Ec Data Protection Directive and Automated Profiling," 17.

not assess any individual characteristics to make predictions and draw conclusions about an individual. While ADM may or may not include profiling, profiling is also possible without basing an automated decision on it.[794]

Data on which profiling or ADM or both are going to be based, can be collected in various ways including manual and semi-automated processes.[795] However, cookies and other tracking technologies that transfer online and off-line human life into digital data points to be stored in massive databases in order to be analysed by machine learning and deep learning models to find out new patterns and correlations which help to sort people and predict and even nudge their future behaviour is what the biggest technology companies' business plans rest upon.

Online Behavioural Advertising (OBA)[796] is at the core of most problematic job ad delivery in terms of discrimination risk[797] and microtargeting[798] is making its effects even more difficult to detect.

---

[794] Article 29 Working Party, "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679". 7.

[795] Lee A. Bygrave, "Article 22 Automated Individual Decision-Making, Including Profiling," in The Eu General Data Protection Regulation (Gdpr): A Commentary, ed. Christopher Kuner, et al. (New York: Oxford University Press, 2020).

[796] Article 29 Working Party, "Opinion 2/2010 on Online Behavioural Advertising ".

[797] Wachter, "Affinity Profiling and Discrimination by Association in Online Behavioural Advertising (May 15, 2019). , Vol. 35, No. 2, 2020, Forthcoming, Available at Ssrn: Https://Ssrn.Com/Abstract=3388639 or Http://Dx.Doi.Org/10.2139/Ssrn.3388639."; Article 29 Working Party, "Opinion 2/2010 on Online Behavioural Advertising ".

[798] EDPS Opinion 3/2018, "European Data Protection Supervisor's Opinion on Online Manipulation and Personal Data."; EDPB Guidelines 8/2020, "On the Targeting of Social Media Users Version 2.0," (2021).

3.1.3.3 The Nature of the Right (Art 22/1)

The nature of Art. 22 has been controversial. The majority of academics[799] and WP29[800] interpret it as qualified prohibition/a passive right or agree that this interpretation offers greater protection,[801] while others argue that it is designed as a right to be exercised actively by the data subject.[802] The first means that data controllers are not allowed to subject data subjects to solely automated decisions unless a condition in Art. 22(2)(a)-(c) is met. The latter, on the other hand, means that data subjects generally have a right to object to being subjected to such a decision unless one of the requirements in Art. 22(2)(a)-(c) is met. In other words, second interpretation burdens data subjects with playing an active role, which requires that the data subject is aware of the ADM with its envisaged consequences and willing/able to intervene.[803] There are three major problems in interpreting Art. 22 as a right to object: (i) it leaves automated decisions legally unchallenged until an objection is entered even if it does not meet any of the requirement set out in Art. 22(2), (ii) automated decisions which do not fall under Art. 22(2)(a)-(c) also do not require safeguards under Art. 22(3) to be provided, (iii) instead of levelling the informational power asymmetry between the controllers and data subjects, especially due to technical complexity of ADM processes, as well as the need of the data subject for such a decision to be made (for example, in cases of e-

[799] Mendoza and Bygrave, "The Right Not to Be Subject to Automated Decisions Based on Profiling."; Maja Brkan, "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the Gdpr and Beyond," International Journal of Law and Information Technology 27, no. 2 (2019).

[800] "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679".

[801] Wachter, Mittelstadt, and Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation."

[802] Luca Tosoni, "The Right to Object to Automated Individual Decisions: Resolving the Ambiguity of Article 22(1) of the General Data Protection Regulation," ibid.11 (2021).

[803] Wachter, Mittelstadt, and Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation."

recruitment, the lower a person's position in the matrix of domination the more likely they may be to submit), it reinforces such asymmetry by remaining blind to it. On this account, we agree with Wachter et. al that interpreting the nature of article 22 as a prohibition provides higher protection to the data subjects including the most vulnerable as it protects by default.[804]

On the other hand, we also agree that the nature of Article 22 remains vague. Analysis regarding the wording and legislative history of the GDPR are to a certain level convincing[805] that Art. 22(1) (The data subject *shall have the right* not to be subject to a decision based solely on automated processing) might have meant to be a right by intentionally not using the word prohibition which was explicitly used for example in Law Enforcement Directive's Article 11 (Member States shall provide for a decision based solely on automated processing, including profiling … to be *prohibited* unless authorised by Union or Member State law) also regulating ADM or Art. 22(4) GDPR (Decisions referred to in paragraph 2 *shall not be based* on special categories of personal data). It is important to note that related Recital 71 of the GDPR states that ADM including profiling *"should be allowed"* where Art. 22(2)(a)-(c) applies which indicates to ADM including profiling is prohibited in general, but it should be allowed where Art. 22(2)(a)-(c) applies. On the other hand, LED Recital 38 indicates to a right[806] and the word prohibition is only used for decisions based on sensitive attributes that lead to discrimination. Even though Recitals are not binding, they are to interpret what EU Regulation means[807] and here in this discussion they just blur the waters.

---

[804] Ibid.

[805] Tosoni, "The Right to Object to Automated Individual Decisions: Resolving the Ambiguity of Article 22(1) of the General Data Protection Regulation."

[806] "…data subject should have the right not to be subject to a decision evaluating personal aspects relating to him or her which is based solely on automated processing…"

[807] Tadas Klimas and Jurate Vaiciukaite, "The Law of Recitals in European Community Legislation," ILSA Journal of Int'l & Comparative Law 15, no. 1 (2008),

Among two scholars who thoroughly compared these two legislations, Brkan concluded that interpreting Art. 22(1) GDPR as a general prohibition aligns it with the wording of Art. 11(1) LED and ensures coherence of Data Protection Law of the EU.[808] On the other hand, Tosoni concluded that this difference between the two texts is intentional as the two instruments also differ in their nature. As a result of the coercive powers allocated to the competent authorities regulated under LED such as restraining personal liberty, automated decisions made under this context are likely to affect their addressees more substantially, thus it is justified to apply more stringent rules to them.[809] This is a very classical interpretation of power. In this regard, state power that is exercised in the public sphere would be subject to restriction, while the power relations in the private sphere would not be interfered with. For instance for women, refused credit applications or job advertisements that they are not shown would lead to them being not able to leave an abusive relationship in which they are trapped by means of economic abuse.[810] Such power imbalance in which WMG are more likely to find themselves would lead to a spiral of violence of different forms.[811] Also, the great power imbalance between platforms and individuals may have detrimental effects on their autonomy which is a prerequisite for personal liberty.

Consistency with the international agreements concluded by the EU is also

---

https://www.researchgate.net/publication/228152770_The_Law_of_Recitals_in_European_Community_Legislation.

[808] Brkan, "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the Gdpr and Beyond."

[809] Tosoni, "The Right to Object to Automated Individual Decisions: Resolving the Ambiguity of Article 22(1) of the General Data Protection Regulation."

[810] Judy L. Postmus et al., "Economic Abuse as an Invisible Form of Domestic Violence: A Multicountry Review," Trauma, Violence, & Abuse 21, no. 2 (2020).

[811] Alina Wernick and Deniz Erden, "Computer Says Hausfrau - Can Automated Credit Scoring Contribute to Gendered Digital Divide?," Encore - The Annual Magazine on Internet and Society Research 2020/2021 (2021).

important for CJEU's interpretation. As the Member States were authorized to ratify the Modernised Convention n. 108 in the interest of the European Union, it was negotiated in parallel with the GDPR and is fully in line with it, CJEU might consider Art. 9(1)(a) in its interpretation of Art. 22.[812] The said provision according to Tosoni establishes an active right to be exercised by the data subject.[813]

Another point is that, the DPD Art. 15 and GDPR Art. 22 have identical texts despite some differences that do not change the meaning . During the time DPD Art. 15 was in effect, Austria, Belgium, France, Germany, Finland, the Netherlands, Portugal, Sweden and Ireland interpreted it as a general prohibition while the UK and Norway did as a right to be actively brought up by the data subject .

In this regard, there is still room for the CJEU, the only body who has the competence to authoritatively interpret the EU law to consider Art. 22 as a right to be actively pursued by the data subject despite the opinion of WP29/EDPB, as well as the majority opinion in the academia.[814] This can be the case, if the CJEU justifies such decision on the fact that in case of ADM and profiling with significant effects would need to undergo a data protection impact assessment (DPIA) under Art. 35 GDPR. In this case, if the high risk is discovered and could not be mitigated the Data Protection Authority (DPA) would also be notified according to Art. 36 GDPR. Thus the CJEU may come to the conclusion that highest protection is provided even for the decisions that do not fall under Art. 22(2)(a)-(c).[815]

---

[812] Tosoni, "The Right to Object to Automated Individual Decisions: Resolving the Ambiguity of Article 22(1) of the General Data Protection Regulation," 21-22.

[813] Ibid.

[814] Ibid.

[815] Ibid.; Bygrave, "Article 22 Automated Individual Decision-Making, Including Profiling."

3.1.3.4 Conditions

In order Art. 22 (1) to be applicable three cumulative conditions that we discuss in detail below have to be met.

3.1.3.4.1 A Decision

What is interesting about Article 22 is that it concerns with a decision while the rest of GDPR provisions are about processing of personal data. Bygrave speculates that the decision here means a particular stance or attitude to be taken towards a person with the intention of acting upon it.[816] There is no requirement specified as to the form of the decision, however according to Bygrave, it must be distinguishable from other processing activities that lead to making the decision. The choice of the business to label the decision as something else also does not matter if de facto it is understood to be a decision.[817] Ivanova argues that processing in training and test phases is not covered by Article 22 as data subjects whose data is used at these stages are not yet subject to any decision which may have legal or similarly significant effects on them.[818] While the question is not resolved yet, applicability of Article 22 to activities at this phase shall be subject to more nuanced evaluation. [819]

---

[816] "Article 22 Automated Individual Decision-Making, Including Profiling."

[817] Ibid.

[818] Ivanova, "The Data Protection Impact Assessment as a Tool to Enforce Non-Discriminatory Ai," 4.

[819] For an analysis of training data under the GDPR see also Wachter and Mittelstadt, "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and Ai."; Finck and Biega, "Reviving Purpose Limitation and Data Minimisation in Personalisation, Profiling and Decision-Making Systems."; Hacker, "A Legal Framework for Ai Training Data—from First Principles to the Artificial Intelligence Act."

3.1.3.4.2 Based Solely on Automated Data Processing

Being "solely" based on automated data processing means no meaningful human involvement in making the decision. If an automated system is producing a recommendation in the direction a decision can be made but the final decision is being made by a human by weighting in other factors than the recommendation made by the automated process then it does not count as a solely automated decision. However, businesses cannot escape Article 22 by appointing a human to approve or stamp such automated recommendation. According to the WP29, human involvement that would deem the decision being not solely automated requires the involved human to have the authority and competence to change the decision as well as considering all the relevant data in their analysis.[820] It is also suggested[821] whether "solely" condition is met would be assessed via a DPIA.

3.1.3.4.3 Legal or similarly significant effects

According to WP29, legal effects occur when an individual's legal rights, legal status or rights under a contract are affected. Similarly significant effects, on the other hand, refer to an impact that still requires protection. WP29 gives an effect on freedom to associate with others as an example to legal affects, while opines that affecting an individual in a manner that leads to exclusion or discrimination is to be where a similarly significant effect gets most extreme.[822] In Recital 71, e-

[820] Article 29 Working Party, "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679"; ibid., 21.

[821] Bygrave, "Article 22 Automated Individual Decision-Making, Including Profiling." Michael Veale and Lilian Edwards, "Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling," Computer Law & Security Review 34, no. 2 (2018); Bygrave, "Article 22 Automated Individual Decision-Making, Including Profiling."

[822] Article 29 Working Party, "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679".

recruitment and automatic refusal of credit application are exemplified as similarly significant effects, in this regard we are in the opinion that job ad delivery on social media would also fall under this category. Unless the targeting based on profiling is not (i) too intrusive in the sense that tracking the individual across websites, applications and devices; (ii) clashing with the expectations of the data subject; (iii) has an impact on the data subject because of the way it is delivered; (iv) utilizes information on the data subject's vulnerabilities, WP29 opines targeted advertisement based on simple categories including age and gender does not amount to similarly significant effect. However, an ADM that excludes an individual from a job opportunity by not delivering an advert based on their gender or age should be considered to have a similarly significant effect on that individual as it involves exclusion and discrimination based on protected attributes. On this account, such practices reduces the universe of those who deemed not suitable to receive such advertisement and this creates social sorting.[823]

An important point left unclear by the WP29 guidance as noted by Veale and Edwards is whether the ADM requires to similarly significantly affect individuals or a group in which the data subject belongs.[824] In the first sight, WP29 acknowledges that possibility of the significance of the effect of an ADM might occur at the group level such as minority groups or adults with vulnerabilities while being insignificant at the individual level. However, the vulnerability of an individual with the gambling problem does not match the problems that we have identified as group level problems. WP29 also refers to situations where an individual may be deprived of an opportunity based on the actions of others such as in the case of an individual's credit card limit being reduced due to the analysis of non-traditional credit data consisting of the behaviour of other individuals sharing properties with them. Similarly, women not being offered certain jobs or

[823] Veale and Edwards, "Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling."

[824] Ibid.

positions based on other women's behaviour of not applying to such jobs or positions can also be considered in the same category.[825]

3.1.3.5 Derogations (Art. 22(2))

In case Art. 22(1) is understood as setting a general prohibition on ADM which carries the abovementioned conditions, Art. 22(2) introduces the first layer of circumstances where this prohibition does not apply, in other words under which conditions fully automated decisions with legal or similarly significant effects are allowed to be made.

3.1.3.5.1 Contract (Art. 22(2)(a))

Mendoza and Bygrave point out how recourse to a standardised contract by the controller to escape the prohibition under Art.22 is made more difficult by adding a necessity assessment under Art. 22(2)(a) different from the previous version under DPD Art. 15.[826] While they[827] found it hard to think of a decision that has to be made without human involvement, WP29[828] extends the scope to pre-contractual processing by giving an example from e-recruitment context that a high volume of applications for a position may justify the necessity for the screening stage to be fully automated. Thus in these cases, safeguards in Art. 22(3) including the right to explanation (if it exists) apply.

---

[825] Bogen and Rieke, "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias."

[826] Mendoza and Bygrave, "The Right Not to Be Subject to Automated Decisions Based on Profiling."

[827] Ibid.

[828] Article 29 Working Party, "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679".

3.1.3.5.2 Authorisation by EU or Member State Law

The level of protection available to the data subjects under this provision will be determined mostly by the national legislation. Here, Art. 22(3) does not apply so that the minimum safeguards of human intervention, expressing opinion, and contesting are not imposed, meaning that it is left to the discretion of the Member States to implement suitable measures. As Regulation (EU) 2018/1725 on the protection of natural persons with regard to the processing of personal data by the EU institutions, bodies, offices and agencies and on the free movement of such data (hereinafter EUDPR) and Directive (EU) 2016/680, which is also known as the Data Protection Law Enforcement Directive (hereinafter LED) expressly mention right to obtain human intervention, this is expected to be the key safeguard for the ADM falling under this scope.[829]

3.1.3.5.3 Data Subject's Explicit Consent

Under DPD consent was not among derogations from Art.15, so it is new under GDPR. Similar to heightened protection in Art. 9(2)(a), here explicit consent of the data subject is required. Article 22(4) also prohibits automated decisions to be based on special categories of personal data, unless the legal basis of the processing is the explicit consent or public interest as laid down in Art. 9. In order to be valid, consent has to be freely given, specific, informed and unambiguous, all of which may turn into a hurdle.

First of all, the freely given component becomes problematic in take-it-or-leave-it situations.[830] In such cases, similar to the necessity test under Art. 22(2)(a),

---

[829] Bygrave, "Article 22 Automated Individual Decision-Making, Including Profiling."

[830] See Noyb complaints against Google, Instagram, WhatsApp and Facebook. <https://noyb.eu/en/noybeu-filed-complaints-over-forced-consent-against-google-instagram-whatsapp-and-facebook> Accessed on 10 September 2022.

the data controller should answer whether ADM is necessary for the offered services as stated in the Art. 7(4). Another criteria suggested to weight in the necessity test is whether the data subject has an interest in the ADM such as in the case of a loan application, whereas it is likely not to be possible to argue that the consent is not freely given for price differentiation. In our opinion these are cursory examples that do not reflect the variety of scenarios that may occur within the same contexts. While some people benefit from the price discrimination, some people may consent to bad deals due to their vulnerabilities especially in the fast-paced world of online contracts.

As the #metoo movement has brought this topic to a more mainstream understanding, where there is a power imbalance, it is hardly possible to recourse to free choice and consent.[831] This is acknowledged by the Article 29 Working Party in their Guidelines on consent under Regulation 2016/679, under the title of "Imbalance of power," acknowledging that when the controller is a public authority or the employer in employment context, relying in consent is not suggested except for very rare situations. WP29 puts it clearly by stating *"Given the dependency that results from the employer/employee relationship, it is unlikely that the data subject is able to deny his/her employer consent to data processing without experiencing the fear or real risk of detrimental effects as a result of a refusal."*[832] Some scholars have stressed how users are actually working for platforms by providing content and their personal data.[833] Moreover, as mentioned above, the only vulnerable category explicitly written in GDPR is the children

[831] Malgieri and Niklas, "Vulnerable Data Subjects."; Olsen, "Statutory Rape: A Feminist Critique of Rights Analysis."; Carole Pateman, The Disorder of Women : Democracy, Feminism, and Political Theory (Stanford, Calif.: Stanford University Press, 1989); MacKinnon, Toward a Feminist Theory of the State; Nicola Lacey, Unspeakable Subjects: Feminist Essays in Legal and Social Theory, 1 ed. (London: Hart Publishing, 1998); Daniel Loick, ""… as If It Were a Thing." a Feminist Critique of Consent," Constellations 27, no. 3 (2020).

[832] "Guidelines on Consent under Regulation 2016/679," WP259.

[833] Siapka and Biasin, "Bleeding Data: The Case of Fertility and Menstruation Tracking Apps."

whose age of consent is affected accordingly which is based on decisional vulnerability.[834] Thus data subjects are mostly trusted for making rational choices once they are sufficiently informed, which is the way to bridge the gap in informational power asymmetry.

Targeting online advertisements, especially real-time bidding is an area where collecting valid consent is difficult and necessary.[835] According to WP29[836] in such online environments controllers may resort to layered consent.[837] Recital 42 requires safeguards to be put in place in order to ensure that the data subject is aware that they are giving consent and know the extent thereof. However, if a non-personalized version of a web service is not available, is the consent still freely given?[838] Another crucial point that determines whether the consent was freely given is that the data subject can withdraw their consent without detriment.

The withdrawal of consent according to Art. 7(3) does not affect the lawfulness of the processing before the consent was withdrawn. If the data controller has another legal ground the data can be retained, even though the processing based

---

[834] Malgieri and Niklas, "Vulnerable Data Subjects."

[835] Legitimate interest as a legal basis for real time bidding and targeting advertisement online was rejected. See the Litigation Chamber decision on the merits 21/2022 of 2 February 2022 of Belgian Data Protection Authority (Case number: DOS-2019-01377) <https://edpb.europa.eu/system/files/2022-03/be_2022-02_decisionpublic_0.pdf> accessed 24 April 2022.

[836] Article 29 Working Party, "Guidelines on Consent under Regulation 2016/679".

[837] See the decision of the CNIL's restricted committee (CNIL, 21 January 2019), which imposed a financial penalty of 50 Million euros against GOOGLE LLC. One of the reasons leading to the infringement decision was that Google's privacy notice was too scattered. <https://www.fieldfisher.com/en/services/privacy-security-and-information/privacy-security-and-information-law-blog/cnil-issues-50-million-euro-fine-against-google-in-the-first-major-gdpr-infringement-case> Accessed on 10 September 2022.

[838] Finck and Biega, "Reviving Purpose Limitation and Data Minimisation in Personalisation, Profiling and Decision-Making Systems," 22.

on consent has to stop. There can be cases though, data which forms a part of a trained model may no longer processed. Currently it seems, in such a case, removal of that data is only possible through retraining the model.[839]

Overall, obtaining freely given, specific, informed and unambiguous consent of data subjects in most ADM scenarios is close to impossible in practice, which makes it very unfortunate that consent is among the derogations.

3.1.3.6 Safeguards

Article 22(3) introduces some safeguards to contract and consent derogations introduced in Article 22(2).

As useful as these safeguards are, a shift of attention occurred fuelling a lively scholarly debate on whether among these safeguards data subjects are granted a right to *ex post* explanations, or whether data controllers are obliged to open the black boxes of their algorithmic systems if requested. We will briefly introduce the actual safeguards listed in article 22(3) and then move to the right to explanation debate.

3.1.3.6.1 GDPR Article 22(3)

According to Article 22(3), data controllers are obliged to grant a data subject who was subjected to a fully automated decision based on the contractual relation they entered into or their explicit consent, at least the right to obtain human intervention, to express their point of view and to contest the decision, in order to safeguard their rights and freedoms, as well as legitimate interests.

The list of safeguards are not exhaustive, so data controllers are welcomed to

[839] Ibid., 24.

provide more protection. For example, Hacker argued, that bias minimization strategies must be considered as one of these safeguards, due to the reference in Recital 71[840] that the controller should use appropriate mathematical or statistical procedures, correct inaccuracies in personal data, minimise the risk of errors in order to prevent discriminatory effects on natural persons.

According to Bygrave, unlike the situation in DPD, a data subject is always going to have the right to obtain human review unless in cases where the fully automated decision is based on a statute. In cases where the data subject exercises their right in Art. 22(1) after a fully automated decision is made (*ex post*), there is only a tiny difference between the level of protection provided by these two paragraphs.[841]

3.1.3.6.2 Right to Explanation (Recital 71) and Meaningful Information about the Logic Involved (Articles 13(2)(f); 14(2)(g); 15(1)(h))

Whether GDPR grants data subjects a right to explanation and if yes what it covers has been central to hottest academic debates about the GDPR. It all started with Goodman and Flaxman[842] claiming that the GDPR brings a right to explanation and Wachter et. al[843] replying to them by arguing that GDPR does not grant a right to explanation even though it would have been nice to have. Then

[840] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 30.

[841] Bygrave, "Article 22 Automated Individual Decision-Making, Including Profiling."

[842] Goodman and Flaxman, "European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"."

[843] Wachter, Mittelstadt, and Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation."

Selbst and Powles[844] accused Wachter et. al[845] for irresponsibly choosing a sensational title that gives the impression of debunking the existence of a right to explanation in the GDPR while failing to do so in their article and criticized their rigid[846] interpretation of the law. Since then both the existence of a right to explanation and if it exists the source and scope of such right have been controversial.

Wachter et. al[847] started with distinguishing between two possible kinds of explanations the GDPR may have meant: (i) system functionality, and (ii) specific decisions. They also distinguished between the timing of such an explanation as to whether it would require *ex ante* or *ex post* explanation of the ADM process. According to the authors, meaningful information about the logic involved in the GDPR includes system functionality but not the specific decisions. However, for most ML models such a distinction would be unnecessary as it is argued in the literature that as long as the input is known and the system functionality (meaning the model) is explainable, then it would give the same output each time it is given the same input and thus the specific definition would also be explainable.[848] According to Selbst and Powles,[849] Wachter et. al[850] were able to argue against the existence of a right to explanation only because they equated it with an *ex post*

---

[844] Selbst and Powles, "Meaningful Information and the Right to Explanation."

[845] Wachter, Mittelstadt, and Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation."

[846] See also Brkan, "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the Gdpr and Beyond."

[847] Wachter, Mittelstadt, and Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation."

[848] Selbst and Powles, "Meaningful Information and the Right to Explanation."; Asghari et al., "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making.".

[849] Selbst and Powles, "Meaningful Information and the Right to Explanation."

[850] Wachter, Mittelstadt, and Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation."

explanation of a specific decision within the framework they themselves created. However, a rather broader approach would be much more preferable for providing individuals with meaningful information as discussed by Asghari et al.[851] that explainability starts even before the system is used. In this regard, also the timing the explanation to be provided as well as the content of the explanation depends on the task the system is performing, context and the target audience.

When we turn to the text of the GDPR and how it has been interpreted in the literature, there is consensus on the facts that Art. 22 (3) does not mention a right to obtain an explanation, only Recital 71 explicitly mentions an explanation to be granted (*to obtain an explanation of the decision reached after such assessment*) and that Recital 71 is not binding.

The wording of the information duty of the controller in Articles 13(2)(f)-14(2)(g) and the right of access of the data subject in Article 15(1)(h) are identical: *"the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject."*

According to the WP29 providing meaningful information about the logic involved requires rationale behind or the criteria relied on reaching a decision to be told to the data subject in a simple format in order to ensure that the data subject understands the reason for the decision. According to Asghari e. al[852] the logic involved could be understood as the structure and the sequence of the data processing which means a global explanation or as it was termed by Wachter et.

[851] Asghari et al., "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making.".

[852] Ibid.

al[853] the system functionality.[854] However, a complex mathematical explanation about the entire technical functioning of the ADM system is not required. Most important is to keep in mind the reason for such information to be given to the data subject which is to strike a balance in the power imbalance between the data controller and the data subject to ensure the autonomy of the latter. In this regard, the data subject should not only be able to understand the decision but also to appeal against it.[855]

Wachter et. al[856] argued that deriving a right to *ex post* explanation of a specific decision from Articles 13(2)(f)- 14(2)(g) would suffer from a timeline problem as these articles are meant to be about providing information at the time when personal data are obtained (Art. 13) and prior to further processing (Art. 14) in the case personal data have not been obtained from the data subject, in which case the ADM would be considered further processing. Despite Wachter et. al[857] also admitting that the same timing problem does not occur with Article 15(1)(h) as the data subject may exercise this right anytime, they still argue that the wording of this article is semantically future oriented because of the use of "*the envisaged consequences."* While Selbst and Powles[858] argued that a general explanation would most of the time provide an explanation also about the specific decision, thus the specific decision and system functionality distinction is not really

---

[853] Wachter, Mittelstadt, and Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation."

[854] "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making.". 18.

[855] Ibid.; Article 29 Working Party, "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679"; Brkan, "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the Gdpr and Beyond."

[856] Wachter, Mittelstadt, and Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation."

[857] Ibid.

[858] Selbst and Powles, "Meaningful Information and the Right to Explanation."

meaningful; Brkan emphasized that even after a specific decision is made some consequences of the decision would still be unknown and require envisioning.[859] According to WP29 envisaged consequences of the processing does not equal to an explanation of a particular decision, however, the data subject should be provided *"with general information (notably, on factors taken into account for the decision-making process, and on their respective 'weight' on an aggregate level)"* aiming at it is useful for challenging the decision.[860] As Brkan puts it, contesting a decision as referred to under Art. 22(3) is closely linked to the substance of the specific decision made about the data subject without which this safeguard would appear to be an empty shell.[861] Mendoza and Bygrave are also in the opinion that right to contest indicates to more than just being able to object and stop the processing but amount to appealing an already made decision.[862]

Thus, it is argued that a right to explanation can actually be derived from reading Articles 13(2)(f), 14(2)(g), and 15(1)(h) together with Article 22 in the light of Recital 71. Reading various provisions of a legislation together with the Recitals has been a way that the CJEU interpreted the data protection legislation . It is established in the case law of the CJEU that *"the preamble to a Community act has no binding legal force and cannot be relied on either as a ground for derogating from the actual provisions of the act in question or for interpreting those provisions in a manner clearly contrary to their wording."* However, given that recourse to the recitals to resolve ambiguity in related legislative provisions

[859] Brkan, "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the Gdpr and Beyond."

[860] Article 29 Working Party, "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679"; ibid., 27.

[861] Brkan, "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the Gdpr and Beyond."

[862] Mendoza and Bygrave, "The Right Not to Be Subject to Automated Decisions Based on Profiling."

is possible in the EU law,[863] relying on Recital 71 in order to resolve the ambiguity of what the related GDPR provisions mean by providing meaningful information to the logic involved in order to contest ADM would stay within this scope and would not amount to derogating from the actual provisions of the GDPR or interpreting Articles 13(2)(f), 14(2)(g), 15(1)(h) and Article 22 in a manner contrary to their wording. It is also noted that, safeguards listed in Art. 22(3) are the bare minimum of what the data controller can provide in order to safeguard the data subject's rights and freedoms and legitimate interests as the article mentions *at least* those to be done.[864]

Finally, the principle of transparency in Article 5(1)(a) and transparent information obligation of the data controller regarding the rights of the data subject in Article 12 reflect the high transparency of data processing aimed by the GDPR which ensures individual transparency or in other words individual data subjects to have a right to know what is being done with their personal data and what the consequences thereof might be in an understandable manner enough to be used to challenge such processing especially in the form of ADM with significant effects.[865] One of these significant effects is indeed the risk of discrimination that stems from ADM as acknowledged in the second paragraph of Recital 71. In order to challenge biased decisions, data subjects would need to understand the reasons

[863] Klimas and Vaiciukaite, "The Law of Recitals in European Community Legislation". 26.

[864] Brkan, "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the Gdpr and Beyond."; Mendoza and Bygrave, "The Right Not to Be Subject to Automated Decisions Based on Profiling."

[865] Brkan, "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the Gdpr and Beyond."; Asghari et al., "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making."; Mendoza and Bygrave, "The Right Not to Be Subject to Automated Decisions Based on Profiling."

behind such behaviour of the model[866] which might be also insightful in showing prima facie discrimination under antidiscrimination law in indirect discrimination cases stemming from ADM.[867] Brkan argues that especially for decisions where the ADM was based on special categories of data, general information about the functioning of the system would not be sufficient to safeguard the rights and freedoms and the legitimate interests of the data subjects without providing that their special category data was involved in that specific decision.[868]

3.1.4 Accountability & Risk Based Approach

The way data protection laws are designed to impose positive obligations requires a constant process of compliance activities for the controllers as they are responsible both for *ex ante* controls and *ex post* harms. By adding accountability in Art. 5(2) among its main principles laid down in Art. 5(1) with which controllers are obliged to comply with, the GDPR goes one step further than usual compliance practices and further asks for the ability to demonstrate such compliance. This is laid down openly in Art. 24 that "the controller shall implement appropriate technical and organisational measures to ensure and to be able to demonstrate that processing is performed in accordance with" the GDPR.

In doing so the controller shall first take into account the nature, scope, context and purposes of processing. In addition, risks for the rights and freedoms of natural persons shall be assessed based on the likelihood and severity. As it is often called as a risk based approach, GDPR requires data controllers to engage with several risk assessments. Legitimate interest assessment (Art. 6 (1)(f)), data protection by

---

[866] Brkan, "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the Gdpr and Beyond."

[867] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."

[868] Brkan, "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the Gdpr and Beyond."

design (Art. 25), security (Art. 32), Data Protection Impact Assessment (Art. 35), as well as codes of conduct (Art. 40), and certification mechanisms (Art. 42) involve further risk assessments. Some scholars warned that, mechanisms brought under accountability principle such as DPIAs, Privacy by Design (PbD), certification schemes are dangerous in the sense that if done with a formalistic approach, they are more likely to become box ticking compliance work and less likely to bring substantive change for the protection of human rights and fundamental freedoms of the data subjects or groups such as WMG.[869]

Others argued, due to the referral in Art. 1(2) to fundamental rights and freedoms of the natural persons without being limited to the protection of personal data as its objective, the risk based approach involves the consideration of all the fundamental rights and freedoms and with respect to all the GDPR provisions even though that provision does not assign a risk assessment obligation to the data controller.[870]

In this part, we will focus on the *ex ante* protection such risk based approach may provide within the context of the algorithmic discrimination. As mentioned in the previous Chapter, seeking protection against algorithmic discrimination under the EU antidiscrimination law could be cumbersome, if at all a situation falling under its scope emerged, due to limitations of protected grounds and material scope of it. And in any case, beyond the will of companies to respect and comply with the law to avoid litigation, protection under EU antidiscrimination law is *ex post.*

Data protection law also have *ex post* protection mechanisms but most

[869] Edwards and Veale, "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For."

[870] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Iii."

importantly it brings in real obligations to consider algorithmic discrimination and to take necessary measures throughout the lifecycle of ADM processes to ensure it does not occur.

3.1.4.1 Responsibility of the Data Controller (Art. 24)

Data controllers are obliged to ensure their data processing activities respect fundamental rights and freedoms of the data subjects which involve taking all the necessary measures to not cause discrimination too. Protection of all the data subject rights are under the data controller's responsibility and also directed to the controllers.

Thus, it is crucial to determine who the data controller is. As we introduced earlier, in the employment context there can be various actors at different stages. Employers can make use of these technologies via procurement, via service contracts, by making use of social media or other platforms via pages, targeted ads, apps, plugins and so on. Thus different kinds of controller, joint controller, controller-processor relationships may occur.

Two problematic examples are job ad delivery and procurement of hiring or other HR management tools from third parties. In the job ad delivery example, according to CJEU case law[871] joint controllership of the platform and the employer may arise. However, according to the antidiscrimination law, the protection is limited to two protected grounds, namely race/ethnicity and sex.

In the procurement, it is up to discussion if the model is being developed in a separate company under scientific research exception and the end product only has aggregate data as output which may fall under statistical use exception. The model

[871] C-210/16 [2018] *Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie Schleswig-Holstein GmbH* ECLI:EU:C:2018:388 para 42-43.

then may have enjoyed some liberty in the way data is used in development and probably cannot be first detected and then hold liable if then it causes discriminatory results in the employer's practices as in this case for the processing activity of the employer, the developer doesn't have a processor status.

3.1.4.2 Privacy by Design and Default (Art. 25)

With the PbD obligation under Art. 25 GDPR, data controllers have to implement technical and organisational measures in order to protect the rights of the data subjects from the very beginning of data collection, preparation and model design, instead of having it as an afterthought.

According to Dreyer and Schulz, even though consideration of threads to the rights of individuals stemming/likely to stem from ADM could help eliminating some of them, PbD concerns only these individual threads and does not oblige group and societal risks to be taken into consideration.[872]

While Art. 25 obliges with the state of the art, which refers to the engineering side of PbD, to be taken into consideration, it is argued that the engineering side is as unclear as the legal side of what steps to be taken in order to ensure the rights of data subjects are protected from harms caused by data processing. This is mostly due to the differentiating definitions of values that are underlying terms like

[872] Stephan Dreyer and Wolfgang Schulz, "The General Data Protection Regulation and Automated Decision-Making: Will It Deliver?: Potentials and Limitations in Ensuring the Rights and Freedoms of Individuals, Groups and Society as a Whole," Discussion Paper Ethics of Algorithms #5 (2019).

privacy[873] and equality in the case of debiasing with fairness metrics.[874]

3.1.4.3 Data Protection Impact Assessment (DPIA) (Art. 35)

Impact assessments are common in many regulatory areas. Technology assessments were used in assessing impacts of new technologies in the 1960s. A common example from the regulatory area is environmental impact assessments.[875] Privacy Impact Assessments (PIA) first emerged in Anglo-Saxon countries in 1990s and have become common practice for many companies in managing their data protection responsibilities.[876] Environmental Impact Statements (EIS) under US law, voluntary Environmental Impact Assessments (EIA), Human Rights Impact Assessments (HRIA) drafted by the United Nations as an integrated part of the UN Guiding Principles on Business and Human Rights (UNGPs), can also be given as examples to its predecessors.[877]

Data Protection Impact Assessment is introduced in the GDPR Article 35 as an accountability measure. In order to assess the impact of an envisioned processing activity on the personal data, especially when new technologies are being used,

---

[873] Jörg Pohle, "Privacy and Data Protection by Design: A Critical Perspective," in Privacy and Cyber Security on the Books and on the Ground, ed. Ingolf Pernice and Jörg Pohle (Berlin, Germany: Alexander von Humboldt Institute for Internet and Society, 2019). For an effort for bridging such gap see, Kobbi Nissim et al., "Bridging the Gap between Computer Science and Legal Approaches to Privacy," *Harvard Journal of Law & Technology* 31 (2018).

[874] Wachter, Mittelstadt, and Russell, "Bias Preservation in Machine Learning: The Legality of Fairness Metrics under Eu Non-Discrimination Law."

[875] Eleni Kosta, "Article 35. Data Protection Impact Assessment," in *The Eu General Data Protection Regulation (Gdpr): A Commentary*, ed. Christopher Kuner, Lee A. Bygrave, and Christopher Docksey (New York: Oxford University Press, 2020).

[876] Niels van Dijk, Raphaël Gellert, and Kjetil Rommetveit, "A Risk to a Right? Beyond Data Protection Risk Assessments," *Comput. Law Secur. Rev.* 32 (2016): 287.

[877] Margot E Kaminski and Gianclaudio Malgieri, "Algorithmic Impact Assessments under the Gdpr: Producing Multi-Layered Explanations," *International Data Privacy Law* 11, no. 2 (2020).

data controllers are obliged to carry out a DPIA before they start processing any personal data, only if this envisioned processing is likely to result in a high risk to the rights and freedoms of natural persons. According to Recital 84, the outcome of the output of the DPIA is to be considered in determining the appropriate measures in order to demonstrate compliance of the processing activity.

In article 35(3), there is an inexhaustive list of examples that are likely to result in high risk; (i) ADM, including profiling, with legal or similarly significant effect to the natural person (ii) Processing large scale of special categories of data or data relating to criminal convictions, (iii) systemic monitoring of a publicly accessible area on a large scale. Then WP29 added nine criteria, by taking into account Article 35(1), 35(3)(a)(b)(c), 35(4) and recitals 71,75 and 91, to be considered when deciding whether a DPIA is required:[878] (i) evaluation or scoring including performance at work, economic situation, reliability or behaviour, (ii) automated decision making with legal or similar significant effect, (iii) systematic monitoring, (iv) systemic monitoring of public areas, (v) sensitive data or data of a highly personal nature, which is broader than special category data as in Articles 9 and 10, including electronic communications and financial data, (vi) data processed on a large scale, which can be determined on the basis number of data subjects as a specific number or a proportion of the given population, the volume and/or the range of data items, duration or permanence, geographical extent (vii) data concerning vulnerable data subjects such as children, employees, mentally ill persons, asylum seekers, patients and the elderly,[879] (viii) innovative use, application of new technological or organizational solutions, (ix) processing which prevents data subjects from exercising a right, using a service or a entering

[878] Of course the fact that WP29/EDPB opinions and recitals are not binding has to be kept in mind.

[879] Here WP29 gives an inexhaustive list by taking power imbalance into account, see Malgieri and González Fuster, "The Vulnerable Data Subject: A Gendered Data Subject?."; Malgieri and Niklas, "Vulnerable Data Subjects."

into a contractual relationship.[880] According to WP29, if the processing activity includes two of the criteria a DPIA is required, however controllers may decide one criteria to be enough to carry out a DPIA.

The first crucial contribution of DPIA to the toolset of data protection law to address algorithmic discrimination is that unlike article 22(1) it is not limited to solely automated decisions, thus it includes a consideration of automation bias in assessing a predictive tool that assists human decision making. Carrying out a DPIA is nevertheless mandatory in Article 22(1) situations. WP29 suggests that it is useful to conduct a DPIA when a controller is not sure whether their processing activity falls within the scope of Art. 22(1), whether an exception is applicable and what safeguards should be implemented.[881] DPIA is suggested by many scholars as a much more effective alternative to Art. 22 in terms of its potential to address algorithmic discrimination.[882] Furthermore, it was suggested that DPIA enhances the Article 22 where it applies, especially in providing the documentation needed

---

[880] Article 29 Working Party, "Guidelines on Data Protection Impact Assessment (Dpia) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679," WP 248 (2017).

[881] "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679".

[882] Bryce Goodman, "A Step Towards Accountable Algorithms ? : Algorithmic Discrimination and the European Union General Data Protection," (2016), https://www.semanticscholar.org/paper/A-Step-Towards-Accountable-Algorithms-%3A-Algorithmic-Goodman/f0b77ffb3d751c7d62dbe003b3e9e441ca10c629; Dreyer and Schulz, "The General Data Protection Regulation and Automated Decision-Making: Will It Deliver?: Potentials and Limitations in Ensuring the Rights and Freedoms of Individuals, Groups and Society as a Whole"; Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."; Kaminski and Malgieri, "Algorithmic Impact Assessments under the Gdpr: Producing Multi-Layered Explanations."; Ivanova, "The Data Protection Impact Assessment as a Tool to Enforce Non-Discriminatory Ai."; Laurens Naudts, "How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them " in *Data Protection and Privacy: The Internet of Bodies*, ed. R. Leenes, et al. (Oxford: Hart Publishing, 2019).

to meet algorithmic transparency requirements stemming from Articles 13, 14, 15 and 22.[883]

An interesting wording is found in Article 35(1) that reads "rights and freedoms of the natural person," which at first glance gives the impression of a consideration of a broader risk than relating to the data subjects affected by the processing. Reading it together with Article 35(9), according to which "the controller shall seek the views of data subjects or their representatives on the intended processing," Naudts suggests that it means the risks to the rights of those whose data have not been processed also has to be measured in a DPIA. On this account, he suggests that DPIA captures the risk of exclusion or underrepsented data sets. On the other hand, reading it with "…risks to the rights and freedoms of data subjects referred to in paragraph 1…." in Article 35 (7), Kosta opines that the obligatory DPIA is limited to situations where the rights and freedoms of those whose data have been processed is affected. In our opinion, it is worth more considering the wording in Article 35(7)(d) which requires rights and legitimate interests of "data subjects and other persons concerned" to be taken into account among envisaged measures. The answer to the question whose rights are to be considered has decisive power in measuring and managing the risk, as well as whether the effect on natural persons whose data have not been or will not be processed by the processing activity under assessment is legally required to be considered in DPIA or not is crucial. If not its effectiveness as an *ex ante* tool for the negative societal effects of algorithmic discrimination on the WMG is diminished. We will come back to DPIA's function as to algorithmic transparency and challenging algorithmic discrimination again, but first what is the risk that has to be assessed in a DPIA?

[883] Kaminski and Malgieri, "Algorithmic Impact Assessments under the Gdpr: Producing Multi-Layered Explanations."

3.1.4.3.1 A risk and a right

WP29 defines risk as "a scenario describing an event and its consequences, estimated in terms of severity and likelihood"[884] as it is already known in various risk management contexts such as information security, environmental law, finance and so on. This is exactly the reason that it is crucial to understand what "a risk to a right" mean, as determining a risk may call for different expertise and methodologies than determining a right. Van Dijk et al provide a thought provoking list of risk-right relations as follows:[885] (i) risk vs right, (ii) right as risk, (iii) risk to a right and (iv) right at risk.

"Risk vs right" appears in governmental and legal contexts. First is exemplified by the authors around security concerns fuelled after 9/11 leading to a trade-off between national security risk and civil liberties, while the latter is exemplified by the proportionality test of the ECtHR which seeks to strike a fair balance between individual rights and public interest. Another risk assessment conducted by the ECtHR falls under the category of *"risk to a right,"* which is developed from *Taskın*[886] and on and requires technical expertise in the form of an impact assessment in order to produce legal evidence so that it can be determined whether a risky practice violated an individual right.

"Right as risk," on the other hand is used in business context within risk governance schemes where the risk is to the business entity in the form of reputational damage and loss of trust. Data protection experts are often advised to bring these concerns forth to achieve initial c-level buy-in. Here the nature of the

---

[884] Article 29 Working Party, "Guidelines on Data Protection Impact Assessment (Dpia) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679". 6.

[885] Dijk, Gellert, and Rommetveit, "A Risk to a Right? Beyond Data Protection Risk Assessments," 290-96.

[886] *Taşkın and Others v. Turkey,* App. No. 46117/99 (2 February 2010).

rights are not the real concern, the real concern is the likelihood of the consequences that may be borne by the entity in case of a violation. It takes a quantitative approach to risk, which relies on breaking down data protection concepts into tasks and targets and making use of technical and organisational expertise more than legal. This is the common trend in approaching DPIAs, which are taken as a part of "broader" information security framework in which they represent the privacy risk. This can be observed in WP29 guidelines whereby some tasks addressed to DPOs in the GDPR were mentioned to be possibly carried out by Chief Information Security Officers (CISO) as well.[887]

Finally, *"right at risk"* is how communities and data subjects experience the risk to a right, they feel under threat. "Right vs risk" and "right as risk" logics leave the actual rightsholders and their experiences mostly out. Van Dijck et. al emphasize that not articulating the risk from the perspective of those who experience the threat would lead to rights losing their meaning, while realization of their rights claims still calls for being operationalized in terms of technology, politics and law.

Future is by definition uncertain, yet risk assessment aims at creating a level of certainty for future events based on statistics and probabilities. Here Van Dijck et. al indicate to a deterministic problem risk-based approach share with those very technologies which are mandatory to be a part of a DPIA. It is the probabilistic nature of them that is designed to better control the future which may end up creating one for individuals and groups based on its own anticipations.[888] This knowledge about the future is derived from what is known today based on the past

---

[887] Article 29 Working Party, "Guidelines on Data Protection Impact Assessment (Dpia) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679"; Kosta, "Article 35. Data Protection Impact Assessment," 672.

[888] Lessig, *Code 2.0*, 306-07; Hildebrandt, "Profiling and the Identity of the European Citizen," 306-07.

data and how it is interpreted by various human actors in the development of a project. How the definition of a target variable such as a good employee is defined by the developers will define the output of a model and for whom it poses a risk of exclusion or disparate treatment. Similarly, how the unwanted future events are defined as risk criteria and how a list of risks created (risk identification) is going to determine what is going to be taken into account and what is going to be left out. Thus, Van Dijck et. al suggest that legal knowledge derived from case law which has a long history dealing with scenarios where those feared future events had occurred should not be left out.[889] This approach may also help defining the risks in a broader societal manner in DPIAs, than it is possible within information security frameworks.

An attempt of bridging the gap between risk regulation and data protection law came from von Grafenstein who looks into the other side of the coin and discusses in detail why a proactive risk-based approach is a better fit for data protection law than a reactive harm-based approach, while also clarifying the ambiguities about object and subject of protection under data protection law by taking purpose limitation principle to the centre in a series of three articles.[890]

First of all, he clarifies the terminology in the sense that risk may be specific or unspecific, determined based on the knowledge about the causality between an event and harm to a specific object of protection; where there is insufficient knowledge as to such causality than it is unspecified harm. If there is insufficient knowledge about both the harm and the likelihood than it is danger not risk (in sociology, and it has been used for the exact opposite concepts in some

[889] Dijk, Gellert, and Rommetveit, "A Risk to a Right? Beyond Data Protection Risk Assessments."
[890] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part I."; "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Ii."; "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Iii."

jurisdictions such as in Germany).[891] Second, von Grafenstein opines that the problems addressed by Dijk et. al are actually easy to solve as both the idea that data protection law provides protection against risks of data processing and how to weigh quantitative measures against the qualitative risk to a right were addressed long before GDPR. First, the object of protection provided by the right to data protection in Article 8 CFREU is individuals' autonomy which is put at risk by processing of their personal data, specifically individuals' ability to autonomously exercise all of their other fundamental rights and freedoms. This protection covers both the unspecified and specified risks. In this regard, unspecified risk to autonomy is considered given, once the personal data is collected. So that the protection starts before this happens as it can never be surely known that data did not turn into information in someone's head and once it does it cannot be taken back. Once the data is collected and the individual is identified, then if there is sufficient proof as to a causal link between a harm and a fundamental right and freedom within the specific context in which the processing happens and the specific way data is used, then the normative substance of the other fundamental rights takes the centre of specific risk assessment.

At this point the ambiguity remains as it is still a difficult task to translate the substance of these rights into measurable risk with quantitative methods without losing touch with their substance. According to von Grafenstein this is ensured by purpose limitation principle. So the causal link between a right and harm is defined by the purpose of processing which remains to be specified by the data controller correctly as a means to legalize their processing activity. On this account, all the other fundamental rights are becoming objects of protection in determining the risk to a right only to the extent the risk is caused by the informational power asymmetry created by the processing. If there is no risk which arose from the informational power asymmetry, then instead of data protection, the protection should be provided by the scope of this other fundamental right.

---

[891] "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part I."

We are in the opinion, in ADM, especially when machine learning is involved informational power asymmetry is evident. Thus, von Grafenstein's approach makes it possible to claim that the risk of the relevant ADM process to all the other fundamental rights of individuals concerned should be a part of risk assessment under DPIA as a result of GDPR Article 1(2).[892] Then the abovementioned discussion about whose rights are in the scope of a DPIA is also answered. In reading Article 1(2) with the definition of data subject, any identified or identifiable natural person whose autonomous exercise of their fundamental rights and freedoms are affected by the informational power asymmetry caused by the processing activity, should be considered. As long as training phase do not include properly anonymized data, which is difficult to achieve as GDPR sets a very high standard,[893] processing taking place at this stage does not preclude a consideration of the risk to all the fundamental rights including right to equality (Art. 20 CFREU) and non-discrimination (Art. 21 CFREU) of natural persons.

According to von Grafenstein, the purpose of the processing set correctly by the controller is the means to express their intention to reach a future event. In this regard discriminatory intent would immediately bring in right to non-discrimination. However, as discussed before, most likely this is not going to be the case. Let's take a scenario from the screening stage of the hiring funnel. The employer wishes to automate their screening stage so that they can weed out those who do not meet required qualifications and top performer traits. In the development stage of this project, they use the data of their current and past employees with the purpose of discovering top performer traits with the most predictive power, in order to be then used in screening new applicants. It would be unrealistic to think that without any further goal, the employer just wishes to discover top performer traits through a machine learning project for the sake of

[892] "This Regulation protects fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data."

[893] C-582/14 [2016] *Patrick Breyer v. Bundesrepublic Deutschland* ECLI:EU:C:2016:779

research or statistics. At this point, perhaps the purpose of using the model also in automating promotion or pay decisions is not yet intended. So, the processing during the training phase will not directly affect current and past employees whose data is going to be processed but there is a likelihood as the outcome may have predictive power on their job performance. So even at the training stage both a specific and an unspecific risk to autonomous exercise of the right to non-discrimination can be captured.

Thus, we could say that von Grafenstein's model facilitates substantiating the risk to a right question. However, the problem remains with the right at question being an individual one, it is questionable whether it brings in any further consideration than what is already there in terms of EU antidiscrimination law, which suffers from many limitations as discussed in the previous chapter in detail. As this solution takes the fundamental rights protection in the EU in its centre, it is enclosed by the limits of protection provided thereby. Thus it may fall short in incorporating "right at risk" concerns.

3.1.4.3.2 Discrimination as risk

According to Ivanova, such interpretation of the data protection law in the light of the non-discrimination principle enshrined in the CFREU Art. 21 could widen the existing scope of equality protection beyond the protected ground and domain based limitations in the Equality Directives. She bases her opinion on the CJEU's case law that "EU data protection legislation must be interpreted and applied in the light of the fundamental rights enshrined in the Charter."[894] This is true that the CJEU has interpreted the DPD in light of the CFREU since the Lisbon Treaty

[894] Ivanova, "The Data Protection Impact Assessment as a Tool to Enforce Non-Discriminatory Ai," 3.

entered into force.[895] However considering that how reluctant CJEU has been applying CFREU Article 21 in its own scope in broadening the protection offered by the EU antidiscrimination law, it is doubtful it would do so in the field of data protection.

As Recital 71 and 75, as well as WP29 guidance on ADM, profiling and DPIA specifically mention discrimination as a high risk against which controllers should take appropriate technical and organizational measures and demonstrate doing so as part of the accountability principle, in addition to the examples to high risk examples given in Article 35 and nine factors listed by WP29 to be considered in weighing the risk to a right, most ADM processes and algorithmic discrimination risk resulting from them are to be considered when conducting a DPIA.

How can a DPIA enhance the protection provided by the EU data protection law and antidiscrimination law, and fill its gaps?

3.1.4.3.2.1 Transparency

First of all, it is widely accepted that DPIA regime enhances the transparency requirements of the GDPR,[896] thus it may fill some of the gaps of the EU antidiscrimination law as we have identified as lack of awareness and problems in

---

[895] See C-362/14 [2015] *Maximillian Schrems v Data Protection Commissioner* ECLI:EU:C:2015:650; C-293/12 and and C-594/12 Joined Cases [2014] *Digital Rights Ireland Ltd* ECLI:EU:C:2014:238

[896] Kaminski and Malgieri, "Algorithmic Impact Assessments under the Gdpr: Producing Multi-Layered Explanations."; Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law."; Emre Bayamlıoğlu, "The Right to Contest Automated Decisions under the General Data Protection Regulation : Beyond the So-Called "Right to Explanation"," *Regulation & Governance* (2021); Brkan, "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the Gdpr and Beyond."

access to evidence in the previous chapter. However, it is also widely mentioned as a problem that DPIAs are not mandatorily made available to the public[897] which remains only a recommendation of the WP29 who considers publication of DPIA as a part of overall transparency and accountability principles.[898]

Article 35(7) requires DPIAs to contain systematic description of the envisioned processing operation and the purposes, as well as the legitimate interest pursued, assessment of the necessity and proportionality of the processing, assessment of risks to the rights of data subjects, measures, safeguards, security measures to demonstrate compliance with the GDPR by taking into account legitimate interests of data subjects and other persons concerned.

According to Casey et. al, Article 35(1) is strikingly similar to the formulation of Article 22(1) and taken together these provisions have a preventive synergy which prevails throughout the GDPR as an understanding of data protection by design.[899] Kaminski et. al are also in the opinion that transparency rights in the GDPR should be understood as a system, which provides multi-layered explanations, which offer *ex ante* and *ex post* intervention opportunities, instead of being examined separately.[900] It is true that statistical nature of algorithmic

---

[897] Dariusz Kloza et al., *Data Protection Impact Assessments in the European Union: Complementing the New Legal Framework Towards a More Robust Protection of Individuals* (2017); Reuben Binns, "Data Protection Impact Assessments: A Meta-Regulatory Approach," *International Data Privacy Law* 7, no. 1 (2017); Alessandro Mantelero, "Ai and Big Data: A Blueprint for a Human Rights, Social and Ethical Impact Assessment," *Computer Law & Security Review* 34, no. 4 (2018).

[898] Article 29 Working Party, "Guidelines on Data Protection Impact Assessment (Dpia) and Determining Whether Processing Is “Likely to Result in a High Risk” for the Purposes of Regulation 2016/679".

[899] Casey, Farhangi, and Vogl, "Rethinking Explainable Machines: The Gdpr's 'Right to Explanation' Debate and the Rise of Algorithmic Audits in Enterprise," 179.

[900] Kaminski and Malgieri, "Algorithmic Impact Assessments under the Gdpr: Producing Multi-Layered Explanations."

discrimination cannot be sufficiently captured by individual rights set out in the GDPR, not only because the scope and nature of these rights are not clear as discussed above and the limitations regarding the individual redress in the previous chapter is also relevant for them, but also use of such rights do not cover demanding design that would mitigate unfair practices and lead to transparency to remain an illusion.[901]

According to Kaminski et. al, individual rights regarding algorithmic systems become a part of more systemic governance through the DPIA regime. According to the authors, DPIAs may provide the source material for individual data subjects exercising their algorithmic due process rights which are laid down in articles 13, 14 and 15 of the GDPR, as the meaningful explanation about the logic involved, significance and the envisaged consequences of the ADM processes are already a part of the necessary DPIA content. According to authors, the group based discriminatory effects which are discovered during DPIA should be included in the information provided to individuals who exercise their algorithmic due process rights.[902] If this approach proves true in practice, in a way the discovery of the DPIA includes group based explanations as to a discriminatory effect to a level of providing sufficient information to render the automated decision contestable[903] to be the only way for such right to be meaningful, then this explanation received by the data subject may involve sufficient information to establish prima facie discrimination before equality bodies or the courts of a member state.

Moreover as it was suggested by Wachter and Mittelstadt, beyond explanations, automated decisions require justifications which involve an explanation of why

---

[901] Edwards and Veale, "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For."

[902] Kaminski and Malgieri, "Algorithmic Impact Assessments under the Gdpr: Producing Multi-Layered Explanations," 132.

[903] Bayamlıoğlu, "The Right to Contest Automated Decisions under the General Data Protection Regulation : Beyond the So-Called “Right to Explanation”."

the controller think it is a correct, lawful and fair decision.[904] Article 35(7)(d) on this account, mandates an assessment of the necessity and proportionality of the processing, the result of which would produce such justification. Which would also be beneficiary for the controller as necessity and proportionality is also required for objective justification of indirect discrimination under EU antidiscrimination law.

Furthermore, pursuant to the prior consultation regime introduced in Article 36 GDPR, when a DPIA results in high risk unless it can be mitigated by the controller, the controller is obliged to consult the supervisory authority (SA). In case the SA comes to the conclusion that the envisaged processing activity infringes the GDPR, especially if the risk is insufficiently identified or mitigated, then the SA may give advice to the controller, alongside with using its powers which include asking for any information, conducting audits, allege infringement, access to all personal data and all the other necessary information, access to any premises of the controller within its investigative powers set out in Article 58(1). Even though the capacity of SAs to identify algorithmic discrimination raises some doubts,[905] they have all the necessary powers to have access to the relevant information and to prevent it from effecting individuals by making use of their corrective powers set out in Article 58(2) including imposing fines and a ban on processing, which we call an *ex ante* kill switch. SA also has the power to bring the infringement before judicial authorities. These activities are included in the annual report of SAs pursuant to Article 59, which may help with raising public awareness. On the other hand, SAs' powers over the DPIA regime is quite sufficient to raise corporate awareness for making the documentation, risk assessment and monitoring of their processing activities correctly.

---

[904] Wachter and Mittelstadt, "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and Ai."

[905] Kaminski and Malgieri, "Algorithmic Impact Assessments under the Gdpr: Producing Multi-Layered Explanations," 133.

Controllers are also obliged to seek the views of data subjects, however only when they find it appropriate. WP29 suggests that the views can be obtained through generic studies, by sending a survey to future customers or a question to employee representatives. It is also suggested that controllers to document why they disagreed with the data subjects' opinion and ground the path they continued, as well as justifying the reasons why they did not sought their opinion in the first place.[906] This can be easily justified as the article 35(9) also provides for data controllers to be able to put forward commercial or public interests, or the security of their operations. The original Commission version of the text did not involve these escape roots, however was added as the European Parliament held the position that this might have burdened the data controllers disproportionately.[907] If they remained mandatory without being watered down, it could have been a very important step in democratizing ADM processes, as well as raising both corporate and public awareness. However, we believe that for companies who wish to legalize their processes, including stakeholder opinions in their DPIAs may still be an option not only to strengthen demonstration of their compliance but also to actually have fairer processes.[908]

3.1.4.3.2.2 Equal treatment by design

Hacker[909] and Ivanova[910] suggest that DPIAs can be utilized in enforcing equal

---

[906] Article 29 Working Party, "Guidelines on Data Protection Impact Assessment (Dpia) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679". 15.

[907] Binns, "Data Protection Impact Assessments: A Meta-Regulatory Approach," 18.

[908] Importance of stakeholder involvement especially on catching downstream effects is also emphasized by NIST in Reva et al., "Towards a Standard for Identifying and Managing Bias in Artificial Intelligence."

[909] Hacker, "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law," 25.

[910] Ivanova, "The Data Protection Impact Assessment as a Tool to Enforce Non-Discriminatory Ai."

treatment by design. Even though, we do not pertain to the logic suggested by Ivanova on anti-discrimination law becoming an obligatory part of DPIA regime with a broadened scope, we agree that an assessment of risk to non-discrimination right likely caused by ADM processes should be included within a DPIA and if the purpose of processing or the automated decision fall under the material scope of antidiscrimination law then it is very expected this regulatory risk to be a part of the assessment.

In this regard, Ivanova's paper involves a profound risk identification and risk assessment methodology for distinguished stages of machine learning and ADM based on direct and indirect discrimination doctrines. Importantly, she involves testing of models for direct and indirect discrimination, as well as documenting whether the practice is falling under one of the exceptions in the case of direct discrimination such as being a genuine occupational requirement, or whether it is objectively justified in case indirect discrimination is captured. Ivanova suggests such a scenario would be taken as medium risk by taking into account the Court may not accept the justification.

We are sharing the opinion that a DPIA provides a structured framework for an early assessment of compliance with the EU antidiscrimination law for controllers when their processing activity falls under its scope. As discussed in the previous Chapter though, this scope is quite limited. On the other hand, despite its reluctance so far, the CJEU is entitled to interpret laws and disputes regulated by the European Union law in the light of the CFREU, which has the same hierarchy in the EU law with the Treaties, and the general principles of law which also include the jurisprudence of the ECtHR among its sources. Thus, it would only be smart for organisations whose practices are likely to materialize discrimination, to involve the CJEU and ECtHR case law on equality and non-discrimination

principles[911] in their risk analysis.

One of the suggestions involved in Ivanova's framework is to include building diverse teams and training of developers for bias and discrimination awareness to be documented in the DPIA among measures taken to minimize the risk of discrimination,[912] which speaks to addressing the ADM as an assemblage consisting also of human practices and the way concepts are organized in these humans' heads being a part of the overall process.[913]

There has been a number of model suggestions on algorithmic processes and AI, such as Algorithmic Impact Statements,[914] Human Impact Statement in Algorithmic Decision-making,[915] model Algorithmic Impact Assessment (AIA) under GDPR[916] and Human Rights, Ethical and Social Impact Assessment (HRESIA). HRESIA model suggested by Mantelero brings the role of right to data protection as serving to the protection of all the other fundamental rights as elaborated by von Grafenstein as well, one step further to also include societal impacts. Mantelero suggests that as a way of providing stronger protection of individual rights and freedoms, within the DPIA framework controllers could voluntarily also consider collective social and ethical values.[917] Even though the model remains voluntary, provides a detailed framework including an *ad hoc* expert committee and a self-assessment tool in the form of a questionnaire.

---

[911] See *Handbook on European Non-Discrimination Law*. for examples from the case law and comparison of two Courts' approaches.

[912] Ivanova, "The Data Protection Impact Assessment as a Tool to Enforce Non-Discriminatory Ai."

[913] See. Section 1.2.3.

[914] Andrew D. Selbst, "Disparate Impact in Big Data Policing," *Georgia law review* 52 (2017).

[915] Katyal, "Private Accountability in the Age of Artificial Intelligence."

[916] Kaminski and Malgieri, "Algorithmic Impact Assessments under the Gdpr: Producing Multi-Layered Explanations."

[917] Mantelero, "Ai and Big Data: A Blueprint for a Human Rights, Social and Ethical Impact Assessment."; ibid.

CAHAI[918] addressed United Nations Guiding Principles on Business and Human Rights (UNGPs)[919] and acknowledged a holistic approach which considers all relevant civil, political, social, cultural and economic rights. The final outcome of CAHAI, "Possible elements of a legal framework on artificial intelligence, based on the Council of Europe's standards on human rights, democracy and the rule of law" suggests a risk-based approach and a Human Rights, Democracy and Rule of Law Impact Assessment (HUDERIA) as a part of a possible legally binding transversal instrument. Also it is worth noting here, gender mainstreaming toolkit of the European Institute for Gender Equality (EIGE) involves a Gender Impact Assessment (GIA) which is a basic method used by governmental structures at the EU and Member State level for gender mainstreaming,[920] this is not a risk based method and does not involve ADM processes, however it can be incorporated in a DPIA.

3.1.4.4 Other instruments in the accountability framework of the GDPR

Data Protection Officers (DPOs), record keeping duties (Art. 30), codes of conduct (Art. 40) and certification (Art. 42) mechanisms have been welcomed by many scholars in the context of addressing algorithmic discrimination.[921]

---

[918] CAHAI, "Ad Hoc Committee on Artificial Intelligence (Cahai): Feasibility Study".

[919] An important note from the CAHAI report is the emphasis on the applicability of international legal instruments to private actors. States have a duty to ensure human rights are respected by private actors by incorporating them in their national laws and providing effective remedies through judicial or non-judicial mechanisms. CAHAI also refers to the UN Guiding Principles on Business and Human Rights as respecting human rights across their operations, services and products is also a part of corporate responsibility schemes ibid..

[920] European Institute for Gender Equality, Gender Impact Assessment: Gender Mainstreaming Toolkit, (Luxembourg: Publications Office of the European Union, 2017), https://eige.europa.eu/gender-mainstreaming/toolkits/gender-impact-assessment.

[921] Dreyer and Schulz, "The General Data Protection Regulation and Automated Decision-Making: Will It Deliver?: Potentials and Limitations in Ensuring the Rights and Freedoms of Individuals,

GDPR consists of many overlapping layers of risk assessments.[922] According to von Grafenstein, being a part of these layers, codes of conduct and certification mechanisms in particular make such risk governance scale by reducing the legal uncertainty and business costs for controllers.[923] Naudts suggests that alongside DPIAs, codes of conduct shift the responsibility back to the data controller,[924] as in exercising algorithmic due process rights and managing privacy notices and their informed consent, data subject is burdened with protecting themselves within these complex assemblages. According to Dreyer and Schulz, controllers and the society at large may benefit from guidance of industry associations on supra-individual rights and freedoms. Moreover, co-regulatory notion of these mechanisms both benefit from a governmental framework and having the expertise and fast respond of the industries especially with respect to new technologies and changes in industry practices, in addition to possibilities of cooperation at the international level without being bound with national laws.[925] For example, if we take Ivanova's DPIA as an enforcement tool of antidiscrimination law,[926] some aspects of it might be too time consuming and costly for one controller, however within a framework of certified codes of conduct more controllers may be willing to commit to incorporating de-biasing, fairness, and antidiscrimination assessments.

Groups and Society as a Whole"; Naudts, "How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them ".

[922] von Grafenstein, "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Iii."

[923] Ibid.

[924] Naudts, "How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them ".

[925] Dreyer and Schulz, "The General Data Protection Regulation and Automated Decision-Making: Will It Deliver?: Potentials and Limitations in Ensuring the Rights and Freedoms of Individuals, Groups and Society as a Whole".

[926] Ivanova, "The Data Protection Impact Assessment as a Tool to Enforce Non-Discriminatory Ai."

Furthermore, certifications can be set on the basis of sectors which is crucial in capturing varying sources and other domain based particularities of discriminatory practices.[927] Codes of conducts and certification mechanism may also achieve more effective stakeholder participation in comparison to DPIAs. According to Naudts, such consultations should include civil society groups such as consumer protection organisations and equality bodies, thus the larger societal effects of machine learning technologies can be taken into account. We would add workers' organisations and trade unions to that for the employment contexts.

Rightfully noted, after the fall of Safe Harbour[928] that was validated with TrustE,[929] co-regulatory efforts as certified codes of conduct and privacy seals may not build much public trust.[930]

Lastly, data protection officers (DPO) are considered a part of the risk-based approach. They take an advisory role in the performance of the DPIA. According to WP29, CISOs may also take this role instead,[931] which would bring forth a more information security approach which may not be sufficient to capture fundamental rights perspective and underlying principles of the data protection law. Controllers may also resort to independent experts outside their organisation, which may bring in expertise in fairness and bias applications, as well as audits.

---

[927] Edwards and Veale, "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For," 80.

[928] C-362/14 [2015] *Maximillian Schrems v Data Protection Commissioner* ECLI:EU:C:2015:650

[929] TrustE continues to verify Privacy Shield which has also lost trust. See C-311/18 [2020] *Data Protection Commissioner v Facebook Ireland Limited and Maximillian Schrems* ECLI:EU:C:2020:559.

[930] Edwards and Veale, "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For."

[931] Article 29 Working Party, "Guidelines on Data Protection Impact Assessment (Dpia) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679".

Finally, outside the GDPR's body of instruments but a crucial supportive instrument is the Whistle-blower Directive (Directive 2019/1937)[932] which includes breaches falling within the scope of the GDPR concerning "protection of privacy and personal data, and security of network and information systems" in its material scope (Art.2(1)(a)(x)). Even if the reporting or public disclosure of information on breaches includes trade secrets, it is considered lawful under the Trade Secrets Directive (Directive 2016/943) provided that acquisition, use or disclosure of the trade secret conforms with the Article 3(1) of the Directive. Whistle-blower schemes have been considered important in achieving accountability of private companies.[933] A report by European Parliamentary Research Service on the governance of algorithmic accountability and transparency pointed at the importance of the role investigative journalism and whistle-blowers play in uncovering questionable uses and outcomes of ADM.[934] The same report suggested similar protection against reveal of trade secrets for investigative journalists who conduct reverse engineering to uncover truth about algorithms.[935] There is a current proposal for a Directive which would cover protection of investigative journalists but reverse engineering ADM processes doesn't seem to be in its scope of protection.[936] However, in its explanatory memorandum, the Whistleblower Directive is cited and it is acknowledged that "whistleblowers are often an important source for investigative journalism." It is not clear whether or which kind of algorithmic discrimination would be considered within the definition of breach under Whistleblower Directive, however such

[932] Directive (EU) 2019/1937 Of The European Parliament and of the Council of 23 October 2019 on the protection of persons who report breaches of Union law.

[933] Katyal, "Private Accountability in the Age of Artificial Intelligence."

[934] Ansgar Koene et al., *A Governance Framework for Algorithmic Accountability and Transparency* (Brussels, European Union: European Parliamentary Research Service, 2019), 69.

[935] Ibid., 70.

[936] Proposal for a Directive of the European Parliament and of the Council on protecting persons who engage in public participation from manifestly unfounded or abusive court proceedings ("Strategic lawsuits against public participation") COM(2022) 177 final. Brussels, 27.4.2022.

mechanisms may powerfully enhance accountability and protection concerning algorithmic discrimination, especially as regards to providing an evidence base for establishing prima facie discrimination in antidiscrimination cases.

3.1.5 Mandate to not-for-profit organization (Art 80)

As we have discussed, data protection is a part of individual human rights protection framework and does not really address social harms such as algorithmic discrimination of WMG at the group level. Art. 80 GDPR in this regard appears as an option similar to class action that is common in the US.[937]

In a recent case between Meta (Facebook) and the Federation of German Consumer Organisations,[938] the CJEU confirmed that consumer organisations have a right to take legal action to enforce GDPR under Art. 80. In combination with the Representative Actions Directive,[939] which has to be transposed by Member States until 25 December 2022 and includes data protection in its scope, it is expected to close some gaps that stem from such individual rights based nature of data protection law.

---

[937] Edwards and Veale, "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For."

[938] C-319/20 [2022] *Meta (Facebook) v. Verbraucherzentrale Bunderverband (vzbv) Meta (Facebook) V. Verbraucherzentrale Bundesverband (Vzbv)* ECLI:EU:C:2022:322.

[939] Directive EU 2020/1828 on representative actions for the protection of the collective interests of consumers, 25 November 2020

3.2 FORTHCOMING

The proposed and agreed Digital Services Act and Artificial Intelligence Act bring extra protection in the area of algorithmic hiring and job advertisement with discriminatory effects, especially where the use of personal data, therefore the protection provided by the GDPR becomes vague, or where GDPR does not allow for evidence to be gathered for an antidiscrimination case, we find some provisions of these forthcoming regulations can be meaningfully enhancing. Both regulations also have been subject to public debate, suggestions and criticism on major shortcomings. We do not aim at covering the legislative process and public debate about these regulations, neither do we aim at investigating their possible effects deeply, the following sections are included in this study for a brief introduction, for the further research is definitely needed that we leave to another study.

3.2.1 Digital Services Act (DSA)

The Digital Services Act (DSA) is welcomed not only in Europe but also in all around the World as a break system to the unleashed powers of online platforms. While content moderation has an effect on how the users engage with platforms, traces of which build their profiles which constitutes the basis they are delivered ads; the scope of such analysis is so wide and outside the scope of this study. However, DSA regulates online advertising on platforms and is also concerned with algorithmic discrimination.

Article 24 of the DSA regulates online advertising transparency and requires online platforms to provide meaningful information to users in real time about the main parameters used to determine to display each ad they are seeing to them in a clear and unambiguous manner.

Article 30 brings additional online advertising transparency to very large online platforms (VLOP), which includes the display period; whether the advertisement

was targeted to particular groups, if so the main parameters used for that purpose; aggregate numbers for the group or groups of recipients who were specifically targeted, alongside with the total number of users the advertisement have reached. This information shall be compiled and made publicly available by the VLOP in a repository that can be accessed with an API until one year after the last time the advertisement was displayed without any personal data related to who was or could have been targeted. Even if, this is an *ex post* measure, that would still not prevent from those who are excluded from a job advertisement to miss life chances, it may help bringing evidence to courts in order to establish prima facie direct or indirect discrimination especially in line with *Feryn*[940] and *Asociaţia Accept*.[941]

However, aggregate numbers may not reflect realities and may not be as granular as the choices given to advertisers. In this regard, even though the group or groups of recipients is not defined in the DSA, the only mention as to the subjects of a group is in the Explanatory Memorandum stating that "specific groups or persons may be vulnerable or disadvantaged in their use of online services because of their gender, race or ethnic origin, religion or belief, disability, age or sexual orientation." This is in line with the single axis protection provided by the EU antidiscrimination law and it resembles a closed system as it does not include the openness in Article 14 ECHR or Article 21 CFREU. On this account, it is unlikely that the aggregate information published by the VLOP will give away enough information to identify intersectional discrimination, proxy discrimination or discrimination based on sexual orientation. For instance, it has been known that Facebook allows for audience choices based on sexual orientation. It may have inferred such data but also it has been allowing 56 different categories of gender choice for user profiles so the data can be provided by the user as well. However, deep down in its system Facebook continues to classify users according to the

[940] C-54/07 [2008] *Feryn* ECLI:EU:C:2008:397

[941] C-81/12 [2013] *Asociaţia Accept* ECLI:EU:C:2013:275

gender binary female or male in order to target them with advertisement.[942] Thus the aggregate data in that sense also will most likely reflect this, making it hard to find hints for discrimination based on sexual orientation.[943]

One of the most important obligations brought on very large online platforms is the mandatory risk assessment under, which includes "any negative effects for the exercise of the fundamental rights to respect for private and family life, freedom of expression and information, the prohibition of discrimination" in Art. 26(1)(b). Under the second paragraph of the said article its determined that this risk assessment to include their "systems for selecting and displaying advertisement" if they influence risks listed in the first paragraph. In this regard, we are in the opinion that the job ad delivery that discriminates against WMG would be captured by this provision.

Article 29 DSA targets recommender systems and mandates VLOP to provide information in their terms and conditions on the parameters used in recommender services as well as options to modify or influence these parameters, VLOP shall also allow users to select and modify these parameters easily on the interface of the VLOP, thus taking control of the order of information presented to them by recommender systems.

More transparency into how platform advertising works is provided Art. 31 Very large online platforms shall provide the Digital Services Coordinator (DSC) access to data that are necessary to monitor and assess compliance with the DSA. Moreover, DSC may also request from VLOPs to provide access to data to vetted researchers for the sole purpose of conducting research that contributes to the identification and understanding of systemic risks such as negative effects of

---

[942] Rena Bivens, "The Gender Binary Will Not Be Deprogrammed: Ten Years of Coding Gender on Facebook," *New Media & Society* 19, no. 6 (2017).

[943] Griffin, "Tackling Discrimination in Targeted Advertising".

online advertisement to the right of non-discrimination. The definition of vetted researcher has been widely criticized as it leaves out researchers working at NGOs such as in AlgorithmWatch who have been conducting outstanding work in revealing algorithmic bias.[944]

Article 36 envisions transparency in online advertising to be furthered beyond the abovementioned provisions through Union level codes of conduct that are to be encouraged and facilitated by the European Commission, with the contribution of organisations representing the recipients of online advertising services as well as civil society organisations.

Another measure that is brought under Article 37 is the crisis protocols which require the freedom of expression and information and the right to non-discrimination to be safeguarded.

All these measures bring us to the recent settlement concerning the alleged housing discrimination ad delivery and signed between Facebook/Meta and Department of Housing and Urban Development in the USA, where Facebook/Meta agreed to develop a "Variance Reduction System" through which it would make targeted audiences more demographically representative of 'eligible audiences,' meaning that it would correct the disproportionality its ad delivery score calculation caused in the audience for sex and estimated race/ethnicity.[945] DSA could lead to such adjustments in job ads delivery by platforms in Europe as well. However, problems remain as the law do not define

[944] Holding platforms accountable: The DSA must empower vetted public interest research to reign in platform risks to the public sphere, AlgorithmWatch, <https://algorithmwatch.org/en/dsa-open-letter-november-2021/> Accessed on 15 September 2022.

[945] Settlement between Facebook/Meta and Department of Housing and Urban Development (On 21 June 2022) <https://about.fb.com/wp-content/uploads/2022/06/June_2022_HUD_Settlement_Agreement.pdf> Accessed on 25 June 2022.

non-discrimination in intersectional terms and given how granular the audience choices can be on these platforms proxy discrimination is very likely to either slip or remain subject to single-axis corrections.[946]

3.2.2 Artificial Intelligence Act (AIA)

In this final section of our legal analysis, we will juxtapose the proposed AIA with risks, limitations and shortcomings we discovered in the previous chapters in order to find out whether this forthcoming regulation fills the gaps.

3.2.2.1 Limitations

While a risk-based approach is welcomed, narrow scope of prohibited AI practices and the limitation of the most provisions in AIA to be exclusively applicable to high risk AI systems only, given the limited number of application areas listed in Annex III is problematic.

Manipulative systems are prohibited, however it requires intent, is limited only to particular vulnerabilities, and only if they cause physical or psychological harm to an individual. However, such harm does not always occur after one event but accumulates, and in the case of WMG they accumulate over years at a collective level rather than individuals. Moreover, the knowledge gained or correlations discovered through manipulative systems on vulnerable or disadvantaged individuals that can classify them emotionally and psychologically over time without causing physical or psychological harm to any individual, thus without being prohibited, may be used by a downstream actor,[947] for example in political targeting during elections, convincing them not to vote such as the strategy used

[946] Griffin, "Tackling Discrimination in Targeted Advertising".

[947] Michael Veale and Frederik Borgesius, *Demystifying the Draft Eu Artificial Intelligence Act* (2021).

during 2016 US elections. Such collective harms are unfortunately not considered.

Art. 5(1)(c) prohibits social scoring but only for the public sector applications, which is not understandable as Veale and Borgesius put it, the private sector applications of social scoring is much wider than public sector both in quantity and quality that includes basic infrastructure.[948]

Nevertheless, Annex III (4) captures our lead examples as employment, workers management and access to self-employment is within its scope. Annex III (4)(a) notably applies to the whole hiring funnel from sourcing (recruitment, selection, job ads), to screening and interviewing (filtering applications, evaluating candidates in the course of interviews or tests). However, does not really mention the selection stage which can be derived from the use of recruitment separately in addition to other actions taken. Annex III (4)(b) captures the platform workers as explicitly mentioned in Recital 36 as well as not fully automated systems used in HR management such decisions on promotion and termination, task allocation, monitoring, performance and behaviour evaluation. This will allow us to better evaluate which gaps the AIA is likely to fill, however, it must be noted that disadvantage within a society is interlinked with many other aspects. For example, systems regarding housing is not among the high risk systems, advertisement of goods and services, price discrimination remain outside, despite how much they may limit life chances of individuals and groups in other areas of their lives.

While perpetuation of historical pattern of discrimination mentioned numerous times in the AIA, there is no mention of intersectional discrimination or *ad hoc* groups. Thus it doesn't broaden the protection provided by the CFREU and anti-discrimination law already.

[948] Ibid.

3.2.2.2 Data

Article 10 of the AIA introduces quality standards for training, validation and testing data sets on the basis of which a high risk AI system is developed. The deployment and use stage is excluded as well as users.

Article 10 (3) mandates these datasets to be relevant, representative, free of errors and complete with added statistical quality expectations. This has been welcomed as it aims at correcting one of the biggest problems regarding the WMG as they are either over or underrepresented in datasets, as well as misrepresented as they are prone to have much more erroneous data.

Article 10(5) tackles another major problem that we have touched throughout this study that it frees bias monitoring, detection and correction from the constraints of special category personal data protection under GDPR. This way it reconciles scientific communities' efforts of debiasing both with requirement for ADMs under GDPR Article 22 derived from Recital 71 to prevent from discriminatory effects on natural persons and the need for testing against indirect discrimination under the EU antidiscrimination law. However, it is only applicable for the users of high risk AI systems, thus it not only excludes developers but also data brokers who might provide such datasets.

Hacker points out the disproportionality between the vague language used in the article and the severity of the fine up to 6 percent of global annual turnover that is applicable to failure in compliance. Another important point is that the accuracy and robustness of the output of the high risk AI systems are regulated under Article 15, which also attaches a high fine to non-compliance.[949] While quality of training and test data is crucial, the separation of these stages reminds

[949] Hacker, "A Legal Framework for Ai Training Data—from First Principles to the Artificial Intelligence Act."

of Lehr and Ohm's criticism of the legal scholars in not focusing enough on the development stage of the model where some of the bias problems can be overcome.[950]

3.2.2.3 Enforcement, oversight and redress

Despite introduction of notified bodies which may affix certifications upon conformity assessments, this remains voluntary as at large all high risk AI systems providers can also self-assess for conformity.[951]

Market surveillance authorities are responsible for oversight and enforcement based on notifications they receive from users and developers of AI systems. Individuals affected by AI systems are not given an option for bringing a complaint or case under the AIA act, neither do social interest groups, thus it is left to users and developers of AI systems to bring complaints against regulators on the right application of the Regulation.[952] In this regard AIA act falls behind the GDPR and the EU antidiscrimination law in its connection with fundamental rights protection, which we have criticized for its reliance on individual redress mechanisms.

Finally, Veale and Borgesius brought up the issue with the aim of AI Act which is to harmonize the laws apply to use of AI systems. They explain that as a result of AI Act taking direct effect, it may lead the CJEU to question national laws which offer extra protection for not being in line with the harmonization of the Union Law from a free movement of goods and services perspective, in other words AI Act leaves it open to interpretation that fragmented laws may need to

---

[950] Lehr and Ohm, "Playing with the Data: What Legal Scholars Should Learn About Machine Learning ".

[951] Veale and Borgesius, *Demystifying the Draft Eu Artificial Intelligence Act*.

[952] Ibid.

disapply for disrupting free trade of AI systems within the EU.

3.2.2.4 Transparency and participation

A database managed by the European Commission for registration of standalone high risk AI systems, which is proposed to be made public for other people. This may enable civil society and journalist to uncover illegal and unfair AI systems and at the minimum to raise awareness.

Malgieri and González Fuster[953] suggest that Florencia Luna's layered vulnerabilities approach[954] match well with the AIA as it refers to people with vulnerabilities. However, we are in the opinion that gender is not intended to be included as a vulnerability layer in the AIA. It seems to be referred rather within the discrimination context in line with prohibition of discrimination under the CFEU and TEU, as well as the secondary legislation of the EU such as antidiscrimination laws. On the other hand vulnerability is considered through age, especially for children (5(2)(2) of the Explanatory Memorandum, Recitals 16, 28, Art. 5(1)(b), 7(2)(f)), disabilities, physical or mental incapacities (Recital 16). Moreover, natural persons applying for or receiving public assistance benefits and services from public authorities (Recital 37), seeking for migration and asylum, as well as passing through border control (Recital 39) also considered to be possibly in a vulnerable position. Finally, vulnerable position in relation to the user of an AI system is acknowledged, in particular due to an imbalance of power, knowledge, economic or social circumstances, or age (Art. 7(2)(f)).

Another important point is that the AIA takes into consideration both decisional vulnerability and vulnerability in the outcomes. However, the second seems to be understood more under antidiscrimination legislation.

---

[953] Malgieri and González Fuster, "The Vulnerable Data Subject: A Gendered Data Subject?."

[954] Luna, "Elucidating the Concept of Vulnerability: Layers Not Labels."

In Art. 69 it is stated that the Commission will encourage voluntary applications related to environmental sustainability, accessibility for persons with a disability, stakeholders participation in the design and development of the AI systems and diversity of development teams in codes of conducts. Also involvement of interested stakeholders and their representative organisations in drawing up codes of conducts is envisioned.

CONCLUSION

Emancipation of WMG within the context of ADM requires both privacy and non-discrimination related frameworks to incorporate discrimination as an information infused harm. We have taken e-recruitment as an exemplary context as employment is the domain which provides the broadest protection under the EU antidiscrimination law, as well as being acknowledged as a context where ADMs would be considered to have legal or similarly significant effects on data subjects, moreover for being listed among the high-risk AI systems in the proposed AI Act. We have chosen a lead example where all the related legislations are providing higher protection, to establish even then there are many questions left for finding ways to effectively address negative effects of automation. Moreover, it has been suggested automated diversity and inclusion programs would remove human bias from hiring processes. However, there are ongoing streams of examples to the opposite of this promise which led us to ask whether algorithmic systems could actually act as emancipatory tools or they are not fit to bring societal change but to amplify what is there.

We have first connected the value attributed to personal data to the broader contexts and the history of oppressive practices such as colonialism, exclusion and segregation in employment, negative stereotypes and their still ongoing effects on WMG.

Automation of hiring processes is an interesting one also because it requires many negative decisions to be made. A bank will extend a loan to anyone who meets the requirements as there is not a set amount of money that they will need to distribute among applicants. A hiring decision for one open spot with many applications though, will inevitably involve many rejections. Thus, e-recruitment should not be understood as one ADM but as a series of decisions which make use of different technologies. In this regard different stages of an ADM process had to be analysed

contextually in order to understand why bias keep creeping into these systems and cause discriminatory harms.

We also gained insight to job ad delivery based on online behavioural advertisement (OBA) which enabled us to connect e-recruitment to the risks of online profiling activities as well as surveillance capitalism. We argued that WMG may have a stronger interest in landing a job and they might change their expressive and associative behaviour at online public spheres which would also cause further harms to democracy and human rights.

In this regard, in Chapter I we first identified the similarities of law and code that from a feminist critical point of view, they are both powerful tools in bringing change and perpetuating the status quo.

Under the international human rights frameworks, we identified the problem of all the related values that underline most important rights and freedoms in democratic societies to be contested depending on the situatedness of the right owner in the society. This is a big problem for automation as well as applying the right safeguards for the risks automation bring by. We suggest taking the standpoint of the most vulnerable individuals in the society when assessing the required safeguards. In this regard, we connected how right to non-discrimination and equality for the WMG is dependent on both their right to privacy and democratic freedoms of expression and association to be ensured. As well as how right to data protection provides *ex ante* protection of these rights as it applies before their own scope does.

Even though both data protection law and antidiscrimination law suffer from similar shortcomings to address algorithmic discrimination as they are based on individual rights protection while the harms ADM causes are rather group based/collective and societal, we agreed to those who argue that data protection law would enhance the reach of antidiscrimination law in the algorithmic

discrimination context. Especially in terms of *ex ante* assessment that is required in the form of a DPIA and transparency requirements even if it is interpreted rigidly to reject a right to explanation, may help overcoming only *ex post* protection provided by the antidiscrimination law and the limited options to collect evidence to establish prima facie discrimination before a court.

We find it positive that forthcoming regulations specifically address algorithmic discrimination, even though many loopholes remain to offer WMG proper protection which would lead to an emancipated future from the baggage of the past discrimination.

It appeared to us in conducting this study, time and time again the most disadvantaged and their experiences were not making it to the main concepts of the law. The subject of protection of the applicable law that we have analysed is not the fluent, intersectional, historically and structurally disadvantaged persons and their needs. It is rather envisioned to be a single axis, static, emancipated person who enjoys their individual rights without a collective background and structural disadvantages. All of the legal framework is designed for this hypothetical person whose rights may sometimes be violated, they would challenge this situation, receive a correction and the life would go back to normal. However, the subjects we sought protection for were diverse, fluent, intersectional, they were experiencing cumulative burdens on a daily basis in most of the contexts. This is the exact reality algorithmic systems are exacerbating and thus unearthing. We understand that the EU antidiscrimination law takes discrimination as a rare event limited to some identities and situations. The AI Act follows a similar root with a broader understanding, but again with a limited scope of application and maybe even more paternalistic as it has very limited options for participation and no option for subjects of protection to activate a due process. The GDPR, despite its flaws of being limited to personal data and overreliance on a hypothetical informed, rational person, who is emancipated from the constraints of time (as they are expected to read and understand all privacy notices they

encounter) and has unlimited cognitive capacity (considering the amount of prior information needed to grasp the individual and societal effects of their consent or objection to new and complex technologies), it still provides a very early intervention chance to individuals and their representatives, acknowledge them to seek redress as groups and also lightens the burden on their shoulders by keeping the controllers responsible for the evaluation and prevention of risks *ex ante*.

An intersectional feminist future demands an understanding of the real human experiences of discrimination though. It is to acknowledge that discrimination is not experienced as a single event on a single identity or character trait once in a lifetime. By treating it this way, both law and technology thus far fall short in addressing the reality of the disadvantaged. A future feminist technology law should recognize these collective and cumulative experiences which require moving on from single axis/single event individual answers to discrimination to maybe seeing them as assemblages or vulnerability layers. There needs to be more mandatory and meaningful stakeholder and domain expert participation in ethical boards and corporate processes, especially in design and *ex ante* prevention mechanisms such as DPIAs and more room for addressing collective, non-individual, continuous, contextual and everchanging experiences of discrimination.

We have found that stakeholder involvement was crucial for the voices of the subjects of protection to be heard. However, further questions await. Who are these stakeholders, how much do they represent the WMG and among the WMG the most marginalized of them? We have presented before how men being more powerful in trade unions led to women's exclusion laws under the name of protection or to the segregation in employment before. Can it be the same again, that because power is not equally distributed also in the civil society, more stakeholder involvement would not mean those at the bottom of the Crenshaw's basement will have a voice? How are these issues to be tackled against algorithmic systems that are globally applicable? How can we reconcile the interests of these

multicultural and multilingual communities that may be adversely affected? The EU regulatory processes which are directed to protecting a multinational and multilingual population may constitute a role model.

However, big companies as stakeholders have already proven to have a powerful voice in these processes in making sure involving the other stakeholders are left to their hands. As much as *ex ante* protective measures are necessary to cope with the pace and scale of ADM, two important problems remain to doubt how meaningful such protection will be. First, currently too much is left to the interpretation of the private companies about the scope of adverse societal effects to be considered. Second, it is difficult to expect from them to consider discrimination as a broader risk than the harms defined and protection provided by the law, which currently falls short.

As an example, when it comes to fairness to be sought by employers through technology, we discussed how some fairness metrics would cause positive action in the sense that CJEU considers reverse discrimination. It is because the Court still applies formal equality based on merits and does not consider those at the bottom layers of Crenshaw's basement may never have a chance to have the same merits with those from the upstairs. This does not mean they would not prove to be a good asset to a company through their work. To be more concrete, for a disabled woman from a poor household with an immigration background, the odds she will get to a top university is very little. She may learn to code at home, she may develop amazing problem solving skills throughout her difficult life. A company who could have benefited from her experience may never know it, because even when they want to give a chance by having a quota for those who are not qualified on the paper but from a certain group to be tried on the job, it may give those who are from the upstairs to bring a reverse discrimination case. This is thoughtful in the sense that the most advanced laws are still blind to structural disadvantage in a way they may even stagnate markets against correcting it by themselves.

This was highlighted as regards the AI Act. Not only it lacks meaningful stakeholder involvement and *ex post* due process for those who are negatively affected to a point of appearing paternalistic, it also may lead to an interpretation that national laws which offer extra protection may need to disapply due to a clash with its aim for EU-wide harmonisation. This also shows that laws which encompass everything may miss nuances and those nuances become very important when it comes to providing protection against discrimination.

Finally, we conclude that as it always has been law, even the most developed in the right direction, is not the best fit to provide emancipation. Neither is the technology. However, the current European law establishes mechanisms that can be furthered with enhanced participation, critical questioning, strategic litigation and continuous dialogue with the lawmakers and tech community. We see here also a window of opportunity, that the algorithms act like archaeologists, digging into the datafied history and previously not datafied deeply embedded binary codes of our societies that are gendered, racial, able-bodied, excavating all the ways showing that our societies are unfair. This is reminiscent of the industrial revolution, a time technological development changed the whole fabric of our societies. It is crucial that masses of users, workers, citizens, women and other marginalized groups, peoples of the south, all the people with layers of vulnerabilities work together to intervene now to change the focus and the language we approach to these new challenges we discussed in this study. There is a cracked window to redefine what we understand from democracy, equality, rights and freedoms. The feminist, anti-colonialist, anti-capitalist and other critical theories developed by the disadvantaged, theorized all the routes that lead to discrimination and imagined alternative futures for decades must not be overlooked in finding remedies.

REFERENCES

Aggarwal, Nikita. "The Norms of Algorithmic Credit Scoring." (2020).

Ali, Muhammad, Piotr Sapiezynski, Miranda Bogen, Aleksandra Korolova, Alan Mislove, and Aaron Rieke. "Discrimination through Optimization: How Facebook's Ad Delivery Can Lead to Biased Outcomes." *Proceedings of the ACM on Human-Computer Interaction* 3 (11/07 2019): 1-30.

Allhutter, Doris, Florian Cech, Fabian Fischer, Gabriel Grill, and Astrid Mager. "Algorithmic Profiling of Job Seekers in Austria: How Austerity Politics Are Made Effective." [In English]. *Frontiers in Big Data* 3 (2020-February-21 2020).

Alpaydın, Ethem. *Machine Learning, Revised and Updated Edition.* Cambridge, Massachusetts: The MIT Press, 2022.

Altman, Andrew. "Legal Realism, Critical Legal Studies, and Dworkin." *Philosophy and Public Affairs* 15, no. 3 (1986): 205-35.

Amnesty International. "Toxic Twitter – Women's Experiences of Violence and Abuse on Twitter." https://www.amnesty.org/en/latest/news/2018/03/online-violence-against-women-chapter-3/.

———. *Xenophobic Machines: Discrimination through Unregulated Use of Algorithms in the Dutch Childcare Benefits Scandal*. London: Amnesty International, 2021. https://www.amnesty.org/en/documents/eur35/4686/2021/en/.

Ananny, Mike. "Toward an Ethics of Algorithms:Convening, Observation, Probability, and Timeliness." *Science, Technology, & Human Values* 41, no. 1 (2016): 93-117.

Ananny, Mike, and Kate Crawford. "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability." *New Media & Society* 20, no. 3 (2018): 973-89.

Angwin, Julia, and Terry Parris Jr. "Facebook Lets Advertisers Exclude Users by Race." https://www.propublica.org/article/facebook-lets-advertisers-exclude-users-by-race.

Angwin, Julia, Ariana Tobin, and Madeleine Varner. "Facebook (Still) Letting Housing Advertisers Exclude Users by Race." https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin.

Article 29 Working Party. "Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679." WP 251(2018). https://ec.europa.eu/newsroom/article29/items/612053

———. "Guidelines on Consent under Regulation 2016/679." WP259(2018).

———. "Guidelines on Data Protection Impact Assessment (Dpia) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679." WP 248(2017).

———. "Guidelines on the Right to Data Portability." WP 242(2017).

———. "Opinion 2/2010 on Online Behavioural Advertising ". WP 171 (2010).

———. "Opinion 03/2013 on Purpose Limitation." WP 203, no. 00569/13/EN (2013). https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf.

———. "Opinion 4/2007 on the Concept of Personal Data " WP 136(2007).

———. "Opinion 05/2014 on Anoymisation Techniques." WP 216(2014).

Asghari, Hadi, Nadine Birner, Aljoscha Burchardt, Daniela Dicks, Judith Fassbender, Nils Feldhus, Freya Hewett, *et al.* "What to Explain When Explaining Is Difficult? An Interdisciplinary Primer on Xai and Meaningful Information in Automated Decision-Making." *Alexander von Humboldt Institute for Internet and Society* (2021). doi:https://doi.org/10.5281/zenodo.6375784.

Atabey, Ayça. *Is Google at Odds with the Gdpr? Evaluation of Google's Personal Data Collection on Mobile Operating Systems in Light of the Principles of Purpose Limitation, Data Minimisation, and Accountability.* On İki Levha, 2020.

Banterle, Francesco. "Data Ownership in the Data Economy: A European Dilemma." 199-225, 2018.

Barlow, John Perry. "A Declaration of the Independence of Cyberspace." (1996). https://www.eff.org/cyberspace-independence.

Barocas, Solon, and Andrew D. Selbst. "Big Data's Disparate Impact." *California Law Review* 104 (2016): 671-732.

Barth, Susanne, and Menno D. T. de Jong. "The Privacy Paradox – Investigating Discrepancies between Expressed Privacy Concerns and Actual Online Behavior – a Systematic Literature Review." *Telematics and Informatics* 34, no. 7 (2017/11/01/ 2017): 1038-58.

Bartlett, Katharine T. "Feminist Legal Methods." *Harvard Law Review* 103, no. 4 (February 1990 1990): 829-88.

Bayamlıoğlu, Emre. "The Right to Contest Automated Decisions under the General Data Protection Regulation : Beyond the So-Called "Right to Explanation"." *Regulation & Governance* (2021).

Beauchamp, Toby. "Artful Concealment and Strategic Visibility: Transgender Bodies and U.S. State Surveillance after 9/11." *Surveillance and Society* 6 (06/26 2009): 356-66.

Berlin, Isaiah Sir. "Two Concepts of Liberty." 1958.

Bieker, F. *The Right to Data Protection: Individual and Structural Dimensions of Data Protection in Eu Law*. T.M.C. Asser Press, 2022.

Binns, Reuben. "Data Protection Impact Assessments: A Meta-Regulatory Approach." *Information Privacy Law eJournal* (2016).

———. "Data Protection Impact Assessments: A Meta-Regulatory Approach." *International Data Privacy Law* 7, no. 1 (2017): 22-35.

———. "Fairness in Machine Learning: Lessons from Political Philosophy." *Decision-Making in Computational Design & Technology eJournal* (2018).

Binns, Reuben, Ulrik Lyngs, Max Van Kleek, Jun Zhao, Timothy Libert, and Nigel Shadbolt. "Third Party Tracking in the Mobile Ecosystem." In *Proceedings of the 10th ACM Conference on Web Science*, 23–31. Amsterdam, Netherlands: Association for Computing Machinery, 2018.

Birhane, Abeba, and Olivia Guest. "Towards Decolonising Computational Sciences." *Kvinder, Køn & Forskning* 29 (02/08 2021): 60-73.

Bivens, Rena. "The Gender Binary Will Not Be Deprogrammed: Ten Years of Coding Gender on Facebook." *New Media & Society* 19, no. 6 (2017): 880-98.

Bogen, Miranda, and Aaron Rieke. "Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias." 2018.

Böhm, Steffen, Ana C. Dinerstein, and André Spicer. "(Im)Possibilities of Autonomy: Social Movements in and Beyond Capital, the State and Development." *Social Movement Studies* 9, no. 1 (2010/01/01 2010): 17-32.

Bolukbasi, Tolga, Kai-Wei Chang, James Y. Zou, Venkatesh Saligrama, and Adam Tauman Kalai. "Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings." Paper presented at the NIPS, 2016.

boyd, danah, and Kate Crawford. "Six Provocations for Big Data." *Computer (Long. Beach. Calif)* 123 (09/21 2011).

Bradley, Harriet, Mark Erickson, Carol Stephenson, and Stephen Williams. *Myths at Work.* Polity, 2000.

Brkan, Maja. "Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the Gdpr and Beyond." *International Journal of Law and Information Technology* 27, no. 2 (2019): 91-121.

Browne, Simone. "Dark Matters : On the Surveillance of Blackness." [In English]. (2015).

Büchi, Moritz, Noemi Festic, and Michael Latzer. "The Chilling Effects of Digital Dataveillance: A Theoretical Model and an Empirical Research Agenda." *Big Data & Society* 9, no. 1 (2022): 20539517211065368.

Bunting, Annie. "Feminism, Foucault, and Law as Power/Knowledge." *Alberta law review* 30 (1969): 829-29.

Buolamwini, Joy, and Timnit Gebru. "Gender Shades: Intersectional Accuracy Disparities in

Commercial Gender Classification." In *Conference on Fairness, Accountability, and Transparency*, edited by Sorelle A. Friedler and Christo Wilson, 1-15: Proceedings of Machine Learning Research, 2018.

Burkert, Herbert. "Information Law: From Discipline to Method." *Berkman Center Research Publication* No. 2014-5, U. of St. Gallen Law & Economics Working Paper No. 2014-02(2014). htpps://ssrn.com/abstract=2402866.

Burrell, Jenna. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 3, no. 1 (2016): 2053951715622512.

Bverfg, (6th November 2019).

Bygrave, Lee A. "Article 22 Automated Individual Decision-Making, Including Profiling." In *The Eu General Data Protection Regulation (Gdpr): A Commentary*, edited by Christopher Kuner, Lee A. Bygrave, Christopher Docksey and Laura Drechsler. New York: Oxford University Press, 2020.

———. "Automated Profiling: Minding the Machine: Article 15 of the Ec Data Protection Directive and Automated Profiling." *Computer Law & Security Review* 17, no. 1 (2001/01/01/ 2001): 17-24.

CAHAI. "Ad Hoc Committee on Artificial Intelligence (Cahai): Feasibility Study." *Council of Europe* CAHAI(2020)23(2020). https://rm.coe.int/cahai-2020-23-final-eng-feasibility-study-/1680a0c6da.

Caliskan, Aylin, Joanna J. Bryson, and Arvind Narayanan. "Semantics Derived Automatically from Language Corpora Contain Human-Like Biases." *Science* 356, no. 6334 (2017): 183-86.

Canhoto, Ana, and James Backhouse. "General Description of the Process of Behavioural Profiling." In *Profiling the European Citizen: Cross-Disciplinary Perspectives*, edited by Mireille Hildebrandt and Serge Gutwirth: Springer, 2008.

Casey, Bryan James, Ashkon Farhangi, and Roland Vogl. "Rethinking Explainable Machines: The Gdpr's 'Right to Explanation' Debate and the Rise of Algorithmic Audits in Enterprise." *Berkeley Technology Law Journal* 34 (2018): 143.

Castets-Renard, Céline. "Human Rights and Algorithmic Impact Assessment for Predictive Policing." In *Constitutional Challenges in the Algorithmic Society*, edited by Amnon Reichman, Andrea Simoncini, Giovanni De Gregorio, Giovanni Sartor, Hans- W. Micklitz and Oreste Pollicino, 93-110. Cambridge: Cambridge University Press, 2021.

Ceccaroni, Federica. "Collective Discrimination without an Identifiable Victim in Eu

Law. Discrimination by Public Speech." *Forum di Quaderni Costituzionali* (2020): 843-69.

Charlesworth, Hilary, Christine Chinkin, and Shelley Wright. "Feminist Approaches to International Law." *American Journal of International Law* 85, no. 4 (1991): 613-45.

Chiusi, Fabio, Sarah Fischer, Nicolas Kayser-Bril, and Matthias Spielkamp, eds. *Automating Society Report 2020*. Germany: AlgorithmWatch gGmbH & Bertelsmann Stiftung, 2020.

Citron, Danielle Keats. "Technological Due Process." *WASH. U. L. REV.* 85, no. 6 (2007): 1248-313.

Citron, Danielle Keats, and Mary Anne Franks. "Criminalizing Revenge Porn ". *Wake Forest Law Review* 49 (2014): 345.

Cohen, Julie E. *Between Truth and Power: The Legal Constructions of Informational Capitalism.* New York: Oxford University Press, 2019.

Collins, Patricia Hill. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment.* Second ed. New York and London: Routledge, 2000.

Commission, European, Directorate-General for Justice, Consumers, J Gerards, and R Xenidis. *Algorithmic Discrimination in Europe : Challenges and Opportunities for Gender Equality and Non-Discrimination Law*. Publications Office, 2021. doi:doi/10.2838/544956.

A Comparative Study on Algorithmic Discrimination between Europe and Northamerica. Italian Equality Network, 2022. https://www.italianequalitynetwork.it/a-comparative-study-on-algorithmic-discrimination-between-europe-and-north-america?action=genpdf&id=1564.

Couldry, Nick, and Ulises A. Mejias. "Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject." *Television & New Media* 20, no. 4 (2019): 336-49.

Craig, Paul, and Gráinne de Búrca. *Eu Law : Text, Cases, and Materials*. 6. edition ed. Oxford: Oxford University Press, 2017.

Crenshaw, Kimberly Williams. "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics." *University of Chicago Legal Forum* 1989, no. 1 (1989).

Criado-Perez, Caroline. *Invisible Women : Data Bias in a World Designed for Men* London: Vintage, 2020.

Custers, Bart, Toon Calders, Bart Schermer, and Tal Zarsky. *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases.* Springer Publishing Company, Incorporated, 2013.

Custers, Bart, Toon Calders, Tal Zarsky, and Bart Schermer. "The Way Forward." In *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases*, edited by Bart Custers, Toon Calders, Bart Schermer and Tal Zarsky: Springer Publishing Company, Incorporated, 2013.

D'Ignazio, Catherine, and Lauren F. Klein. *Data Feminism.* Cambridge, Massachusetts: Massachusetts Institute of Technology, 2020.

Datta, Amit, Michael Carl Tschantz, and Anupam Datta. "Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice, and Discrimination." *ArXiv* abs/1408.6491 (2014).

Dewey, John. "Ethics." In *The Collected Works of John Dewey V. 5; 1908, Ethics : The Middle Works, 1899-1924*, edited by Jo Ann Boydston. Carbondale, United States: Southern Illinois University Press, 2008.

Dijck, José van. "Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology." *surveillance and society* 12 (2014): 197-208.

Dijk, Niels van, Raphaël Gellert, and Kjetil Rommetveit. "A Risk to a Right? Beyond Data Protection Risk Assessments." *Comput. Law Secur. Rev.* 32 (2016): 286-306.

Dinar, Christina. "The State of Content Moderation for the Lgbtiqa+ Community and the Role of the Eu Digital Services Act." (2021). https://eu.boell.org/sites/default/files/2021-06/HBS-e-paper-state-platform-moderation-for-LGBTQI-200621_FINAL.pdf.

Doward, Jamie, and Alice Gibbs. "Did Cambridge Analytica Influence the Brexit Vote and the Us Election?" *The Guardian*, 2017.

Dreyer, Stephan, and Wolfgang Schulz. "The General Data Protection Regulation and Automated Decision-Making: Will It Deliver?: Potentials and Limitations in Ensuring the Rights and Freedoms of Individuals, Groups and Society as a Whole." *Discussion Paper Ethics of Algorithms #5* (2019). doi:DOI 10.11586/2018018.

Duball, Joseph. "Changing Course: Google Maneuvers toward New User-Tracking Norms." (2022). https://iapp.org/news/a/changing-course-google-maneuvers-toward-new-user-tracking-norms/.

Dunn, John Dale. "Democracy: The Unfinished Journey, 508 Bc to Ad 1993." *Contemporary Sociology* 22 (1993): 680.

Dwork, Cynthia, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard S. Zemel. "Fairness through Awareness." *ArXiv* abs/1104.3913 (2012).

Eckersley, Peter. "How Unique Is Your Web Browser?", Berlin, Heidelberg, 2010.

EDPB Guidelines 8/2020. "On the Targeting of Social Media Users Version 2.0." 2021.

EDPS Opinion 3/2018. "European Data Protection Supervisor’s Opinion on Online Manipulation and Personal Data." 2018.

Edwards, Lilian, and Michael Veale. "Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For." *Duke Law & Technology Review* 16, no. 18 (2017): 18-84.

Ellis, Evelyn, and Philippa Watson. *Eu Anti-Discrimination Law.* Second ed. New York: Oxford University Press, 2012.

Elson, Diane. "Gender Justice, Human Rights, and Neo-Liberal Economic Policies." In *Gender Justice, Development, and Rights*, edited by Maxine Molyneux and Shahra Razavi, 0: Oxford University Press, 2002.

Equality, European Institute for Gender. *Gender Impact Assessment: Gender Mainstreaming Toolkit*. Luxembourg: Publications Office of the European Union, 2017. https://eige.europa.eu/gender-mainstreaming/toolkits/gender-impact-assessment.

Eren, Ezgi, Lukas Hondrich, Linus Huang, Basileal Imana, Matthias C. Kettemann, Joanne Kuai, Marcela Mattiuzzo, *et al.* "Increasing Fairness in Targeted Advertising: The Risk of Gender Stereotyping by Job Ad Algorithms.", edited by N. Birner, S. Hod, M. C. Kettemann, A. Pirang and F. Stock2021.

Eubanks, Virginia. "Automating Inequality : How High-Tech Tools Profile, Police, and Punish the Poor." [In English]. (2018).

"European Commission (2014) ". *Joint Report on the application of Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin ('Racial Equality Directive') and of Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation ('Employment Equality Directive')* COM(2014) 2 final (Brussels). https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52014DC0002&from=EN.

European, Commission, Justice Directorate-General for, Consumers, and S. Fredman. *Intersectional Discrimination in Eu Gender Equality and Non-Discrimination Law.* Publications Office, 2016. doi:doi/10.2838/241520.

European Union: European Commission. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Better Working Conditions for a Stronger Social Europe: Harnessing the Full Benefits of Digitalisation for the Future of Work, 9 December 2021, Com(2021) 761 Final* European Union: European Commission, 2021.

———. *Communication from the Commission to the European Parliament, the Council, the European Economic Tnd Social Committee and the Committee of the Regions - a Union of Equality: Gender Equality Strategy 2020-2025, 5 March 2020, Com(2020) 152 Final.* European Union: European Commission, 2020.

———. *Proposal for a Directive of the European Parliament and of the Council on Improving Working Conditions in Platform Work, 9 December 2021, Com(2021) 762 Final.* European Union: European Commission, 2021.

Fennessy, Caitlin, and Nicole Sakin. "How Google and Apple Are Shaking up Adtech." (2021). https://iapp.org/news/a/how-google-and-apple-are-shaking-up-adtech/.

Ferrer, X., T. v. Nuenen, J. M. Such, M. Coté, and N. Criado. "Bias and Discrimination in Ai: A Cross-Disciplinary Perspective." *IEEE Technology and Society Magazine* 40, no. 2 (2021): 72-80.

Ferretti, Federico. "Not-So-Big and Big Credit Data between Traditional Consumer Finance, Fintechs, and the Banking Union: Old and New Challenges in an Enduring Eu Policy and Legal Conundrum." *Global Jurist* 18, no. 1 (2018).

Fia, Tommaso. "An Alternative to Data Ownership: Managing Access to Non-Personal Data through the Commons." *Global Jurist* 21, no. 1 (2021): 181-210.

Finck, Michèle, and Asia J. Biega. "Reviving Purpose Limitation and Data Minimisation in Personalisation, Profiling and Decision-Making Systems." *ArXiv* abs/2101.06203 (2021).

Fineman, Martha. "The Vulnerable Subject: Anchoring Equality in the Human Condition." *Yale J. of Law & Feminism* 20 (05/10 2008).

Floridi, Luciano. "Open Data, Data Protection, and Group Privacy." *Philosophy & Technology* 27, no. 1 (2014/03/01 2014): 1-3.

Foucault, Michel. *Discipline and Punish: The Birth of the Prison.* London: Penguin Books, 1991.

———. *The History of Sexuality. Volume 1: An Introduction.* [in In English; translated from French.] New York: Pantheon Books, 1978.

Foulds, James, Rashidul Islam, Kamrun Naher Keya, and Shimei Pan. "An Intersectional Definition of Fairness." *Journal Name: IEEE International Conference on Data Engineering* (2020): Medium: X.

Fraser, Nancy. "On Justice." *New Left Review* MAR/APR 2012, no. 74 (2012).

Fredman, Sandra. *Women and the Law.* New York: Oxford University Press, 1997.

Gabbatt, Adam. "Facebook Charged with Housing Discrimination in Targeted Ads." https://www.theguardian.com/technology/2019/mar/28/facebook-ads-housing-discrimination-charges-us-government-hud.

Gallie, W. B. "Essentially Contested Concepts." *Proceedings of the Aristotelian Society* 56 (1955): 167-98.

Ganesh, Maya, Jeff Deutch, and Jennifer Schulte. *Privacy, Visibility, Anonymity: Dilemmas in Tech Use by Marginalised Communities.* 2016. doi:10.13140/RG.2.2.17371.36648.

Ganty, Sarah. "Poverty as Misrecognition: What Role for Antidiscrimination Law in Europe?". *Human Rights Law Review* 21, no. 4 (2021): 962-1007.

Gellert, Raphaël, Katja de Vries, Paul de Hert, and Serge Gutwirth. "A Comparative Analysis of Anti-Discrimination and Data Protection Legislations." In *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases*, edited by Bart Custers, Toon Calders, Bart Schermer and Tal Zarsky, 61-89. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.

Gerards, Janneke H., and Frederik J. Zuiderveen Borgesius. "Protected Grounds and the System of Non-Discrimination Law in the Context of Algorithmic Decision-Making and Artificial Intelligence." *EngRN: Electronic* (2020).

Gillespie, Tarleton, Pablo J. Boczkowski, and Kirsten A. Foot, eds. *Media Technologies: Essays on Communication, Materiality, and Society*. United States of America: The MIT Press, 2014.

Gitelman, Lisa, ed. *"Raw Data" Is an Oxymoron*. Cambridge, Massachusetts: The MIT Press, 2013.

Gollatz, Kirsten, Felix Beer, and Christian Katzenbach. "The Turn to Artificial Intelligence in Governing Communication Online." In *HIIG Workshop Report*: Alexander von Humboldt Institute for Internet and Society (HIIG), 2018.

Goodman, Bryce. "A Step Towards Accountable Algorithms ? : Algorithmic Discrimination and the European Union General Data Protection." 2016.

———. "A Step Towards Accountable Algorithms ? : Algorithmic Discrimination and the European Union General Data Protection." (2016). https://www.semanticscholar.org/paper/A-Step-Towards-Accountable-Algorithms-%3A-Algorithmic-Goodman/f0b77ffb3d751c7d62dbe003b3e9e441ca10c629.

Goodman, Bryce, and Seth Flaxman. "European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"." *AI Mag.* 38 (2017): 50-57.

Griffin, Rachel. "Tackling Discrimination in Targeted Advertising." Verfassungsblog, https://verfassungsblog.de/targeted-ad/.

Gutwirth, Serge, and Paul De Hert. "Regulating Profiling in a Democratic Constitutional State." In *Profiling the European Citizen: Cross-Disciplinary Perspectives*, edited by Mireille Hildebrandt and Serge Gutwirth: Springer, 2008.

Hacker, Philipp. "A Legal Framework for Ai Training Data—from First Principles to the Artificial Intelligence Act." *Law, Innovation and Technology* 13, no. 2 (2021/07/03 2021): 257-301.

———. "Teaching Fairness to Artificial Intelligence: Existing and Novel Strategies against Algorithmic Discrimination under Eu Law." *Common Market Law Review* 55 (April 18, 2018 2018).

Haggerty, Kevin D., and Richard V. Ericson. "The Surveillant Assemblage." *The British journal of sociology* 51 4 (2000): 605-22.

Hallinan, Dara. "Opinions · Data Protection without Data: Could Data Protection Law Apply without Personal Data Being Processed?". *European Data Protection Law Review* 5, no. 3 (2019).

Handbook on European Non-Discrimination Law. 2018 Edition. Luxembourg: Publications Office of the European Union, 2018.

Hao, Karen. "We Read the Paper That Forced Timnit Gebru out of Google. Here's What It Says." MIT Technology Review, https://www.technologyreview.com/2020/12/04/1013294/google-ai-ethics-research-paper-forced-out-timnit-gebru/.

Haraway, Donna. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14, no. 3 (Autumn, 1988 1988).

Harding, Sandra. "Rethinking Standpoint Epistemology: What Is "Strong Objectivity"?". In *Feminist Epistemologies (Thinking Gender)*, edited by E. Potter L. Alcott. New York: Routledge, 1993.

Hardt, Moritz, Eric Price, and Nathan Srebro. "Equality of Opportunity in Supervised Learning." *ArXiv* abs/1610.02413 (2016).

Hartmann, Heidi. "Capitalism, Patriarchy, and Job Segregation by Sex." *Signs* 1, no. 3 (1976): 137-69.

Helberger, Natali. "Profiling and Targeting Consumers in the Internet of Things – a New Challenge for Consumer Law." 2016.

Hijmans, Hielke. "Article 1 Subject-Matter and Objectives." In *The Eu General Data Protection Regulation (Gdpr): A Commentary*, edited by Christopher

Kuner, Lee A. Bygrave and Christopher Docksey. New York: Oxford University Press, 2020.

Hildebrandt, Mireille. "Defining Profiling: A New Type of Knowledge?". In *Profiling the European Citizen: Cross-Disciplinary Perspectives*, edited by Mireille Hildebrandt and Serge Gutwirth, 17-45. Dordrecht: Springer Netherlands, 2008.

———. "Profiling and the Identity of the European Citizen." In *Profiling the European Citizen: Cross-Disciplinary Perspectives*, edited by Mireille Hildebrandt and Serge Gutwirth. Dordrecht: Springer Netherlands, 2008.

———. "Profiling and the Rule of Law." *Identity in the Information Society* 1, no. 1 (2008/12/01 2008): 55-70.

Hoffmann, Anna Lauren. "Where Fairness Fails: Data, Algorithms, and the Limits of Antidiscrimination Discourse." *Information, Communication & Society* 22 (2019): 900 - 15.

Hofmann, Jeanette, Christian Katzenbach, and Kirsten Gollatz. "Between Coordination and Regulation: Finding the Governance in Internet Governance." *New Media & Society* 19 (2017): 1406 - 23.

Hughes, Melanie. "Diversity in National Legislatures around the World." *Sociology Compass* 7 (01/01 2013).

Hummel, Patrik, Matthias Braun, and Peter Dabrock. "Own Data? Ethical Reflections on Data Ownership." *Philosophy & Technology* 34, no. 3 (2021/09/01 2021): 545-72.

Hurley, Mikella, and Julius Adebayo. "Credit Scoring in the Era of Big Data." *Yale J.L. & Tech* 18, no. 1 (2017).

Information Commissionar's Office (ICO). "Big Data, Artificial Intelligence, Machine Learning and Data Protection." 20170301 Version: 2.0(2017). https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf.

Ivanova, Yordanka. "The Data Protection Impact Assessment as a Tool to Enforce Non-Discriminatory Ai." *SSRN Electronic Journal* (01/01 2020).

Janeček, Václav. "Ownership of Personal Data in the Internet of Things." *Computer Law & Security Review* 34, no. 5 (2018/10/01/ 2018): 1039-52.

Jarovsky, Luiza. "Dark Patterns in Personal Data Collection: Definition, Taxonomy and Lawfulness." (2022). doi:http://dx.doi.org/10.2139/ssrn.4048582, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4048582.

Joseph, Matthew, Michael Kearns, Jamie H. Morgenstern, Seth Neel, and Aaron Roth. "Rawlsian Fairness for Machine Learning." *ArXiv* abs/1610.09559 (2016).

Kadiri, Aisha P.L. "Data and Afrofuturism: An Emancipated Subject?". *Internet Policy Review* 10, no. 4 (2021).

Kaminski, Margot E, and Gianclaudio Malgieri. "Algorithmic Impact Assessments under the Gdpr: Producing Multi-Layered Explanations." *International Data Privacy Law* 11, no. 2 (2020): 125-44.

Kaminski, Margot E. "The Right to Explanation, Explained." *Berkeley Technology Law Journal* 34 (2018): 189.

Kappelhof, Joost. "Survey Research and the Quality of Survey Data among Ethnic Minorities." 2017.

Katyal, Sonia K. "Private Accountability in an Age of Artificial Intelligence." *The Cambridge Handbook of the Law of Algorithms* (2020).

———. "Private Accountability in the Age of Artificial Intelligence." *UCLA L. REV.* 66 (2019).

Kennedy, Duncan. *A Critique of Adjudication.* Harvard University Press, 1997.

Keyes, Os. "The Misgendering Machines: Trans/Hci Implications of Automatic Gender Recognition." Paper presented at the The ACM on Human-Computer Interaction, 2018.

———. "Os Keyes on Avoiding Universalism and 'Silver Bullets' in Tech Design." By Kerry Mackereth and Eleanor Drage. *The Good Robot Podcast* (2022).

Khan, Shmyla. "Surveillance as a Feminist Issue." Privacy International, 2017.

Kim, Pauline. "Data-Driven Discrimination at Work." *William & Mary Law Review* 48 (2017 2017): 857-936

Kim, Pauline T., and Sharion Scott. "Discrimination in Online Employment Recruiting." *Saint Louis University law journal* 63 (2018): 7.

Kitchin, Rob, and Gavin McArdle. "What Makes Big Data, Big Data? Exploring the Ontological Characteristics of 26 Datasets." *Big Data & Society* 3, no. 1 (2016): 2053951716631130.

Klimas, Tadas, and Jurate Vaiciukaite. "The Law of Recitals in European Community Legislation." *ILSA Journal of Int'l & Comparative Law* 15, no. 1 (2008).

https://www.researchgate.net/publication/228152770_The_Law_of_Recitals_in_European_Community_Legislation.

Kloza, Dariusz, Niels Dijk, Raphaël Gellert, István Böröcz, Alessia Tanas, Eugenio Mantovani, and Paul Quinn. *Data Protection Impact Assessments in the European Union: Complementing the New Legal Framework Towards a More Robust Protection of Individuals.* 2017.

Köchling, Alina, and Marius Claus Wehner. "Discriminated by an Algorithm: A Systematic Review of Discrimination and Fairness by Algorithmic Decision-Making in the Context of Hr Recruitment and Hr Development." *Business Research* 13, no. 3 (2020/11/01 2020): 795-848.

Koene, A., L. Dowthwaite, and S. Seth. "Ieee P7003tm Standard for Algorithmic Bias Considerations." Paper presented at the 2018 IEEE/ACM International Workshop on Software Fairness (FairWare), 29-29 May 2018 2018.

Koene, Ansgar, Chris Clifton, Yohko Hatada, Helena Webb, and Rashida Richardson. *A Governance Framework for Algorithmic Accountability and Transparency.* Brussels, European Union: European Parliamentary Research Service, 2019.

Kokott, Juliane, and Christoph Sobotta. "The Distinction between Privacy and Data Protection in the Jurisprudence of the Cjeu and the Ecthr." *International Data Privacy Law* 3, no. 4 (2013): 222-28.

Koldinská, Kristina. "Eu Non-Discrimination Law and Policies in Reaction to Intersectional Discrimination against Roma Women in Central and Eastern Europe." In *European Union Non-Discrimination Law and Intersectionality : Investigating the Triangle of Racial, Gender and Disability Discrimination*, edited by Anna Lawson and Dagmar Schiek: Burlington, 2011.

Koops, Bert-Jaap, and Ronald E. Leenes. "Privacy Regulation Cannot Be Hardcoded. A Critical Comment on the 'Privacy by Design' Provision in Data-Protection Law." *International Review of Law, Computers & Technology* 28 (2014): 159 - 71.

Korff, Douwe. "New Challenges to Data Protection Study - Working Paper No. 2: Data Protection Laws in the Eu: The Difficulties in Meeting the Challenges Posed by Global Social and Technical Developments." *Information Privacy Law eJournal* (2010).

Kosta, Eleni. "Article 35. Data Protection Impact Assessment." In *The Eu General Data Protection Regulation (Gdpr): A Commentary*, edited by Christopher Kuner, Lee A. Bygrave and Christopher Docksey. New York: Oxford University Press, 2020.

Kreimer, Margareta. "Labour Market Segregation and the Gender-Based Division of Labour." *European Journal of Women's Studies* 11, no. 2 (2004): 223-46.

Kroll, Joshua, Joanna Huey, Solon Barocas, Edward Felten, Joel Reidenberg, David Robinson, and Harlan Yu. "Accountable Algorithms." *University of Pennsylvania Law Review* 165 (02/01 2017): 633-705.

Krys, Kuba, Karolina Hansen, Cai Xing, Piotr Szarota, and Miao-miao Yang. "Do Only Fools Smile at Strangers? Cultural Differences in Social Perception of Intelligence of Smiling Individuals." *Journal of Cross-Cultural Psychology* 45, no. 2 (2014): 314-21.

Kuner, Christopher, Lee A. Bygrave, and Christopher Docksey, eds. *The Eu General Data Protection Regulation (Gdpr): A Commentary*. New York: Oxford University Press, 2020.

Lacey, Nicola. *Unspeakable Subjects: Feminist Essays in Legal and Social Theory.* [in en] 1 ed. London: Hart Publishing, 1998. doi:10.5040/9781472561916.

Lahuerta, Sara Benedi. "Enforcing Eu Equality Law through Collective Redress: Lagging Behind?". *Common Market Law Review* (2018): 783-817.

Lambrecht, Anja, and Catherine Tucker. "Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of Stem Career Ads." *SSRN Electronic Journal* (01/01 2016).

Langenkamp, Max, A. Costa, and Chris Cheung. "Hiring Fairly in the Age of Algorithms." *ArXiv* abs/2004.07132 (2020).

Lehr, David, and Paul Ohm. "Playing with the Data: What Legal Scholars Should Learn About Machine Learning ". *U.C. Davis Law Review* 51, no. 2 (2017): 653-717.

Lepri, Bruno, Nuria Oliver, Emmanuel Letouzé, Alex Pentland, and Patrick Vinck. "Fair, Transparent, and Accountable Algorithmic Decision-Making Processes." *Philosophy & Technology* 31, no. 4 (2018/12/01 2018): 611-27.

Leslie, David, Christopher Burr, Mhairi Aitken, Michael A. Katell, Morgan Briggs, and Camilo A. RamÌrez Rincon. "Human Rights, Democracy, and the Rule of Law Assurance Framework for Ai Systems: A Proposal." *ArXiv* abs/2202.02776 (2022).

Lessig, Lawrence. *Code 2.0.* CreateSpace, 2009.

Levy, Karen. "Chilling Effects and Unequal Subjects: A Response to Jonathon Penney's Understanding Chilling Effects." *Minnesota Law Review* 106 (2022): 392-99.

Litowitz, Douglas E. "Gramsci, Hegemony, and the Law." *BYU Law Review* 2000 (2000): 515-51.

Loftus, Joshua R., Chris Russell, Matt J. Kusner, and Ricardo Silva. "Causal Reasoning for Algorithmic Fairness." *ArXiv* abs/1805.05859 (2018).

Loick, Daniel. "“… as If It Were a Thing.” a Feminist Critique of Consent." *Constellations* 27, no. 3 (2020): 412-22.

Luna, F. "Identifying and Evaluating Layers of Vulnerability - a Way Forward." [In eng]. *Dev World Bioeth* 19, no. 2 (Jun 2019): 86-95.

Luna, Florencia. "Elucidating the Concept of Vulnerability: Layers Not Labels." *International Journal of Feminist Approaches to Bioethics* 2, no. 1 (2009): 121-39.

Lynskey, O. *The Foundations of Eu Data Protection Law*. Oxford University Press, 2015.

Mackenzie, Catriona, and Natalie Stoljar. *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self.* Vol. 17: Oxford University Press, 2000.

MacKinnon, Catharine A. *Feminism Unmodified : Discourses on Life and Law.* [in English] Cambridge, Mass.: Harvard University Press, 1987.

———. "Feminism, Marxism, Method, and the State: Toward Feminist Jurisprudence." *Signs* 8, no. 4 (1983): 635-58.

———. *Toward a Feminist Theory of the State.* Cambridge, Massachusetts: New York University Press, 1989.

Malgieri, Gianclaudio. "Automated Decision-Making in the Eu Member States: The Right to Explanation and Other 'Suitable Safeguards' for Algorithmic Decisions in the Eu National Legislations." *Computer Law & Security Review*, 2019 Forthcoming (August 17, 2018).

———. "'Ownership' of Customer (Big) Data in the European Union: Quasi-Property as Comparative Solution?", 2016.

———. "Right to Explanation and Algorithm Legibility in the Eu Member States Legislations." *SSRN Electronic Journal* (01/01 2018).

Malgieri, Gianclaudio, and Gloria González Fuster. "The Vulnerable Data Subject: A Gendered Data Subject?". *SSRN Electronic Journal* (01/01 2021).

Malgieri, Gianclaudio, and Jędrzej Niklas. "Vulnerable Data Subjects." *Computer Law & Security Review* 37 (2020/07/01/ 2020): 105415.

Mann, Steve, and Joseph Ferenbok. "New Media and the Power Politics of Sousveillance in a Surveillance-Dominated World." *Surveillance & Society* 11, no. 1/2 (2013): 18-34.

Mantelero, Alessandro. "Ai and Big Data: A Blueprint for a Human Rights, Social and Ethical Impact Assessment." *Computer Law & Security Review* 34, no. 4 (2018/08/01/ 2018): 754-72.

Marks, Susan. "The European Convention on Human Rights and Its 'Democratic Society'." *British Yearbook of International Law* 66, no. 1 (1996): 209-38.

Marshall, Brandeis. "Algorithmic Misogynoir in Content Moderation Practice." (2021). https://eu.boell.org/sites/default/files/2021-06/HBS-e-paper-Algorithmic-Misogynoir-in-Content-Moderation-Practice-200621_FINAL.pdf.

Mathiesen, Thomas. "The Viewer Society:Michel Foucault's `Panopticon' Revisited." *Theoretical Criminology* 1, no. 2 (1997): 215-34.

Melkas, Helina, and Richard Anker. "Occupational Segregation by Sex in Nordic Countries: An Empirical Investigation." *Int'l Lab. Rev.* 136 (1997): 341.

Mendoza, Isak, and Lee A. Bygrave. "The Right Not to Be Subject to Automated Decisions Based on Profiling." In *Eu Internet Law: Regulation and Enforcement*, edited by Tatiana-Eleni Synodinou, Philippe Jougleux, Christina Markou and Thalia Prastitou: Springer, 2017.

Mendus, Susan. "Loosing the Faith: Feminism and Democracy." In *Democracy the Unfinished Journey: 508 Bc to Ad1993*, edited by John Dunn. New York: Oxford University Press, 1992.

Metwaly, Kareem M., Aerin Kim, Elliot Branson, and Vishal Monga. "Car - Cityscapes Attributes Recognition a Multi-Category Attributes Dataset for Autonomous Vehicles." *ArXiv* abs/2111.08243 (2021).

Michelman, Frank I., and Kathleen M. Sullivan. "The Supreme Court, 1985 Term." *Harvard Law Review* 100, no. 1 (1986): 1-311.

Mill, John Stuart. "On Liberty." In *On Liberty*, edited by David Bromwich, George Kateb and Jean Bethke Elshtain. Rethinking the Western Tradition. New Haven: Yale University Press, 2003.

Minow, Martha, and Donald C. Langevoort. "The Supreme Court, 1986 Term." *Harvard Law Review* 101, no. 1 (1987): 7-370.

Mittelstadt, Brent. "From Individual to Group Privacy in Big Data Analytics." *Philosophy & Technology* 30 (12/01 2017).

Mulligan, Deirdre K., Colin Koopman, and Nick Doty. "Privacy Is an Essentially Contested Concept: A Multi-Dimensional Analytic for Mapping Privacy." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374 (2016).

Mulvey, Laura. "Visual Pleasure and Narrative Cinema." *Screen* 16, no. 3 (1975): 6-18.

Murphy, Jeffrie G., and Jules L. Coleman. *Philosophy of Law : An Introduction to Jurisprudence.* [in English] Boulder: Westview Press, 1990.

Murray, A. *Information Technology Law: The Law and Society.* 3 ed.: Oxford University Press, 2016.

Naudts, Laurens. "How Machine Learning Generates Unfair Inequalities and How Data Protection Instruments May Help in Mitigating Them ". In *Data Protection and Privacy: The Internet of Bodies*, edited by R. Leenes, R. van Brakel, S. Gutwirth and P. De Hert. Oxford: Hart Publishing, 2019.

Nissenbaum, Helen. "Privacy in Context - Technology, Policy, and the Integrity of Social Life." 2009.

———. "Respect for Context as a Benchmark for Privacy Online: What It Is and Isn't." In *Social Dimensions of Privacy: Interdisciplinary Perspectives*, edited by Beate Roessler and Dorota Mokrosinska, 278-302. Cambridge: Cambridge University Press, 2015.

Nissim, Kobbi, Aaron Bembenek, Alexandra Wood, Mark Bun, Marco Gaboardi, Urs Gasser, David O'Brien, Salil P. Vadhan, and Thomas Steinke. "Bridging the Gap between Computer Science and Legal Approaches to Privacy." *Harvard Journal of Law & Technology* 31 (2018): 687.

Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism.* New York: New York University Press, 2018.

Nouwens, Midas, Ilaria Liccardi, Michael Veale, David Karger, and Lalana Kagal. "Dark Patterns after the Gdpr: Scraping Consent Pop-Ups and Demonstrating Their Influence." In *Proceedings of the 2020 Chi Conference on Human Factors in Computing Systems*, 1–13: Association for Computing Machinery, 2020.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.* Penguin Books, 2016.

Olsen, Frances. "Statutory Rape: A Feminist Critique of Rights Analysis." *Texas Law Review* 63, no. 3 (1984): 387-432.

Orwat, Carsten. "Risks of Discrimination through the Use of Algorithms. A Study Compiled with a Grant from the Federal Anti-Discrimination Agency." 2020.

Overdorf, Rebekah, Bogdan Kulynych, Ero Balsa, Carmela Troncoso, and Seda F. Gürses. "Questioning the Assumptions Behind Fairness Solutions." *ArXiv* abs/1811.11293 (2018).

Pasquale, Frank. *Black Box Society : The Secret Algorithms That Control Money and Information.* First Harvard University Press paperback edition. ed. Cambridge, Massachusetts: Harvard University Press, 2016.

Pateman, Carole. *The Disorder of Women : Democracy, Feminism, and Political Theory.* [in English] Stanford, Calif.: Stanford University Press, 1989.

Pech, Laurent. *The Concept of Chilling Effect: Its Untapped Potential to Better Protect Democracy, the Rule of Law, and Fundamental Rights in the Eu*. Open Society European Policy Institute, 2021. https://www.opensocietyfoundations.org/publications/the-concept-of-chilling-effect#publications_download.

Pendergrast, Kelly. "The Next Big Cheap." In *Real Life*, 2019.

Penney, Jonathon. "Understanding Chilling Effects." *Minnesota Law Review* 106 (2021): 1451-530.

Penney, Jonathon W. "Chilling Effects: Online Surveillance and Wikipedia Use." *Berkeley Technology Law Journal* 31, no. 1 (2016): 117-82.

———. "Internet Surveillance, Regulation, and Chilling Effects Online: A Comparative Case Study." *Internet Policy Review* 6, no. 2 (2017). doi:10.14763/2017.2.692, https://policyreview.info/articles/analysis/internet-surveillance-regulation-and-chilling-effects-online-comparative-case.

Phillips, P., Hahn Carina, Fontana Peter, Yates Amy, Greene Kristen, Broniatowski David, and Przybocki Mark. "Four Principles of Explainable Artificial Intelligence." NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, Gaithersburg, MD, 2021.

Pohle, Jörg. "Datenschutz Und Technikgestaltung: Geschichte Und Theorie Des Datenschutzes Aus Informatischer Sicht Und Folgerungen Für Die Technikgestaltung." Humboldt-Universität zu Berlin, 2018.

———. "Privacy and Data Protection by Design: A Critical Perspective." In *Privacy and Cyber Security on the Books and on the Ground*, edited by Ingolf Pernice and Jörg Pohle, 134-41. Berlin, Germany: Alexander von Humboldt Institute for Internet and Society, 2019.

Polan, Diane. "Toward a Theory of Law and Patriarchy ". In *Feminist Legal Theory : Foundations*, edited by D. Kelly Weisberg. Philadelphia: Temple University Press, 1993.

Post, Paul, and Rikki Holtmaat. "A False Start: Discrimination in Job Advertisements." *European Gender Equality Law Review*, no. 2/2014 (2015): 12-22.

Postmus, Judy L., Gretchen L. Hoge, Jan Breckenridge, Nicola Sharp-Jeffs, and Donna Chung. "Economic Abuse as an Invisible Form of Domestic Violence: A Multicountry Review." *Trauma, Violence, & Abuse* 21, no. 2 (2020): 261-83.

Prainsack, Barbara. "Logged Out: Ownership, Exclusion and Public Value in the Digital Data and Information Commons." *Big Data & Society* 6, no. 1 (2019): 2053951719829773.

Prietl, Bianca. "Big Data: Inequality by Design?" Paper presented at the Weizenbaum Conference 2019 "Challenges of Digital Inequality - Digital Education, Digital Work, Digital Life", Berlin, 2019.

Prince, Anya, and Daniel B. Schwarcz. "Proxy Discrimination in the Age of Artificial Intelligence and Big Data." *Iowa Law Review* 105 (2020): 1257-318.

Purtova, Nadezhda. "Do Property Rights in Personal Data Make Sense after the Big Data Turn?: Individual Control and Transparency." *ISN: Property Protection (Topic)* (2017).

———. "The Law of Everything. Broad Concept of Personal Data and Future of Eu Data Protection Law." *Law, Innovation and Technology* 10, no. 1 (2018/01/02 2018): 40-81.

———. "Property in Personal Data: A European Perspective on Instrumentalist Theory of Propertization." *Research Policy - RES POLICY* (01/01 2010).

Raghavan, Manish, Solon Barocas, Jon Kleinberg, and Karen Levy. "Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 469–81. Barcelona, Spain: Association for Computing Machinery, 2020.

renan barzilay, Arianne, and Anat Ben-David. "Platform Inequality: Gender in the Gig-Economy." *Seton Hall Law Review* 47 (02/01 2017): 393-431.

Reva, Schwartz, Vassilev Apostol, Greene Kristen, Perine Lori, Burt Andrew, and Hall Patrick. "Towards a Standard for Identifying and Managing Bias in

Artificial Intelligence." Special Publication (NIST SP), National Institute of Standards and Technology, Gaithersburg, MD, 2022.

Rifkin, Janet. "Toward a Theory of Law and Patriarchy." *Harv. Women's L.J.* 3 (1980): 83.

Ritter, Jeffrey, and Anna Mayer. "Regulating Data as Property: A New Construct for Moving Forward." (03/06 2018).

Roberts, Jessica L. "Protecting Privacy to Prevent Discrimination." *William and Mary Law Review*, 2015, 2097+.

Romei, Andrea, and Salvatore Ruggieri. "A Multidisciplinary Survey on Discrimination Analysis." *The Knowledge Engineering Review* 29, no. 5 (2014): 582-638.

Rouvroy, Antoinette, and Yves Poullet. "The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Inportance of Privacy for Democracy." In *Reinventing Data Protection?*, edited by Serge Gutwirth, Yves Poullet, Paul De Hert, Cécile de Terwangne and Sjaak Nouwt, 45-76. New York, NY: Springer, 2009.

Sánchez-Monedero, Javier, Lina Dencik, and Lilian Edwards. "What Does It Mean to 'Solve' the Problem of Discrimination in Hiring? Social, Technical and Legal Perspectives from the Uk on Automated Hiring Systems." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 458–68. Barcelona, Spain: Association for Computing Machinery, 2020.

Scales, Ann C. "The Emergence of Feminist Jurisprudence: An Essay." *Yale Law Journal* 95 (1986): 1373.

Schermer, Bart. "Risks of Profiling and the Limits of Data Protection Law." In *Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases*, edited by Bart Custers, Toon Calders, Bart Schermer and Tal Zarsky: Springer Publishing Company, Incorporated, 2013.

Schiek, Dagmar. "On Uses, Mis-Uses and Non-Uses of Intersectionality before the Court of Justice (Eu)." *International Journal of Discrimination and the Law* 18, no. 2-3 (2018): 82-103.

———. "Organising Eu Equality Law around the Nodes of 'Race', Gender and Disability." In *European Union Non-Discrimination Law and Intersectionality : Investigating the Triangle of Racial, Gender and Disability Discrimination*, edited by Dagmar Schiek and Anna Lawson. London and New York: Routledge, 2011.

Schneider, Elizabeth Marie. "The Dialectic of Rights and Politics: Perspectives from the Women's Movement [1986]." *New York University Law Review* 61 (1986): 318-32.

Schüll, Natasha Dow. "Data for Life: Wearable Technology and the Design of Self-Care." *BioSocieties* 11, no. 3 (2016/09/01 2016): 317-33.

Selbst, Andrew D. "Disparate Impact in Big Data Policing." *Georgia law review* 52 (2017): 3373.

———. "An Institutional View of Algorithmic Impact Assessments." *Harvard Journal of Law & Technology* 35, no. 1 (2021): 117-91.

Selbst, Andrew D., and Solon Barocas. "The Intuitive Appeal of Explainable Machines." *Fordham Law Review* 87 (2018): 1085.

Selbst, Andrew D., and Julia Powles. "Meaningful Information and the Right to Explanation." *International Data Privacy Law* 7, no. 4 (2017): 233-42.

Sheard, Natalie. "Employment Discrimination by Algorithm: Can Anyone Be Held Accountable?" [In English]. *UNIVERSITY OF NEW SOUTH WALES LAW JOURNAL* 45, no. 2 (2022): 617-48.

Shephard, Nicole. "Big Data and Sexual Surveillance." 2016.

Siapka, Anastasia, and Elisabetta Biasin. "Bleeding Data: The Case of Fertility and Menstruation Tracking Apps." *Internet Policy Review* 10, no. 4 (2021).

Solove, Daniel J. "Conceptualizing Privacy." *California Law Review* 90, no. 4 (2002): 1087-155.

———. *The Digital Person: Technology and Privacy in the Information Age.* New York, USA: New York University Press, 2004. doi:doi:10.18574/nyu/9780814708965.

———. "'I've Got Nothing to Hide' and Other Misunderstandings of Privacy." 44 (07/12 2007).

———. "The Meaning and Value of Privacy." In *Social Dimensions of Privacy: Interdisciplinary Perspectives*, edited by Beate Roessler and Dorota Mokrosinska, 71-82. Cambridge: Cambridge University Press, 2015.

———. "Privacy and Power: Computer Databases and Metaphors for Information Privacy." *Stanford Law Review* 53 (09/24 2001).

Steinmüller, W., B. Lutterbeck, C. Malimann, U. Harbort, G. Kolb, and J. Schneider. "Grundfragen Des Datenschutzes: Gutachten Im Auftrag Des Bundesministeriums Des Innern [Bt-Drs. Vi/3816 Anlage 1]." 1971.

Strubell, Emma, Ananya Ganesh, and Andrew McCallum. "Energy and Policy Considerations for Deep Learning in Nlp." *ArXiv* abs/1906.02243 (2019).

Stubbs, Margot. "Feminism and Legal Positivism." *Australian Journal of Law and Society* 3 (1986).

Suárez-Gonzalo, Sara. "Personal Data Are Political. A Feminist View on Privacy and Big Data." *Recerca Revista de pensament i anàlisi* 24 (10/01 2019): 173-92.

Sweeney, Latanya. "Discrimination in Online Ad Delivery." *Commun. ACM* 56, no. 5 (2013): 44–54.

Tambe, Prasanna, Peter Cappelli, and Valery Yakubovich. "Artificial Intelligence in Human Resources Management: Challenges and a Path Forward." *California Management Review* 61, no. 4 (2019): 15-42.

Taylor, Linnet, Luciano Floridi, and Bart van der Sloot, eds. *Group Privacy: New Challenges of Data Technologies*: Springer, 2017.

Theilen, Jens T., Andreas Baur, Felix Bieker, Regina Ammicht Quinn, Marit Hansen, and Gloria González Fuster. "Feminist Data Protection: An Introduction." *Internet Policy Review* 10, no. 4 (7 December 2021 2021).

Tobler, Christa. *Limits and Potential of the Concept of Indirect Discrimination.* Luxembourg: : Office for Official Publications of the European Communities, 2008.

Tosoni, Luca. "The Right to Object to Automated Individual Decisions: Resolving the Ambiguity of Article 22(1) of the General Data Protection Regulation." *International Data Privacy Law* 11, no. 2 (2021): 145-62.

Tufekci, Zeynep. *Twitter and Tear Gas: The Power and Fragility of Networked Protest*. United States of America: Yale University Press, 2017. https://www.twitterandteargas.org/downloads/twitter-and-tear-gas-by-zeynep-tufekci.pdf.

Unger, Roberto Mangabeira. *Knowledge & Politics.* [in English] New York: Free Press, 1975.

van der Veer, Sabine N, Lisa Riste, Sudeh Cheraghi-Sohi, Denham L Phipps, Mary P Tully, Kyle Bozentko, Sarah Atwood, *et al.* "Trading Off Accuracy and Explainability in Ai Decision-Making: Findings from 2 Citizens' Juries." *Journal of the American Medical Informatics Association* 28, no. 10 (2021): 2128-38.

Veale, Michael, and Reuben Binns. "Fairer Machine Learning in the Real World: Mitigating Discrimination without Collecting Sensitive Data." *Big Data and Society* 4, no. 2 (November 20, 2017 2017).

Veale, Michael, and Frederik Borgesius. *Demystifying the Draft Eu Artificial Intelligence Act.* 2021. doi:10.31235/osf.io/38p5f.

Veale, Michael, and Lilian Edwards. "Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling." *Computer Law & Security Review* 34, no. 2 (2018/04/01/ 2018): 398-404.

Verma, S., and J. Rubin. "Fairness Definitions Explained." Paper presented at the 2018 IEEE/ACM International Workshop on Software Fairness (FairWare), 29-29 May 2018 2018.

Vogiatzoglou, Plixavra, and Peggy Valcke. "Chapter 1: Two Decades of Article 8 Cfr: A Critical Exploration of the Fundamental Right to Personal Data Protection in Eu Law ". In *Research Handbook on Eu Data Protection Law*, edited by Eleni Kosta, Ronald Leenes and Irene Kamara: Edward Elgar Publishing, 2022.

von Grafenstein, Maximilian. *The Principle of Purpose Limitation in Data Protection Laws: The Risk-Based Approach, Principles, and Private Standards as Elements for Regulating Innovation.* [in en] Schriften Zur Rechtswissenschaftlichen Innovationsforschung. 1 ed. Vol. 12, Baden-Baden: Nomos Verlagsgesellschaft mbH & Co. KG, 2018. doi:10.5771/9783845290843.

———. "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part I." *European Data Protection Law Review* 6, no. 4 (2020).

———. "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Ii." *European Data Protection Law Review* 7, no. 2 (2021).

———. "Refining the Concept of the Right to Data Protection in Article 8 Ecfr – Part Iii." *European Data Protection Law Review* 7, no. 3 (2021).

Wachter, Sandra. "Affinity Profiling and Discrimination by Association in Online Behavioural Advertising (May 15, 2019). , Vol. 35, No. 2, 2020, Forthcoming, Available at Ssrn: Https://Ssrn.Com/Abstract=3388639 or Http://Dx.Doi.Org/10.2139/Ssrn.3388639." *Berkeley Technology Law Journal* 35, no. 2 (2020, Fourthcoming 2020).

———. "Privacy: Primus Inter Pares — Privacy as a Precondition for Self-Development, Personal Fulfilment and the Free Enjoyment of Fundamental Human Rights." (2017). https://ssrn.com/abstract=2903514.

Wachter, Sandra, and Brent Mittelstadt. "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and Ai." *Columbia Business Law Review* 2019, no. 2 (04/23 2019).

Wachter, Sandra, Brent Daniel Mittelstadt, and Chris Russell. "Bias Preservation in Machine Learning: The Legality of Fairness Metrics under Eu Non-Discrimination Law." *SSRN Electronic Journal* (2021).

Wachter, Sandra, Brent Mittelstadt, and Luciano Floridi. "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation." *International Data Privacy Law* 7, no. 2 (2017): 76-99.

Wachter, Sandra, Brent Mittelstadt, and Chris Russell. "Counterfactual Explanations without Opening the Black Box: Automated Decisions and the Gdpr." *Harvard journal of law & technology* 31 (04/01 2018): 841-87.

———. "Why Fairness Cannot Be Automated: Bridging the Gap between Eu Non-Discrimination Law and Ai ". (2020).

Wajcman, Judy. "Feminist Theories of Technology." *Cambridge Journal of Economics* 34, no. 1 (2009): 143-52.

———. *Technofeminism.* [in English] Cambridge: Polity, 2004.

Walby, S. *Theorizing Patriarchy.* Basil Blackwell, 1990.

Ward, Angela. "The Impact of the Eu Charter of Fundamental Rights on Anti-Discrimination Law: More a Whimper Than a Bang?". *Cambridge Yearbook of European Legal Studies* 20 (2018): 32-60.

Weisberg, D. Kelly. *Feminist Legal Theory : Foundations.* [in English] Philadelphia: Temple University Press, 1993.

Wernick, Alina, and Deniz Erden. "Computer Says Hausfrau - Can Automated Credit Scoring Contribute to Gendered Digital Divide?". *Encore - The Annual Magazine on Internet and Society Research* 2020/2021 (2021): 36-46.

West, Robin. "Introduction to the Research Handbook on Feminist Jurisprudence." In *Research Handbook on Feminist Jurisprudence*, edited by Robin West and Cynthia Grant Bowman: Edward Elgar Publishing Limited, 2019.

Westin, Alan. *Privacy and Freedom*. New York: Ig Publishing, 1967.

White, Jefferson, and Dennis M. Patterson. *Introduction to the Philosophy of Law : Readings and Cases.* [in English] New York: Oxford University Press, 1999.

Wishik, Heather Ruth. "To Question Everything: The Inquiries of Feminist Jurisprudence." *Berkeley Journal of Gender, Law and Justice* 1 (1985): 64.

Wulf, Jessica. *Automated Decision-Making Systems and Discrimination: Understanding Causes, Recognizing Cases, Supporting Those Affected* Berlin: AlgorithmWatch, 2022. https://algorithmwatch.org/en/autocheck/.

Xenidis, Raphaële. "Tuning Eu Equality Law to Algorithmic Discrimination: Three Pathways to Resilience." *Maastricht Journal of European and Comparative Law* 27, no. 6 (2020): 736-58.

Xenidis, Raphaële, and Linda Senden. "Eu Non-Discrimination Law in the Era of Artificial Intelligence: Mapping the Challenges of Algorithmic Discrimination." 2020.

Xiang, Alice, and Inioluwa Deborah Raji. "On the Legal Compatibility of Fairness Definitions." *ArXiv* abs/1912.00761 (2019).

Yeung, Karen. "Algorithmic Regulation: A Critical Interrogation." *Regulation & Governance* 12, no. 4 (2018): 505-23.

Young, Cate. "This Is What I Mean When I Say “White Fem- Inism”.", 2014.

Zarsky, Tal Z. "Incompatible: The Gdpr in the Age of Big Data." *Seton Hall L. Rev.* 47 (2016): 995.

Žliobaitė, Indrė, and Bart Custers. "Using Sensitive Personal Data May Be Necessary for Avoiding Discrimination in Data-Driven Decision Models." *Artificial Intelligence and Law* 24, no. 2 (2016/06/01 2016): 183-201.

Zuboff, Shoshana. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. First ed. New York: PublicAffairs, 2019.

CASES

Aydin v. Turkey App no 23178/94 (ECtHR, 25 September 1997)

Big Brother Watch And Others v. The United Kingdom (Applications nos. 58170/13, 62322/14 and 24960/15) 25 May 2021 https://hudoc.echr.coe.int/eng?i=001-210077

C 644/19 [2020] *FT v Universitatea 'Lucian Blaga' Sibiu, GS and Others, HS, Ministerul Educaţiei Naţionale* ECLI:EU:C:2020:810

C-1/95 [1997] *Hellen Gerster v Freistaat Bayern* ECLI:EU:C:1997:452

C-101/01 [2003] *Lindqvist* ECLI:EU:C:2003:596;

C-129/96 [1997] *Inter-Environnement Wallonie ASBL v Région wallonne* ECLI:EU:C:1997:628;

C-13/05 [2006] *Sonia Chacón Navas v Eurest Colectividades SA,* ECLI:EU:C:2006:456

C-144/04 [2005] *Werner Mangold v Rüdiger Helm.* ECLI:EU:C:2005:709;

C-152/84 [1986] *Marshall v Southampton and South-West Hampshire Area Health Authority (Teaching)* ECLI:EU:C:1986:84;

C-158/97 [2000] *Badeck and others v Landesanwalt beim Staatsgerichtshof des Landes Hessen* ECLI:EU:C:2000:163

C-161/18 [2019] *Villar Láiz* ECLI:EU:C:2019:382; C-527/13 [2015] *Cachaldora Fernández* ECLI:EU:C:2015:215

C-167/97 [1999] *Regina v Secretary of State for Employment, ex parte Nicole Seymour-Smith and Laura Perez* ECLI:EU:C:1999:60

C-171/88 [1989] *Rinner-Kühn v FWW Spezial-Gebäudereinigung* ECLI:EU:C:1989:328

C-171/88 [1989] *Rinner-Kühn v FWW Spezial-Gebäudereinigung* Opinion of Darmon AG ECLI:EU:C:1989:158

C-177/88 [1990] *Dekker v Stichting Vormingscentrum voor Jong Volwassenen* ECLI:EU:C:1990:383;

C-184/89 [1991] *Nimz v Freie und Hansestadt Hamburg* ECLI:EU:C:1991:50

C-188/15 [2017] *Bougnaoui and Association de défense des droits de l'homme (ADDH) v Micropole SA* ECLI:EU:C:2017:204

C-196/02 [2005] *Vasiliki Nikoloudi v Organismos Tilepikinonion Ellados AE* ECLI:EU:C:2005:141

C-201/02 [2004] *Wells v Secretary of State for Transport, Local Government and the Regions* ECLI:EU:C:2004:12.

C-229/08 [2010] *Wolf v Stadt Frankfurt am Main* ECLI:EU:C:2010:3

C-249/97 [1999] *Gruber v Silhouette International Schmied GmbH & Co. KG.* ECLI:EU:C:1999:405

C-25/17 [2018] *Jehovan todistajat* ECLI:EU:C:2018:551

C-256/01 [2004] *Allonby v Accrington & Rossendale College* ECLI:EU:C:2004:18

C-267/06 [2008] *Tadao Maruko v Versorgungsanstalt der deutschen Bühnen* ECLI:EU:C:2008:179;

C-267/12 [2013] *Frédéric Hay v. Crédit agricole mutuel de Charente-Maritime et des Deux-Sèvres* ECLI:EU:C:2013:823

C-28/08 [2010] *European Commission v The Bavarian Lager Co. Ltd.* ECLI:EU:C:2010:378

C-281/97 [1999] *Andrea Krüger v Kreiskrankenhaus Ebersberg* ECLI:EU:C:1999:396

C-300/06 [2007] *Ursula Voß v Land Berlin* ECLI:EU:C:2007:757

C-303/06 [2008] *S. Coleman v Attridge Law and Steve Law* ECLI:EU:C:2008:415

C-311/18 [2020] *Data Protection Commissioner v Facebook Ireland Limited and Maximillian Schrems* ECLI:EU:C:2020:559.

C-32/93 [1994] *Webb v EMO Air Cargo* ECLI:EU:C:1994:300

C-33/89 [1990] *Maria Kowalska v Freie und Hansestadt Hamburg* ECLI:EU:C:1990:265;

C-343/92 [1994] *De Weerd, Roks and Others v Bestuur van de Bedrijfsvereniging voor de Gezondheid, Geestelijke en Maatschappelijke Belangen and Others* ECLI:EU:C:1994:71

C-356/09 [2010] *Pensionsversicherungsanstalt v Kleist* ECLI:EU:C:2010:703

C-362/14 [2015] *Maximillian Schrems v Data Protection Commissioner* ECLI:EU:C:2015:650

C-363/12 [2014] *Z. v A Government department and The Board of management of a community school* ECLI:EU:C:2014:159

C-409/95 [1997] *Marschall v Land Nordrhein-Westfalen* ECLI:EU:C:1997:533

C-423/15 [2016] *Nils-Johannes Kratzer v R+V Allgemeine Versicherung AG* ECLI:EU:C:2016:604

C-434/16 [2017] *Peter Nowak v Data Protection Commissioner* ECLI:EU:C:2017:994

C-441/14 [2016] *Dansk Industri (DI) v Estate of Karsten Eigil Rasmussen* ECLI:EU:C:2016:278

C-447/09 [2011] *Prigge and Others* ECLI:EU:C:2011:573

C-450/93 [1995] *Kalanke v Freie Hansestadt Bremen* ECLI:EU:C:1995:322

C-450/93 [1995] *Kalanke v Freie Hansestadt Bremen* Opinion of the Tesauro AG ECLI:EU:C:1995:105 at 3065.

C-527/13 [2015] *Lourdes Cachaldora Fernández v Instituto Nacional de la Seguridad Social (INSS), Tesorería General de la Seguridad Social (TGSS)* ECLI:EU:C:2015:215

C-54/07 [2008] *Feryn* ECLI:EU:C:2008:397

C-553/11 [2012] *Rintisch v Klaus Eder* ECLI:EU:C:2012:671

C-555/07 [2010] *Seda Kücükdeveci v Swedex GmbH & Co. KG.* ECLI:EU:C:2010:21

C-582/14 [2016] *Patrick Breyer V. Bundesrepublic Deutschland* ECLI:EU:C:2016:779

C-6/64 [1964] Flaminio Costa v E.N.E.L. ECLI:EU:C:1964:66

C-703/17 [2019] *Krah v Universität Wien*, Opinion of Leger AG, ECLI:EU:C:2019:450

C-81/12 [2013] *Asociaţia Accept* ECLI:EU:C:2013:275

C-91/92 [1994] *Paola Faccini Dori v Recreb Sri* ECLI:EU:C:1994:292

C-141/12 and C-372/12 Joined Cases [2014] *YS (C-141/12) v Minister voor Immigratie, Integratie en Asiel, and Minister voor Immigratie, Integratie en Asiel (C-372/12) v M, S.* ECLI:EU:C:2014:2081

C-258/15 [2016] *Salaberria Sorondo v Academia Vasca de Policía y Emergencias* ECLI:EU:C:2016:873

C-293/12 and C 594/12 Joined Cases [2014] *Digital Rights Ireland Ltd* Opinion of Villalón AG, ECLI:EU:C:2013:845

C-293/12 and C-594/12 Joined Cases [2014] *Digital Rights Ireland Ltd* ECLI:EU:C:2014:238

C-319/20 [2022] *Meta (Facebook) v. Verbraucherzentrale Bunderverband (vzbv) Meta (Facebook) V. Verbraucherzentrale Bundesverband (Vzbv)* ECLI:EU:C:2022:322.

C-335/11 and C-337/11 [2013], *HK Danmark v Dansk almennyttigt Boligselskab (C-335/11), and HK Danmark v Dansk Arbejdsgiverforening (C-337/11)* ECLI:EU:C:2013:222

C-668/15 [2017] *Jyske Finans A/S v Ligebehandlingsnævnet, acting on behalf of Ismar Huskic* ECLI:EU:C:2017:278

C-136/17 [2019] *GC,AF, BH, ED v Commission nationale de l'informatique et des libertés* (CNIL) ECLI:EU:C:2019:773

C-157/15 [2017] *Achbita, Centrum voor gelijkheid van kansen en voor racismebestrijding v G4S Secure Solutions NV* ECLI:EU:C:2017:203

C-16/19 [2021] *VL v Szpital Kliniczny im. dra J. Babińskiego Samodzielny Publiczny Zakład Opieki Zdrowotnej w Krakowie* ECLI:EU:C:2021:64

C-210/16 [2018] *Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie Schleswig-Holstein GmbH* ECLI:EU:C:2018:388

C-212/13 [2014] *Ryneš v Úřad pro ochranu osobních údajů* ECLI:EU:C:2014:2428

C-354/13 [2014] *Fag og Arbejde (FOA) v Kommunernes Landsforening (KL)* ECLI:EU:C:2014:2463

C-414/16 [2018] *Vera Egenberger v Evangelisches Werk für Diakonie und Entwicklung eV* ECLI:EU:C:2018:257

C-415/10 [2012] *Meister v Speech Design Carrier Systems GmbH* ECLI:EU:C:2012:217

C-434/15 [2017] *Asociación Profesional Élite Taxi v Uber Systems Spain SL* ECLI:EU:C:2017:981

C-443/15 [2016] *David L. Parris v Trinity College Dublin and Others* ECLI:EU:C:2016:897

C-443/16 [2017] *Francisco Rodrigo Sanz v Universidad Politécnica de Madrid* ECLI:EU:C:2017:109

C-486/08 [2010] *Zentralbetriebsrat der Landeskrankenhäuser Tirols v Land Tirol* ECLI:EU:C:2010:215

C-507/18 [2020] *NH v Associazione Avvocatura per i diritti LGBTI — Rete Lenford* ECLI:EU:C:2020:289.

C-83/14 [2015] *CHEZ Razpredelenie Bulgaria AD v Komisia za zashtita ot diskriminatsia* ECLI:EU:C:2015:480.

Case 109/88 [1989] *Handels- og Kontorfunktionærernes Forbund I Danmark v Dansk Arbejdsgiverforening, acting on behalf of Danfoss* ECLI:EU:C:1989:383

Case 148/78 [1979] *Ratti* ECLI:EU:C:1979:110

Case 170/84 [1986] *Bilka-Kaufhaus GmbH v Karin Weber von Hartz* ECLI:EU:C:1986:204

Case 26-62 [1963] *Van Gend & Loos v Netherlands Inland Revenue Administration* ECLI:EU:C:1963:1

Case 352/85 [1988] *Bond van Adverteerders and others v The Netherlands* State ECLI:EU:C:1988:196

Case 41-74 [1979] *Van Duyn v Home Office* ECLI:EU:C:1974:133

Case 43-75 [1976] *Defrenne v Sabena* ECLI:EU:C:1976:56

Centrum För Rättvisa v. Sweden (Application no. 35252/08) 25 May 2021 https://hudoc.echr.coe.int/eng?i=001-210078

Delfi AS v. Estonia (Application no. 64569/09) 16 June 2015 https://hudoc.echr.coe.int/eng?i=001-155105

Ezelin v. France (Application no. 11800/85) 26 April 1991 https://hudoc.echr.coe.int/eng?i=001-57675

Handyside v. The United Kingdom (Application no. 5493/72) 7 December 1976 https://hudoc.echr.coe.int/eng?i=001-57499

Kiyutin v. Russia App no 2700/10 (ECtHR, 10 March 2011).

Klass and Others v. Germany (Application no. 5029/71) 6 September 1978 https://hudoc.echr.coe.int/eng?i=001-57510

Lingens v. Austria (Application No . 9815/82) 8 July 1986 https://hudoc.echr.coe.int/eng?i=001-57523

Litigation Chamber decision on the merits 21/2022 of 2 February 2022 of Belgian Data Protection Authority (Case number: DOS-2019-01377) <https://edpb.europa.eu/system/files/2022-03/be_2022-02_decisionpublic_0.pdf> accessed 24 April 2022.

M.M. v. The United Kingdom (Application no. 24029/07) 13 November 2012

M.S.S. v. Belgium and Greece (App no 30696/09) 21 January 2011

Magyar Kétfarkú Kutya Párt v. Hungary (Application no. 201/17) 20 January 2020 https://hudoc.echr.coe.int/eng?i=001-200657

Rotaru v. Romania (Application no. 28341/95) 4 May 2000

Taşkın and Others v. Turkey, App. No. 46117/99 (2 February 2010).

Young, James and Webster v. The United Kingdom (Application no. 7601/76; 7806/77) 13 August 1981 https://hudoc.echr.coe.int/eng?i=001-57604