

Customer ratings as a vector for discrimination in employment relations? Pathways and pitfalls for legal remedies

Marco Biagi Conference “Assessing Worker performance in a Changing Technological and Societal Environment”, Modena, 19-21 March 2018

ROSSANA DUCATO*, MIRIAM KULLMANN† and MARCO ROCCA‡

“[phone beeps]
[gasps] ‘Two stars?! Two stars?’
‘Wasn’t a meaningful encounter’”

--NOSEDIVE, BLACK MIRROR SEASON 3, EP. 1

1. Introduction

Screenwriters¹ and novelists² have already explored the dystopian prospects of a society where human interactions are constantly submitted to ratings. This is inspired by the growing presence in our daily lives of the possibility of rating work performances, mostly through the well-known system of 1-5 “stars”. This practice is particularly widespread in the gig-economy, with the case of Uber being the most visible example³. In the words of the well-known UK *Aslam and Ferrar v Uber* case⁴, this “amounts to a performance management/disciplinary procedure”, as customer ratings are directly used by the platform to decide upon the relationship with the given driver, up to the deactivation of his or her account. However, the use made by Uber of customer’s ratings has also been described as a vehicle for workplace discrimination, in light of the high potential for biases to creep into evaluations for drivers.⁵

* Ph.D., Postdoctoral researcher at Université Catholique de Louvain and Université Saint-Louis – Bruxelles (Belgium). Her research is supported by the Innoviris grant 2016-BB2B-9.

† Ph.D., Assistant Professor at Vienna University of Economics and Business (WU) (Austria).

‡ Ph.D., Postdoctoral researcher at Hasselt University, Lecturer at Université de Liège and Université Libre de Bruxelles (Belgium).

¹ Shteyngart, *Super Sad True Love Story*, Random House, 2010.

² Brooker, *Nosedive*, Black Mirror, Season 3 Episode 1, 21 October 2016.

³ Dzieza, “The Rating Game”, *The Verge*, 28 October 2015, available at <https://www.theverge.com/2015/10/28/9625968/rating-system-on-demand-economy-uber-olive-garden>.

⁴ *Aslam, Ferrar and Others v. Uber* [2017] IRLR 4 (ET).

⁵ Rosenblat, Levy, Barocas, and Hwang, Discriminating Tastes: Uber’s Customer Ratings as Vehicles for Workplace Discrimination (2017), 9 *Policy and Internet* 3, p. 256-279.

Far from being inextricably intertwined with the specific nature of work on demand via app, the phenomenon has a high potential of spilling-over in other sectors, and notably to those characterised by customer-facing employees, from call centres to restaurants⁶. In fact, in our own work experience we are directly confronted with students' evaluations, which rate our performance as teachers. These evaluations, we are told, are relevant for the renewal of teaching contracts, promotions or future applications, although studies have highlighted the potential biases of these evaluations too.⁷

The growing presence of customer ratings is part of the trend towards the “scored society” described by Pasquale and Citron⁸, and shares many of the risks highlighted by both computer and social sciences⁹. We argue that the use of customer ratings in employment-related decision represents a lesser and less refined version of more complex techniques of algorithmic management and/or big data application in the employment context.¹⁰ Hence, the use of customer ratings in employment decisions risks producing discriminatory outcomes, by exacerbating biases existing in society¹¹ and “reproducing many of the same troubling dynamics that have allowed discrimination to persist in society”.¹² What is more, the use of ratings in decisions related to employment provides an easy way to “launder” discrimination, by deploying an instrument which appears as neutral¹³. In this perspective, the use of customer ratings to evaluate worker performance and employment opportunities is indeed worrisome.

Studies on the gig economy already proved that it is not free from discriminatory practices¹⁴. However, taking Uber as a particularly visible case study, no specific data is available to correlate ratings and protected grounds, although Uber seems to collect, at least in the US market, data related to the background of its drivers.¹⁵ However, we argue, the potential of this kind of ratings to reflect biases and prejudices held (even implicitly) by customers can be inferred by looking at comparable effects in other areas. Indeed, biases have been found to shape customer behaviours in online marketplaces, resulting in lower offer prices and decreased response rates¹⁶. Furthermore, past studies have shown disparities in the tipping habits of taxi customers, highlighting how customers are consistently more likely to leave

⁶ Liu, “A new job search app allows restaurant employees and employers to rate each other”, *The Star*, 23 October 2017, available at https://www.thestar.com/life/food_wine/2017/10/28/a-new-job-search-app-allows-restaurant-employees-and-employers-to-rate-each-other.html.

⁷ MacNell, Driscoll and Hunt, “What’s in a Name: Exposing Gender Bias in Student Ratings of Teaching” (2015), *Innovative Higher Education*, 40, p. 291.

⁸ Citron and Pasquale, “The Scored Society”, 89 *Washington Law Review* 1, 2014.

⁹ Pope and Sydnor, “Implementing Anti-Discrimination Policies in Statistical Profiling Models”, 3 *American Economic Journal: Economic Policy* 3, 2011, 206; O’Neil, *Weapons of Math Destruction*, Penguin Random House, 2016.

¹⁰ Lee et al., “Working with Machines: The Impact of Algorithmic and Data Driven Management on Human Workers”, *Proc. 33rd Association for Computing Machinery Conference on Human Factors in Computing Systems*, 2015, p. 1612; Musiani, “Governance by Algorithms”, 2 *Internet Policy Review*, 2013.

¹¹ Zarsky, “Understanding Discrimination in the Scored Society”, 89 *Washington Law Review* 1, 2014, p. 1393

¹² Barocas & Selbst, “Big Data’s Disparate Impact” 104 *Calif. L. Rev.* 671, 2016, p. 732.

¹³ *Ibidem*, p. 1382; Pasquale, *The Black Box Society*, Harvard University Press, 2015, p. 41.

¹⁴ Edelman et al., “Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment”, *American Economic Journal: Applied Economics*, 2016; McAfee & Brynjolfsson, *Machine Platform, Crowd*, p. 202-203.

¹⁵ Vaccaro, MIT Professor Working With Uber to Address Racial Bias, *Boston Globe*, 15 February 2017, available at <https://www.bostonglobe.com/business/2017/02/15/mit-professor-working-withuber-address-racial-bias/MMCyaxlBTofPPVr7gojXOI/story.html>.

¹⁶ See Rosenblat, Levy, Barocas, and Hwang, *supra* note 5, p. 8-10.

lower tips to drivers of minority background¹⁷. As Uber has long discouraged tipping¹⁸, the rating given to the driver can be seen as a proxy for tipping and, as such, seems exposed to the very same kind of behaviour. Thus, we will proceed on the assumption that a similar argument can be made for other sectors exposed to the practice of customer's ratings.

As it was highlighted by Sunstein, “in a competitive market that contains private racism and sexism, then, the existence of third-party pressures can create significant spheres of discrimination”¹⁹, so that bending to the discriminatory preferences of customers might be a rational strategy²⁰ and, as such, impossible to correct without the intervention of regulation. In this sense, we agree with AG Maduro that “market alone will not cure discrimination”²¹. Therefore, a specific attention to the risks posed by the use of customer's ratings in employment relations seems warranted, one that would notably consider the fitness for purpose of non-discrimination law in light of these developments.

These findings have inspired the topic of the present paper. Our analysis focuses on the potential discriminatory effects of the use of customer ratings in employment-related decision. In this perspective, we propose a legal analysis of this phenomenon grounded in EU non-discrimination law. To do so, we draw from the recent case law concerning a seemingly disconnected subject, notably discrimination at the workplace on the ground of religion. Interestingly, the Court of Justice has ruled in this context that “the willingness of the employer to take account of the particular wishes of the customer” cannot constitute a “genuine and determining occupational requirement” capable of justifying a discriminatory practice.²² Hence, we propose an analogy between practices adopted by an employer to satisfy its customers' preferences and choices grounded on biased customer ratings. To do so, we first consider the applicability of EU anti-discrimination law to such a situation, then (Section 3) we turn to the interrelated issues stemming from data protection law. Finally (Section 4) we assess the obstacles and potential outcome of such a challenge, and we propose alternative pathways to limit the potential for discrimination of the use of customer ratings in employment-related decisions.

2. Discrimination and customers' wishes

We now engage with the question to what extent EU equality law is fit to protect workers who have been discriminated because their employer relies for a work-related decision on ratings given by customers who have experienced the service of the respective worker. In

¹⁷ Ayres et al., “To Insure Prejudice: Racial Disparities in Taxicab Tipping”, 114 *Yale Law Journal*, 2005.

¹⁸ Rosenblat & Stark, “Algorithmic Labor and Information Asymmetries: A Case Study of Uber's Drivers”, *International Journal of Communication* 10, 2016, 3758–3784, p. 3775.

¹⁹ Sunstein, “Why markets don't stop discrimination”, in Sunstein (ed.), *Free markets and social justice*, Oxford University Press, 1997, p. 153.

²⁰ *Ibidem*, p. 154: “Suppose that society includes a number of traders who want to interact with blacks or women as well as a number of traders who do not. Suppose, too, that it is difficult for traders who are themselves indifferent to race and sex distinctions to tell which traders are discriminators (a frequent phenomenon in competitive markets). If these conditions are met, *discrimination is likely to be a rational strategy for traders*. New companies will do poorly if they refuse to discriminate. All this suggests that *markets will frequently have great difficulty in breaking down discrimination, if discriminatory tastes are even somewhat widespread*” (emphasis added).

²¹ Opinion of AG Maduro in *Feryn* (C-54/07), para. 18.

²² Case C-188/15 *Boungaoui* ECLI:EU:C:2017:204, para. 40.

many service sectors, employers rely on customer-experiences to rate the performance of their (self-) employed workers. As we saw, the most visible example is constituted by Uber, which allows its customers to rate drivers based on the routes they take, for slow traffic, or for refusing to speed.²³ However, ratings could also be based on, or related to, other criteria, such as sex, age, or race and religion. Current EU equality law – as employment law in general – rests upon the historically developed binary divide between workers and self-employed. Atypical workers, amongst which platform workers, do not always fit easily within either the established worker or self-employed concepts. So, what is the relevance of EU equality laws which are mostly anchored to workers?²⁴ This question, which touches upon the personal scope of employment and equality law, is even more relevant since businesses increasingly rely on customer ratings to decide on who may continue to work for an online platform or not. Given that customer-ratings may have wide-ranging implications for an individual where a business (heavily) relies on them, there is no compelling reason to reserve protection resulting from the application of the non-discrimination/equal treatment principle to a selected group of privileged workers, to the exclusion of self-employed workers.²⁵ Workers, independent or not, are equally exposed to the biases of their customers.

Where an employer relies on its customers to judge the quality of the service provided by his workers, the question arises as to what extent an employer can be held liable for using its customer's rating. Is the customers view attributable to an employer in such a case? Usually, customer ratings are processed through mobile applications or other web-based tools, using specifically designed algorithms. An algorithm can be described as “a formally specified sequence of logical operations that provides step-by-step instructions for computers to act on data and thus automate decisions”.²⁶ The ‘behaviour’ of a digital platform is specified by decisions taken at various levels in the company. Therefore, an algorithm may be designed to take into account requirements defined by the company's customers or any other stakeholder in or outside the company. In the end, there is always a conscious (human) decision that feeds into the algorithm (including its design) and the data model the algorithm will use to solve a particular problem. Algorithms play an important role not only in identifying useful patterns in datasets, but also in making decisions that rely on these patterns.²⁷

There is a widespread, though erroneous, belief that algorithms are free of subconscious biases.²⁸ Mittelstadt et al. highlight the fact that operational parameters are specified by software engineers and “configured by users with desired outcomes in mind that privilege

²³ Cherry, "People Analytics and Invisible Labour", 61 *Saint Louis University Law Journal* 1, 2017, p. 12 (footnote 63).

²⁴ Kullmann, "Platform Work, Algorithmic Decision-Making, and EU Gender Equality Law", 34 *IJCLIR* 1, 2018, p. 3. An exception is Directive 2010/41/EU of 7 July 2010 on the application of the principle of equal treatment between men and women engaged in an activity in a self-employed capacity and repealing Council Directive 86/613/EEC (L 180/1).

²⁵ Fudge, "Self-employment, Women, and Precarious Work: The Scope of Labour Protection", in Fudge & Owens (eds), *Precarious Work, Women, and the New Economy: The Challenge to Legal Norms*, Hart, 2006, p. 219.

²⁶ Barocas & Selbst, *supra* note 12, p. 674.

²⁷ *Ibid*, p. 677.

²⁸ New York Times, "When Algorithms Discriminate" (9 July 2015); The Atlantic, "When Discrimination Is Baked Into Algorithms" (6 September 2015); Isaac & Dixon, "Why big-data analysis of police activity is inherently biased", 2017, available at <http://theconversation.com/why-big-data-analysis-of-police-activity-is-inherently-biased-72640>.

some values and interests over others”.²⁹ Emphasizing the fact that technology is not autonomous, Felix Stalder highlights the influential role of human beings in designing the algorithmic models applied.³⁰ In addition, an algorithm can only be as good as the data it works with, that is the data model the algorithm analyses to take decisions can be biased.³¹ An algorithm may therefore “result in disproportionately adverse outcomes concentrated within historically disadvantaged groups in ways that look a lot like discrimination”.³² Underlining the fact that “[a]lgorithms are not immune from the fundamental problem of discrimination”, Frank Pasquale stresses that “negative and baseless assumptions congeal into prejudice”.³³

Therefore, it is the employer who has a say in how customer ratings do look like and whether and to what extent such ratings should feed into the worker’s performance feedback. By allowing customer ratings to feed into the employer’s decision-making on how the worker performs, even though it is perhaps the algorithm taking the ultimate decision, the employer externalises his right to give directions and instructions as part of the employment relationship. Hence, the customer becomes part of the firm, playing an active role in the service encounter and the worker bearing the responsibility for any failure or success.³⁴ As ratings are mostly made available in relation to service jobs where customer satisfaction is important, customers (can) therefore play a powerful role in determining the terms, conditions, and privileges of employment.³⁵ Consequently, customer ratings can be seen as “good customer service”.³⁶

But is the employer, according to EU equality law, liable for taking discriminatory decisions regarding worker performance, where he relies for his decision on a customer’s point of view? So far, EU case law has only cursorily dealt with the role of customers in relation to employment relationships. Nevertheless, there is limited EU case law where businesses rely on customer preferences, which shows the possible tension between business interests, i.e. satisfying customer needs, and the right of individuals not to be treated differently based on, for instance, their ethnic origin or religion. In *Feryn*, an employer declared “publicly that [he] will not recruit employees of a certain ethnic or racial origin”.³⁷ The reason the employer gave for not recruiting ‘immigrants’ was because its customers were reluctant to give them – as fitters – access to their private residences.³⁸ As the employer’s statement is likely to strongly dissuade certain candidates from submitting their candidature and, consequently, to hinder their access to the labour market, the CJEU found that this constitutes direct discrimination in respect of recruitment within the meaning of Directive 2000/43/EC.³⁹

²⁹ Mittelstad *et al.*, “The ethics of algorithms: Mapping the debate”, 3 *Big Data & Society* 1, 2016, p. 1.

³⁰ Stalder, “Algorithmen, die wir brauchen”, 2017, available at https://netzpolitik.org/2017/algorithmen-die-wir-brauchen.

³¹ Barocas & Selbst, p. 680ff.

³² *Ibid.*, p. 673.

³³ Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information*, HUP, 2015, p. 38.

³⁴ Wang, “When the Customer is King: Employment Discrimination as Customer Service” 23 *Virginia Journal of Social Policy & the Law* 250, 2016, p. 265.

³⁵ *Ibid.*, p. 250.

³⁶ *Ibid.*, p. 262.

³⁷ Case C-54/07, *Feryn*, ECLI:EU:C:2008:397.

³⁸ *Ibid.*, para 16.

³⁹ *Ibid.*, para 25.

In addition, the CJEU ruled that for the assessment whether someone has been discriminated, it is irrelevant whether there has been an identifiable victim.⁴⁰ A lack of an identifiable complainant cannot lead to the conclusion that there is no direct discrimination.⁴¹ Furthermore, in *Bouagnaoui*, in which the customer requested the employer “that there should be ‘no veil next time’”, it was ruled that a customer’s preference could not be regarded as a genuine and determining occupational requirement justifying a discrimination.⁴² AG Sharpston emphasised that a “customer’s attitude may itself be indicative of prejudice based on one of the ‘prohibited factors’, such as religion”, and therefore “it seems [...] particularly dangerous to excuse the employer from compliance with an equal treatment requirement in order to pander to that prejudice”.⁴³ She further stresses that Directive 2000/78/EC confers protection in employment against adverse treatment based on one of the prohibited factors (religion or belief, disability, age or sexual orientation, Art. 1). “It is not about losing one’s job in order to help the employer’s profit line.”

From the few cases at EU level, it can be derived that a customer’s wish or preference seems to be irrelevant for establishing whether it was the employer who directly or indirectly discriminates his workers. Indeed, as AG Kokott in her opinion in the *Achbita* case wrote, it certainly may be the case that an undertaking takes, even must take, into consideration the wishes of third parties. Otherwise, it would be unable to sustain its presence in the market. Nevertheless, an undertaking cannot “pander blindly and uncritically to each and every demand and desire expressed by a third party”.⁴⁴ Thus, it is still the employer who relies on a customer’s wish and by doing that has the power to discriminate a worker because of his or her sex, religion, or age. Based on the few cases, under EU law, the discrimination by a customer’s rating is attributable to the employer and compliance with EU non-discrimination laws can be enforced against him.

A difficulty to establish a discriminatory decision or behaviour may exist where the customer rating does not criticise the worker on grounds that are protected by EU law. Existing directives protect individuals based on the grounds of racial or ethnic origin (2000/43/EC), religion or belief, disability, age or sexual orientation (2000/78/EC) and sex (2006/56/EC). A much wider range of potential grounds on which someone can be discriminated can be found in Art. 21 EUCFR, referring to colour, social origin, genetic features, languages, political and any other opinion, membership of a national minority, property, and birth. Relevant in relation to customer ratings through digital platforms seems to be that it still is the business that treats individuals differently if it relies on its customer’s prejudices or biases. Other than Directive 2000/78/EC and Directive 2006/54/EC, Directive 2000/43/EC is not restricted to discrimination in situations of employment and occupation, including working conditions such as pay (Art. 157 TFEU) and dismissal, and has thus a much broader scope.

⁴⁰ *Ibid*, para 23.

⁴¹ *Ibid*, para 25.

⁴² According to the CJEU, a genuine and determining occupation requirement “cannot, however, cover subjective considerations, such as the willingness of the employer to take account of the particular wishes of the customer” (para 40).

⁴³ Para 133.

⁴⁴ Para 90.

Customer ratings might, to use an example mentioned by Zarsky, be implicitly discriminatory, that is discrimination through masking (making discrimination undetectable, or defensible)⁴⁵, subconscious discriminatory motivations (long learned biases) and relying upon tainted datasets or tools (datasets or data collection models which systematically discriminate, e.g. due to relying on workers' previous achievements, overrepresented negative information).⁴⁶ While EU law only knows two types of discrimination, the question is whether this concerns direct or indirect discrimination.

An individual who thinks that he has been discriminated against, needs to establish relevant facts. According to the CJEU in the *Feryn* case, where there are facts from which it may be presumed that there has been discrimination, the defendant has to prove that there has been no violation of the principle of non-discrimination.⁴⁷ Statements made by an employer that it will not recruit workers of a certain racial or ethnic origin are such facts. Discrimination has the reputation of being particularly hard to substantiate and perhaps even more so in employment situations. Aware of this problem, the EU legislature has adopted measures to assist applicants claiming to be victims of discrimination on the grounds of, for instance, sex, age or origin. Nevertheless, the shift in the burden of proof does not go so far as to uphold its complete reversal. Employers enjoy a long-standing freedom to recruit the people of their choice, which must not be completely disregarded.⁴⁸ The shift of the burden of proof means that it is then for the employer to present sufficient evidence to prove that it has not breached the equal treatment principle, by showing, for instance, that the actual recruitment practices do not correspond to the statements at stake.⁴⁹

In practice, even though the burden of proof can shift to the employer, it still is, in some cases, difficult for an individual to establish enough facts to trigger the reversal.

Establishing the relevant facts can be problematic, especially where this would involve (sensitive) information of third parties, such as other applicants or co-workers. Here, the *Kelly* case is of interest, where an unsuccessful applicant believed that his application was not accepted because of an infringement of the principle of equal treatment. According to the CJEU, the applicant is not entitled to information held by the course provider on the qualifications of the other applicants for the course in question, in order that he may establish “facts from which it may be presumed that there has been direct or indirect discrimination”.⁵⁰ Thus, it cannot be excluded that a refusal of disclosure by the defendant, in the context of establishing such facts, could risk compromising the achievement of the objective pursued by that directive and thus depriving that provision in particular of its effectiveness.⁵¹

Furthermore, in *Accept*, the CJEU ruled that a *prima facie* case of discrimination on grounds of sexual orientation may be refuted with evidence, including a reaction by the defendant distancing itself from public statements on which the appearance of discrimination is based

⁴⁵ Barocas & Selbst, p. 692.

⁴⁶ Zarsky, *supra* note 11, p. 1389.

⁴⁷ Case C-54/07 *Feryn* ECLI:EU:C:2008:397, para 30.

⁴⁸ Advocate General, Case C-415/10 *Galina Meister* ECLI:EU:C:2012:217. Cf. Art. 16 EUCFR.

⁴⁹ Case C-54/07 *Feryn* ECLI:EU:C:2008:397, para 32. See Ellis and Watson, *EU Anti-discrimination law*, Oxford University Press, 2014, pp. 162-163.

⁵⁰ Case C-104/10 *Kelly* ECLI:EU:C:2011:506.

⁵¹ *Ibid*, paras 34, 38.

and the existence of express provisions concerning its recruitment policy. However, it is unnecessary for a defendant to prove that persons of a particular sexual orientation have been recruited in the past, as this would interfere with their right to privacy.⁵²

These cases, where an individual assumes he has been discriminated against, show that there might be difficulties for individuals to get access to their ratings to (try to) prove the alleged discrimination, as this might touch upon the right to privacy of third parties involved, among which customers and co-workers. Nevertheless, in a case where a worker has been dismissed, in a number of EU Member States the employer will have to motivate its decision, unless the worker's fixed-term contract automatically ends without being renewed. Collective agreements may specify the procedures for workers' ratings, providing the necessary information based on which the worker can know what kind of information will be included in the employer's decision-making. Moreover, data could be anonymised so that no personal data will be disclosed.

3. Data protection challenges of customer ratings

The system of customers' ratings is currently implemented by almost every platform as a way to build trust within the community and preserve the attractiveness of the platform.⁵³ How many people used to accept a ride from a stranger before the advent of the sharing economy and its rating mechanisms? The scoring system, usually in the form of the well-known 5-stars icon, has allowed the flourishing of new business activities, providing an additional source of information that traditionally was supplied by professionals through advertising or face-to-face meetings. In principle, ratings may help to reduce the inherent information asymmetry within the parties, but in particular they can potentially promote the overall transparency over the transaction: they are third party reviews, reporting previous experiences with that buyer or provider of a service. In other words, rating systems have codified the word of mouth.⁵⁴ Indeed, users can acquire a privileged amount of information that they would not be able to get from the counterparty, thus overcoming the difficulties, described by behavioural science⁵⁵, related to informed choices.⁵⁶ As Busch underlined: "Considering the limits of human attention, such consolidated ratings can help to mitigate the problem of information overload which could be caused by a large number of confusing and

⁵² Case C-81/12 *Asociația Accept* ECLI:EU:C:2013:275, paras. 57-58.

⁵³ Goldman, "The Regulation of Reputational Information", in Szoka & Marcus (eds.), *The next digital decade: essays on the future of the Internet*, TechFreedom, 2010, available at <https://ssrn.com/abstract=1754628>; Busch, "Crowdsourcing Consumer Confidence: How to Regulate Online Rating and Review Systems in the Collaborative Economy", in De Franceschi (ed.), *European Contract Law and the Digital Single Market*, Intersentia, 2016, available at <https://ssrn.com/abstract=2799489>; Thierer et al., "How the Internet, the Sharing Economy, and Reputational Feedback Mechanisms Solve the 'Lemons Problem'", 70 *University of Miami Law Review* 3, 2016, available at <https://ssrn.com/abstract=2610255>; Botsman, *Who can you trust? How technology brought us together and why it might drive us apart*, Penguin, 2017; Hausemer et al., *Exploratory study of consumer issues in online peer-to-peer platform markets. Final report*, European Commission (EUR 2017.4058 EN), available at http://ec.europa.eu/newsroom/document.cfm?doc_id=45245.

⁵⁴ Dellarocas, "The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms" 49 *Management Science* 10, 2003, 1407-1424.

⁵⁵ Luth, *Behavioural economics in consumer policy: the economic analysis of standard terms in consumer contracts revisited*, Intersentia, 2010.

⁵⁶ Busch, *supra* note 53, p. 12; Stemler, "Feedback Loop Failure: Implications for the Self-Regulation of the Sharing Economy", 18 *Minnesota Journal of Law, Science & Technology*, 2017, available at <https://ssrn.com/abstract=2754768>, p. 683.

contradictory reviews. The use of consolidated ratings thus takes into consideration the problems of bounded attention and bounded rationality and increases the salience (i.e. the cognitive accessibility) of the most important information”.⁵⁷

In addition, in many platforms – typically those of the sharing economy or supposedly so – users rate each other. Hence, scoring systems constitute an incentive for users to follow the rules of the community and to behave accordingly: participants are nudged to act in a way to keep the digital platform a safe and trustworthy place⁵⁸. As Uber states: “Feedback is a two-way street – important to ensuring a high quality experience for the leading internet-enabled platforms, marketplaces and online companies you love. At Uber, riders rate their experience at the end of every trip, and drivers do the same. Uber regularly reviews that feedback and, through this process, we’re able to create and maintain a safe and respectful environment for riders and drivers in more than 200 cities around the world”.⁵⁹

The role of ratings is pivotal because it can influence the choices not only of the users-potential-buyers but of the platform as well. Therefore, it can produce significant consequences on the weakest subjects, such as the providers of the service (e.g. the Uber driver). Indeed, bad scores and comments given the users of the platform may severely affect the digital reputation of the worker and, as a consequence, his/her activity: customers could decide not to accept a service from someone with a low score or, in the worst scenario, workers could be excluded from the platform. For example, if Uber’s drivers have an average score under the threshold of 4.6, their account can be deactivated, which basically means that their contract with the platform is terminated unilaterally.⁶⁰

Hence, despite their potential benefits and the promise towards more transparency, ratings mechanisms are far from being a perfect instrument to assess the quality of a performance: customers may complain fraudulently just to get a refund or a discount or they may give an inaccurate score because of a mistake or a personal bias.⁶¹ Such a problem, which inherently depends on personal behaviours, is not counterbalanced by a mechanism to “review the reviews”.

In our view, the problem with scores is twofold and largely depends on how the platform designs the system.⁶² First, such ratings are highly subjective, but the score expressed in stars (from 1 to 5) give a false sense of objectivity. In other words, it turns a qualitative opinion about an overall experience into a quantitative one, generating a problem of oversimplification or, worst, inaccuracy. Furthermore, because of how the system is built, customers are not often aware of the difference between the score, for example, of 4 or 5 stars. It is true that, in some platforms as Uber, customers are also allowed to leave a

⁵⁷ Busch, *supra* note 53, p. 12. By “consolidated ratings”, the Author refers to the score resulting from the aggregation of different reviews given by the customers and expressed in “stars” or other kinds of indicators.

⁵⁸ Bolton, Greiner, and Ockenfels, “Engineering Trust: Reciprocity in the Production of Reputation Information”, 59 *Management Science*, 2013, 265-285. Stemler, *supra* note 56, p. 683-684.

⁵⁹ <https://www.uber.com/newsroom/feedback-is-a-2-way-street>.

⁶⁰ California Labor Commission, *Uber Technologies v. Bernick*, 2015, available at <https://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=1988&context=historical>.

⁶¹ Rosenblat, Levy, Barocas, and Hwang, *supra* note 5, p. 8 ff.

⁶² For an overview of different types of rating systems, see Pettersen, “Rating mechanisms among participants in sharing economy platforms”, 22 *First Monday* 12, 2017, available at <http://firstmonday.org/ojs/index.php/fm/article/view/7908/6586>.

comment, where he/she can offer details about the reason of the score. However, as previously mentioned, to get 4 instead of 5 stars can really make a difference for an Uber's driver.

Secondly, as shown by Dambrine et al., some platforms in the sharing economy do not facilitate the access to the reputation scores nor provide efficient means to challenge the accuracy of the reviews nor, in some cases, allow the erasure of bad comments and scores.⁶³ This is also the case of Uber, which – on the basis of information available – does not provide any ability to challenge or respond to ratings.⁶⁴ Indeed, the Drivers Privacy Statement currently in force does not offer many details on how to handle bad evaluations: it vaguely states that drivers can “request access to, correction of, or a copy of [their] information by contacting Uber at help.uber.com or visiting our Privacy Policy Help Page”.⁶⁵ However, the help page does not provide any insight about ratings issues and the direct request to the company does not seem to be particularly successful.⁶⁶

Such a situation of opaqueness is likely to conflict with data protection rules⁶⁷. For instance, scores and reviews given by customers and referring to platform workers can be qualified as “personal data”, as they are a typology of information relating to an identified or identifiable individual.⁶⁸ Furthermore, as the Art. 29 Data Protection Working Party (hereinafter “WP29”) has affirmed, the concept of personal data can include not only “objective” information but also “subjective” information, like opinions or assessments.⁶⁹ Therefore, ratings may fall under this notion. As a result, the processing of this information should be guided by the data protection fundamental principles (lawfulness, purpose limitation, data minimisation, notice and consent, etc.) and data subjects shall be entitled to exercise their rights, in particular the rights of access, to rectification and to erasure. With specific reference to the processing in the context of employment, the GDPR leaves a certain amount of discretionary power to the Member States that can provide for more specific rules (Art 88.1, GDPR). In any case, such rules “shall include suitable and specific measures to safeguard the data subject's human dignity, legitimate interests and fundamental rights, with particular regard to the transparency of processing, the transfer of personal data within a group of undertakings, or a group of enterprises engaged in a joint economic activity and monitoring systems at the work place” (Art 88.2, GDPR).

In the light of the limitations often enforced by platforms through technological and contractual means, we have identified three main points of frictions with the GDPR.

⁶³ Dambrine *et al.*, “User Reputation: Building Trust and Addressing Privacy Issues in the Sharing Economy”, June 2015, available at https://fpf.org/wp-content/uploads/FPF_SharingEconomySurvey_06_08_15.pdf, p. 9 ff.

⁶⁴ *Ibid.*, p. 11.

⁶⁵ <https://privacy.uber.com/policy/>.

⁶⁶ See here for example: <https://uberpeople.net/threads/can-you-dispute-ratings.74676/>.

⁶⁷ For the purpose of this analysis we will take into consideration the provisions of Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, GDPR) (L 119/1), which will be applicable on May 2018.

⁶⁸ See now Article 4(1), GDPR.

⁶⁹ WP29, *Opinion 4/2007 on the concept of personal data*, WP136, 20 June 2007, p. 6.

1) *Consent*. Informed consent has always been one of the backbones of the European data protection framework and its role has been further emphasised in the GDPR. According to Art. 4(11) GDPR, consent must be: a) freely given; b) specific; c) informed; and d) unambiguous. In other words, it has to be a “clear affirmative act”⁷⁰ and therefore the platform shall implement only opt-in mechanisms (e.g. no pre-ticked boxes or implicit consent). A preliminary problem deals with the highly questionable obtainment of the consent when the worker accepts the Terms of Service and privacy policy of the platform: is this consent actually freely given as required by 7(4), GDPR?⁷¹ In its Opinion on consent, the WP29 affirmed that “freely given” means that the data subject (the worker, in our case) has to be able to “exercise a real choice” and that “there is no risk of deception, intimidation, coercion or significant negative consequences if he/she does not consent”.⁷² As an example of controversial situation, the WP29 refers precisely the employment relationship, where the data subject is under the influence of the data controller (the employer) and he/she is not in the position of refusing the processing especially if it is a condition for the employment.⁷³ In the context at stake, the consent is unlikely to be freely given: platforms’ workers have to accept the rating system if they want to offer their services through the platform and therefore they are obliged to consent to the dissemination of their personal data (scores) which will be accessible to any user of the platform. However, platforms may invoke another legal ground rather than consent, such as the legitimate interest (Art. 6(1)(f) GDPR).⁷⁴

2) *Right to access, rectification, object and erasure*. This quartet of principles is designed to give the data subject the necessary legal tools to control his/her information flow. In this sense, the right of access is a precondition for the others: only if the data subject is aware of what data are processed, by whom and how, he/she can successfully pursue his/her other rights guaranteed under the GDPR⁷⁵. For the purpose of this analysis, some provisions of Art. 15 GDPR results to be particularly interesting. In fact, the data subject is entitled to receive from the data controller the following information: 1) the existence of the right to request from the controller rectification or erasure of personal data or restriction of processing of personal data concerning the data subject or to object to such processing; 2) where the personal data are not collected from the data subject (which will be the case of rating assigned by customers), any available information as to their source; 3) the existence of automated decision-making, including profiling, 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.

First of all, as several studies indicate, the ability to access is often poorly implemented in the

⁷⁰ Recital 32 GDPR.

⁷¹ According to Art. 7(4), GDPR “when assessing whether consent is freely given, utmost account shall be taken of whether, *inter alia*, the performance of a contract, including the provision of a service, is conditional on consent to the processing of personal data that is not necessary for the performance of that contract”.

⁷² WP29, *Opinion 15/2011 on the definition of consent*, WP187, 13 July 2011, p. 12. See now, WP29, *Guidelines on Consent under Regulation 2016/679* (WP259), 28 November 2017.

⁷³ WP29, *Opinion 8/2001 on the processing of personal data in the employment context* (WP48), 13 September 2001. See now, WP29, *Opinion 2/2017 on data processing at work* (WP249), 8 June 2017, p. 6-7.

⁷⁴ On the concept of legitimate interest, see WP29, *Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC* (WP217), 9 April 2014.

⁷⁵ Ausloos & Dewitte, “Shattering One-Way Mirrors. Data Subject Access Rights in Practice”, in *International Data Privacy Law* (Forthcoming), available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3106632.

online environment and filing an access requests may be result in a task far from being trivial.⁷⁶ In addition, as already noticed, there are no well-developed tools that allow an effective and timely rectification of inaccurate data (see Art. 16, GDPR) nor their cancellation (Art. 17 GDPR)⁷⁷.

Secondly, the data controller shall communicate to the data subject the source of information: however, ratings given by customers are usually anonymised and the worker of the platform is therefore unable to defend himself from the author of the bad review. Indeed, the right of access may be limited if it can affect the rights and liberties of other individuals (recital 63 GDPR), such as for example the right to privacy of other co-workers or users of the platforms (i.e., ratings are also data referring to the customer who gave them).⁷⁸ Furthermore, the scope of application of the right to access is limited to data subjects' personal data: this means that the right cannot be exercised in order to get access, for example, to information (such as ratings and evaluations) related to the "colleagues".⁷⁹ This could constitute a serious impediment if the worker wants to discover and demonstrate whether he/she was subject to any discriminatory behaviour by platforms or by other users.

Finally, the GDPR has introduced a specific provision which aims at fostering transparency and some authors have welcomed this innovation as the "right of explanation" of the algorithm logic.⁸⁰ This leads us to the third point of friction with the GDPR.

3) *The right not to be subject to automated individual decision making.* The new Art. 22 expressly recognises that the data subject has the right to not be subject to a decision which is based only on an automated processing of his/her data, such as profiling. Furthermore, the data subject shall receive "meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing" (Art. 14(2)(g) and Art. 15(1)(h) GDPR). This means that any kind of profiling directed to measure data subject's performance at work (and that can result in his/her exclusion from the platform) cannot be based only on the aggregation of different scores given by customers without the possibility of any human intervention.⁸¹ The data subject in particular shall have the right to receive "specific information [...] express his/her own point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision" (recital 71 GDPR). However, reading *a contrario* such provisions, if the processing is semi-automated – or not

⁷⁶ Ibidem.

⁷⁷ Dambrine *et al.*, *supra* note 63.

⁷⁸ See, for instance, the previous paragraph.

⁷⁹ As affirmed by the Italian Data Protection Authority, who denied the possibility of some employees of a local health unit, which were involved in a procedure for the economic recognition of personal productivity, to get access to *note di qualifica* ("assessment of performance") and evaluations related to other concurrent employees. See, Italian Data Protection Authority, decision of 6 February 2001, available at: <http://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/39236> [Italian only].

⁸⁰ Goodman & Flaxmax, "European Union regulations on algorithmic decision-making and a 'right to explanation'", 38 *AI Magazine* 3, 2016; Selbst & Powles, "Meaningful Information and the Right to Explanation", 7 *International Data Privacy Law* 4, 2017, p. 233-242, available at <https://ssrn.com/abstract=3039125>. *Contra*, Wachter, Mittelstadt, and Floridi, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation", *International Data Privacy Law*, 2017, available at <https://ssrn.com/abstract=2903469>.

⁸¹ Even if we agree that the automated decision is necessary for entering into a contract between the data subject and the data controller, the latter shall implement suitable measure to protect the data subject's right, "at least, the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision" (Art. 22(3) GDPR).

automated at all - the data subject cannot enjoy such rights. Hence, if the platform demonstrates that the exclusion of a user from the system is partially based on customers ratings but the final decision is taken by a person (e.g. from HR), the worker could only rely on Art. 16, GDPR and the problems already mentioned. However, in its recent guidelines on automated individual decision making and profiling, the WP29 has clarified that: “the controller cannot avoid the Art. 22 provisions by fabricating human involvement. For example, if someone routinely applies automatically generated profiles to individuals without any actual influence on the result, this would still be a decision based solely on automated processing. To qualify as human intervention, the controller must ensure that any oversight of the decision is meaningful, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision. As part of the analysis, they should consider all the available input and output data”.⁸² It is worth noting that the mentioned “token gesture” is usually what happens in most of the cases related to the gig-economy, where the worker is excluded from the platform if his/her ratings tragically falls under a predetermined threshold.

4. Obstacles and pathways

As we stated in our Introduction, we argue that an analogy can be drawn between the decision of the CJEU in *Boungaoui*, notably the part concerning customer preferences, and the use of customer ratings in employment-related decisions. However, it should be noted that in *Boungaoui* the Court of Justice only denied that customer preferences could represent a genuine occupational requirement.⁸³ This is only relevant if a *direct* discrimination is alleged.⁸⁴ Similarly, as we saw before, pandering to (alleged) customers’ preferences could not excuse a direct discrimination, as it was the case in *Feryn*. It is by no means clear that such a reasoning would be extended in a case of indirect discrimination. Indeed, an indirect discrimination only needs to be objectively justified by a legitimate aim, while the means of achieving that aim should be appropriate and necessary⁸⁵. This warrants a particularly cautious approach to the applicability of this case law to the phenomenon here at stake.

Considering that the appearance of neutrality is a characteristic of the “scored society”⁸⁶, it is indeed likely that judges might assess the situation as one of indirect discrimination. As stated by Morozov: “Here we run into the perennial problem of algorithms: their presumed objectivity and quite real lack of transparency”.⁸⁷ A *prima facie* case of indirect discrimination needs to be demonstrated through significant statistical data showing the adverse impact of the measure on persons characterised by one of the protected grounds. However, the ability of the individual worker to have access to data related to customer ratings runs against important obstacles, as these ratings are related to *other* workers, and as such are beyond the scope of data to which he or she can demand access. Furthermore, to provide a meaningful comparison, and hence statistical proof of discrimination, the worker would need to receive

⁸² WP29, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679* (WP251), 3 October 2017, p. 10.

⁸³ *Boungaoui*, para. 40-41.

⁸⁴ Directive 2000/43/EC, Art. 4 ; Directive 2000/78/EC, Art. 4 ; Directive 2006/54/EC, Art. 14(2).

⁸⁵ Directive 2000/43/EC, Art. 2(b) ; Directive 2000/78/EC, Art. 2(b) ; Directive 2006/54/EC, Art. 2(b).

⁸⁶ Pasquale, *supra* note 13, p. 35.

⁸⁷ Morozov, *To save everything click here*, Allen Lane, 2013, p. 248.

data related to ratings of individuals characterised by the same protected criterion, hence covering an information which the employer might not be entitled to collect in first place⁸⁸.

If such an obstacle can indeed be overcome, and a *prima facie* case of indirect discrimination is hence established, it is arguable that the employer would not be able to hide behind the black box of customer ratings to justify its decision. This would in fact make it impossible the application of non-discrimination law by mere obfuscation, something which has been ruled out by the Court of Justice in *Danfoss*⁸⁹. The question would then turn in one of justification.

The precedents of *Bougnaoui* and *Achbita*⁹⁰ have shown that the Court is willing to accept employers' organisational choices as legitimate aims, on the basis of the freedom to conduct a business enshrined in Article 16 of the EU Charter of Fundamental Rights. In these cases, the Court has found the wish of the employer to "project an image of neutrality towards the customers" to be a legitimate objective. Thus, as highlighted by Vickers, the fact of responding to customers' preferences, which could not be used to justify a direct discrimination, was instead deemed a legitimate objective in the context of an indirect one.⁹¹ Furthermore, the Court did not in any way consider whether the choice in favour of such an image was in any way necessary for the business at stake.

One can only wonder whether a similar fate would await the choice of using customer ratings for assessing the performance of workers. After all, Uber stresses the importance of customer ratings to keep "a top tier service for riders".⁹² The wish to project a certain image seems sufficiently broad a concept to be employed also in this context, as an "image of neutrality" is not per se distinguishable from any other image an employer might be willing to project -apart from those which might be explicitly illegal. An "image of neutrality" does not appear to be more legally protected than, for instance, an image of punctuality, cleanliness, friendliness and so forth, all of which could justify the decision to use customer ratings to assess workers' performances.

The fact that customer ratings are only used to assess customer-facing workers might also fulfil the requirement of proportionality, if one follows the reasoning of the Court in *Achbita*, where it stated: "[...] what must be ascertained is whether the prohibition on the visible wearing of any sign or clothing capable of being associated with a religious faith or a political or philosophical belief covers only G4S workers who interact with customers. If that is the case, the prohibition must be considered strictly necessary for the purpose of achieving the aim pursued".⁹³ Moreover, the distributed nature of some services prone to use customer

⁸⁸ Ellis and Watson, *supra* note 49, p. 155; Zarsky, *supra* note 11, p. 1403

⁸⁹ CJEU, C-109/88, *Handels- og Kontorfunktionærernes Forbund i Danmark v Dansk Arbejdsgiverforening, agissant pour Danfoss*, ECLI:EU:C:1989:383, para. 13: "[...] in a situation where a system of individual pay supplements which is completely lacking in transparency is at issue, female employees can establish differences only so far as average pay is concerned. They would be deprived of any effective means of enforcing the principle of equal pay before the national courts if the effect of adducing such evidence was not to impose upon the employer the burden of proving that his practice in the matter of wages is not in fact discriminatory". See also CJEU, C-318/86, *Commission v France*, ECLI:EU:C:1988:352, para. 26-27.

⁹⁰ Case C-157/15, *Achbita*, ECLI:EU:C:2017:203.

⁹¹ Vickers, *Achbita* and *Bougnaoui*: One Step Forward and Two Steps Back for Religious Diversity in the Workplace (2017), 8 *European Labour Law Journal* 3, 1-27, p. 20.

⁹² Uber's Welcome Packet ("What Uber looks for"), quoted in *Uber BV, Uber London Limited and Uber Britannia v Aslam & Others*, UKEAT/0056/17/DA.

⁹³ *Achbita*, para. 42.

ratings (from the gig-economy to call centres) might also suggest that using this form of control is indeed necessary for maintaining a control over the quality of the service provided, and hence to ensure the desire *image* of the company.

Rosenblat, Levy, Barocas and Solon, analysing this issue in the context of US law, conclude that non-discrimination law is ill equipped to fight discriminatory practices grounded in customer ratings⁹⁴. Our conclusion in the context of EU non-discrimination law is presently similar. Even if sufficient data to prove an indirect discrimination can be presented, we argue that an employer, following the reasoning of the CJEU in *Achbita* and *Boungaoui*, would be able to justify the use of customers' ratings on the basis of their equal application to all workers, their importance for the image of the company and their application only to customer-facing workers. This would be more likely the case in the case of work on-demand via mobile applications, for instance, where an Uber driver has been demanded to drive a customer. In the end, EU non-discrimination law would prove more useful to protect against the single employer holding discriminatory views (or pandering to the same views from a specific customer) than against widespread prejudices among customers. We wish to stress that this conclusion is based on a limited number of precedents, so that it is still a possibility that new decisions by the CJEU might go in a different direction.

That being said, we argue that, in light of the obstacles for *ex-post* remedies, which seem difficult to overcome without sacrificing the privacy of other workers, *ex ante* solutions should be explored to preserve the effectiveness of non-discrimination. The insidiousness of the alternative has been highlighted by AG Sharpston in her Opinion in the *Boungaoui* case, when she pointed out that accepting "the argument, 'but we need to do X because otherwise our customers won't like it' [...] [w]here the customer's attitude may itself be indicative of prejudice based on one of the 'prohibited factors'" would be akin to excusing "the employer from compliance with an equal treatment requirement in order to pander to that prejudice".⁹⁵

The data protection framework could offer a set of rights and remedies which may ensure in principle a fairer balance of interests of all the subjects involved in the processing, but they cannot eliminate the risk of discrimination perpetrated by the platform via its customers ratings.⁹⁶ Indeed, as affirmed by the European Data Protection Authority during the Computer, Privacy and Data Protection Conference, held in Brussels in January 2018: "Data protection tools alone will not create a better world. We need to go beyond mere compliance".⁹⁷

Keeping this exhortation in mind, we present a proposal that aims to reduce the risk of discrimination and enhance the transparency, through a more effective IT implementation of data protection principles. The main problem, as we have seen, is that the battering-ram able to demonstrate a discrimination, the right of access, is often limited by its very nature (data subjects can access only their personal data) or by the other competing rights and freedoms. In the light of the principles of data protection by design and data protection by default, now

⁹⁴ See Rosenblat, Levy, Barocas, and Hwang, *supra* note 5, p. 14.

⁹⁵ Opinion of AG Sharpstone in *Boungaoui*, para. 133.

⁹⁶ As pointed out by Rosenblat, Levy, Barocas, and Hwang: "Through a rating system, consumers can directly assert their preferences and biases in ways that companies would be prohibited from doing directly". *Ibid.*, p. 10.

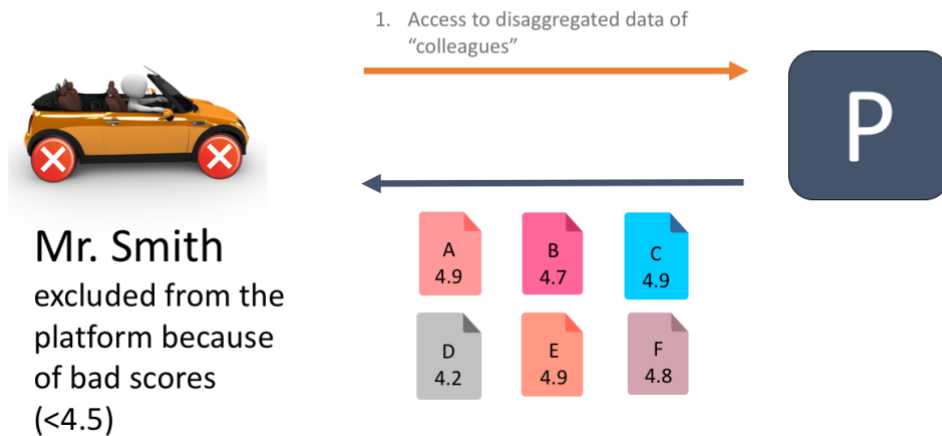
⁹⁷ Buttarelli, *Closing speech at the Computer, Privacy and Data Protection Conference*, Brussels, 25 January 2018.

expressly codified at Art. 25 GDPR, we propose a solution to burst the “rating bubble”, by keeping the privacy of all the subjects and, at the same time, ensuring the effective exercise of the rights recognised under the GDPR. It is a proposal that presuppose the collaboration of the platform: the latter should design the system in order to provide the relevant information to the data subject without disclosing the identity of the customer who left a bad review. First of all, if customers are not satisfied with the service and leave a bad review, they should not be able to save the score without – at least, succinctly - explain the reasons of the complaint. Secondly, if the worker wants to have access to his/her data, he/she has to be entitled to see the consolidated score in a disaggregated form, in order to identify the bad reviews. Since the latter should be annotated with a short comment from the customer, the worker should be able to address them and explain his/her point of view. In addition, he/she shall always be entitled to ask how the consolidated rating has been calculated. Therefore, he/she could ask for the rectification or the erasure of the score from his/her consolidated rating. While this verification is pending, the platform will have to recognise the worker his/her right to the restriction of the processing (Art. 18 GDPR), in order that the bad score will not affect his/her consolidated rating until the issue is clarified.



Source for the icons: Pixabay. CC0 License

In any case, the platform should not be entitled to terminate the contract with the worker without giving to him/her the possibility to challenge the bad score. In addition, if the platform still takes a serious decision against the worker, the latter should be able to have access to the de-identified scores of the “colleagues” in order to challenge the platform’s decision.



Source for the icons: Pixabay, CCO License

A further alternative might be found in systems of auditing for scoring algorithms⁹⁸, which could be transposed in the context of customer ratings. Citron and Pasquale notably propose that scoring systems should be subject to licensing and audit requirements when they enter critical settings like employment, insurance, and health care.⁹⁹ Transposing this in the context of the phenomenon here at stake, employers would have to disclose data related to ratings to independent auditors/independent agencies. These independent actors would then be able to assess the potential biases of the ratings themselves, as well as of the resulting decisions¹⁰⁰, and hence their suitability as a ground on which to base employers' choices. To provide an incentive to allow for the auditing to take place¹⁰¹, one might imagine a shift of the burden of proof for those employers basing their decisions on an un-audited rating system. A similar "stick" could also be used to underpin our previous proposal dealing with access to ratings. Alternatively, to preserve the effectiveness of non-discrimination law, legislators should take steps to ensure that customer ratings are treated as an intrinsically insufficient ground for employment-related decisions. As such, these decisions should *also* be based on grounds *other* than that customer ratings.

⁹⁸ Sandvig, Hamilton, Karahalios & Langbort, "Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms", The 64th Annual Meeting of the International Communication Association, 2014; Pasquale, *supra* note 13, p. 150-151.

⁹⁹ Citron and Pasquale, *supra* note 8, p. 21-22.

¹⁰⁰ Including, for automated decisions, the way in which algorithms using customer ratings are designed.

¹⁰¹ See Zarsky, *supra* note 11, p. 1388.