Main Article

etui.

Transfer
2019, Vol. 25(4) 465–481

DOI: 10.1177/1024258919876416
journals.sagepub.com/home/trs
SAGE

Algorithms, artificial intelligence and automated decisions concerning workers and the risks of discrimination: the necessary collective governance of data protection

Adrián Todolí-Signes
University of Valencia, Spain

Summary
Big data, algorithms and artificial intelligence now allow employers to process information on their employees and potential employees in a far more efficient manner and at a much lower cost than in the past. This makes it possible to profile workers automatically and even allows technology itself to replace human resources personnel in making decisions that have legal effects on employees (recruitment, promotion, dismissals, etc.). This entails great risks of worker discrimination and defencelessness, with workers unaware of the reasons underlying any such decision. This article analyses the protections established in the EU General Data Protection Regulation (GDPR) for safeguarding employees against discrimination. One of the main conclusions that can be drawn is that, in the face of the inadequacy of the GDPR in the field of labour relations, there is a need for the collective governance of workplace data protection, requiring the participation of workers' representatives in establishing safeguards.

Résumé
Les « big data », ou mégadonnées, les algorithmes et l'intelligence artificielle permettent à l'heure actuelle aux employeurs de traiter les informations relatives à leurs salariés et à leurs salariés potentiels de manière bien plus efficace et à un coût bien moindre que par le passé. Ils peuvent ainsi profiler automatiquement les travailleurs et il est même possible que la technologie puisse remplacer elle-même le personnel des ressources humaines pour prendre des décisions qui ont un impact juridique sur les salariés (recrutement, promotion, licenciement, etc.). Ce phénomène entraîne des risques considérables de discrimination et de vulnérabilité de la part des travailleurs, ceux-ci ne pouvant connaître les raisons qui sous-tendent de telles décisions. Cet article analyse les protections instaurées par le règlement général sur la protection des données (RGPD) de l'UE

Corresponding author:
Adrián Todolí-Signes, Faculty of Law, University of Valencia, Av. Tarongers s/n, 46022 Valencia, Spain.
Email: adrian.todoli@uv.es

pour protéger les travailleurs contre la discrimination. L'une des principales conclusions que l'on peut tirer est que, face au caractère inadéquat du RGPD dans le domaine des relations du travail, il existe un besoin de gouvernance collective de la protection des données sur le lieu de travail, qui exige la participation des représentants des travailleurs dans la sauvegarde des intérêts des travailleurs.

Zusammenfassung

Big Data, Algorithmen und künstliche Intelligenz ermöglichen es Arbeitgebern heutzutage, Informationen über ihr Personal und ihre potenziellen Mitarbeiter deutlich effizienter und zu weitaus geringeren Kosten als bisher zu verarbeiten. Damit wird nicht nur eine automatische Profilerstellung von Arbeitnehmern möglich; diese Technologie kann sogar Mitarbeiter im Personalwesen ersetzen bei Entscheidungen mit direkten rechtlichen Folgen für Beschäftigte (Personaleinstellung, Beförderungen, Entlassungen usw.). Dies beinhaltet beträchtliche Risiken dafür, dass Arbeitnehmer diskriminiert werden und sich nicht wehren können, da sie über die Gründe für solche Entscheidungen im Unklaren gelassen werden. Der vorliegende Artikel analysiert die Schutzmechanismen, die in der EU-Datenschutzgrundverordnung (DSGVO) festgelegt wurden und Arbeitnehmer vor Diskriminierung schützen sollen. Eine der wichtigsten Schlussfolgerungen lautet, dass angesichts der Unzulänglichkeiten der DSGVO im Bereich der Arbeitsbeziehungen kollektive Entscheidungs- und Steuerungsmechanismen (*collective governance*) für den Schutz von Arbeitsplatzdaten benötigt werden, welche die Mitwirkung von Arbeitnehmervertretern bei der Festlegung von Schutzmaßnahmen erfordern.

Keywords

Data protection, big data, profiling, algorithm-based surveillance, automated decisions, discrimination, worker surveillance and monitoring, management power, algorithms, artificial intelligence, machine learning

The introduction of technology: a change of paradigm

From automatic data handling to automated processing

In the information age, much of the work done by human resources' (HR) experts consists of gathering as much information as possible about workers in order to improve decision-making (recruitment, promotion, dismissals, part-time/full-time contract, geographical mobility, payment of bonuses, etc.) (Grensing-Pophal, 2009: 42; Sameen and Cornelius, 2013). Indeed, it is crucial for a company to gather and compile as much information as possible in order to gain a thorough understanding of a worker's skills, knowledge, capabilities, attitudes, etc. so as to be able to make decisions best suiting its interests (Jackson Lewis, 2016).

In the same way that a company wants to know as much as possible about consumers in order to know what product to offer them or what exact advertisement will convince them to buy its products, employers want to gather as much information as they can about their (potential) employees in order to know whether they will be productive, how well they will fit into the company's environment or what in particular will motivate them to stay or to work harder (Ajunwa et al., 2017; De Stefano, 2018; Moore et al., 2018b).

The common denominator of these situations is the collection of information (by means of online reputations through customer evaluation, wearables, video cameras, etc.). Gathering information to make better decisions is nothing new.[1] Indeed, for years, companies have been using interviews in selection processes, group dynamics, performance evaluations, etc. to make such work-related decisions. In recent years, however, HR experts have concentrated on gathering information via new technologies (Facebook, Linkedin and now online reputation) (Ouridi et al., 2016: 240–249; Wolf et al., 2014). While it is of course true that technology potentially allows companies to access larger amounts of data in a very economical way (Daws, 2016), an HR manager was always the person who, once that information had been gathered, had to process it and arrive at a decision. This meant that, despite the increasing amount of information available, there was a natural limit to using that information, namely, the human capacity to process such data.

However, the latest technologies are changing even this. Thanks to algorithms, big data and artificial intelligence, not only is there a reduction in the cost of access to information (that which until now was available thanks to Facebook and Linkedin and other public data), but there is also an unprecedented reduction in the cost of processing this information to make it useful, in turn facilitating decision-making based on this information (automated decisions).

The use of new technologies to assess and monitor workers has thus fundamentally changed i) how information is collected and from which sources; ii) how that information is processed; and iii) how decisions are made.

i. *Increase in the amount of information available.* Technologies such as video surveillance, GPS or wearables (e.g. bracelets that monitor a worker's heart rate and his or her attention and activity status) are leading to an increase in the amount of information available.

Likewise, digital reputation systems (customer ratings) make it possible to obtain information about employees' behaviour in a much cheaper way (Thierer et al., 2015: 7). Employers are even beginning to measure workers' emotions (Moore, 2018a: 18).

ii. *Increase in the capacity to process that information*. All this information needs to be processed. Here again, new technologies represent an important step forward in the capacity to carry out this action. In the case of video surveillance, for example, until now a supervisor had to spend hours looking at video surveillance footage to check whether a worker had committed any kind of irregularity. Face and shape recognition systems now allow the automated signalling of any irregularity, reporting it immediately it occurs. This has considerably lowered the cost of monitoring workers.

In the case of wearables, having an HR employee monitoring the heart rate of all workers (or their location if done by means of GPS) would be excessively expensive and, therefore, impracticable. However, by means of automated systems (and algorithms) it is possible, and very inexpensive, to set up alarms informing an HR manager when a worker is inactive for a long period. This means that there is no longer any need for an HR manager to monitor the information or perform surveillance tasks. Instead the manager will simply be automatically 'alerted' when a situation warrants observation.

---

1 For the history of worker surveillance and monitoring methods from the beginning of industrialisation to the present day, see Ajunwa et al. (2017: 107–108).

Some companies in the USA are developing devices fitted with microphones, not with the intention of recording workers' conversations, but to know the worker's mood according to his or her tone of voice. This device can also be used to measure a worker's interactions with colleagues in order to know which of them they interact with and for how long (the Week Staff, 2015).[2]

By the same token, in the case of online reputation, analysing and systematising information and evaluations collected from customers can be excessively costly, while a computerised rating system allows information to be categorised and averages and alerts to be obtained when a worker's behaviour deviates from the acceptable standards. Reducing the need for the interaction of an HR manager obviously gives rise to more economical monitoring methods.

iii. *Capacity for automated decision-making.* The last step in maximising monitoring efficiency, automated decision-making does away with any need for human intervention, with artificial intelligence taking over HR tasks, including decision-making. There are several levels available, ranging from simplified to more complex.

The simplified level basically consists of automating the process in question (promotion, bonus payments or dismissals) by establishing a command in a computer programme (if X happens, perform Y). Hence, it would be possible to develop an automated process such that, if a worker's activity (measured by heart rate) decreases for more than three hours, an email is automatically sent with a letter of dismissal.[3] According to the inspection report of the Labour Inspectorate of Valencia, if a Deliveroo rider is not in motion (detected by GPS), he or she automatically receives a warning message, being told to get moving again[4] (as a 'mental whip' (Moore, 2018a: 23)). Or, for example, if a worker's average online reputation – customer ratings – drops below 4.6 out of 5, he or she is 'automatically' disconnected from the platform[5] (or prevented from entering the workplace by automatically deactivating his or her credentials).

A more complex system would use artificial intelligence (AI). Determined by a firm's programming decisions, true artificial intelligence could, of course, take many more factors into account when making a decision to promote, dismiss, etc. one of the company's employees.

In short, the reduced cost of these three levels allows companies easily to step up the monitoring of workers: the cheaper monitoring is, the more measures an employer will take to protect its legitimate business interests. At present, European labour legislation grants employers the power to choose which forms of worker surveillance and monitoring they deem appropriate. However,

2 More examples in Ajunwa (2019).

3 This can also be performed by time control. In Amazon's logistic centres, the time it takes a warehouse assistant to transport packages from one place to another is monitored by means of a wearable. If it takes him or her longer than stipulated, a notification is sent to warn the assistant. See: https://www.thesun.co.uk/news/6055021/rushed-amazon-warehouse-staff-time-wasting/ (accessed 24 July 2019).

4 In accordance with the message 'sabemos que has recogido el pedido, pero vemos que no te mueves, ponte en movimiento' [we know that you have picked up the order, but we can see that you are not moving, so get going], Spanish Labour Inspection Report no. 460016685/17/sms, dated 5 December 2017. A summary can be found at: https://adriantodoli.com/2017/12/18/comentario-a-la-resolucion-de-la-inspeccion-de-trabajosobre-delivero-son-laborales-y-no-autonomos/ (accessed 24 July 2019).

5 For example, the transport company Lyft has a rule whereby if a driver has an average rating below 4.6 (out of 5) he or she is automatically deactivated. Other decisions are also made; for example, if a user rates a driver with less than a 3, the algorithm will prevent that driver from providing that customer with a service again. In this respect, see: 'We go the extra mile for safety'. Available at: www.lyft.com/safety (accessed 17 April 2018).

these regulations were enacted at a time when surveillance and monitoring were limited by their very nature – in short, because they were expensive.

In contrast to what is commonly believed, technology rarely enables monitoring that was not previously possible in any shape or form. What technology does allow is for monitoring to be performed at a lower cost. Similar to worker monitoring, installing several video cameras in a shop has the same effect as having a greater number of security guards. Video cameras allow the same monitoring to be performed in a more efficient and cheaper way. Hence, with the appearance of these new technologies and the lower cost of monitoring, there is now a need to reconsider this unilateral employer power.

While we have no intention whatsoever of questioning the legitimacy of an employer to be able to monitor the work carried out by its employees, the starting hypothesis of this article is that cheaper monitoring and new methods may give rise to unjustified or abusive interference in workers' fundamental rights and freedoms, in principle justifying legal censorship. In the face of these risks, we advocate the collective governance of workers' data as a way of minimising the discrimination risks and potential violations of fundamental rights at work.

The risks of automated processing: big data and discrimination

Big data not only consist of the accumulation of data and information, but also refer to the set of tools and computer systems (algorithms, machine learning) that analyse these data in search of recurrent patterns and correlations to make predictions (Garriga-Dominguez, 2018: 112; Goñi-Sein, 2017: 16–19). The objective is to profile citizens or workers in order to classify them using parameters introduced within the algorithm itself. The main problem is the possibility of such profiles classifying workers, either directly or indirectly, according to discriminatory categories (Ajunwa et al., 2017; Bodie, 2016; Hildebrandt, 2012). According to many experts, there is an extremely high risk of this occurring.

1. Technology seems capable of inferring certain personal characteristics on the basis of data not immediately related thereto. In other words, even if collecting data on trade union membership, religion, gender, sexual orientation or disability is forbidden, algorithms are capable of deriving this information through other data (Crawford and Schultz, 2014). For example, religion or race can be statistically very closely related to the post code or the district where the person lives. Thus, making decisions based on housing location may ultimately result in a decision based on race (Mittelstandt et al., 2016). Similarly, it is possible to predict political or trade union affiliation according to the time spent reading certain news items on Facebook or Google, and not others. In many cases, the capabilities of an algorithm to make statistical inferences are unknown, meaning that it is 'impossible' to know whether the algorithm itself is making decisions based on discriminatory information or not (Hardt, 2014).
2. In addition, the very construction of the algorithm requires data biased by discriminatory parameters. The algorithm takes reality as a learning factor when processing data, meaning that the results obtained from these data may perpetuate existing biases in our society. For example, the fact that seven out of 10 Fortune 500 company directors are white men[6] may

6 See: http://fortune.com/2017/06/09/white-men-senior-executives-fortune-500-companies-diversity-data/ (accessed 24 July 2019). In Spain nine of 10 company directors of the IBEX 35 are men. See: https://www.elperiodico.com/es/economia/20170204/espana-mujeres-consejos-administracion-ibex35-2016-5784962 (accessed 24 July 2019).

lead an algorithm to understand that a white man is 'more likely' to fit in better as a director in one of these companies – because this is statistically 'confirmed' by the data it possesses.[7]

3. When an algorithm is in command, minorities will tend to be at a disadvantage. The science of statistics itself accords greater value to decisions made with more available information. As there are always fewer data available on minorities (race, religion, sexual orientation, etc.), this will lead the algorithm to understand that making a decision in favour of a minority group is riskier than making one in favour of a majority group (Hardt, 2014). In other words, to select a candidate from a minority group the algorithm will demand (by default) more qualities, aptitudes, knowledge, etc. than if it selects someone from a majority group, simply due to the fact that it is easier to predict (statistically) the behaviour of a candidate belonging to the latter group.[8]

In short, automated data processing exponentially increases the chances of workers' rights being violated.[9] Regardless of whether a decision is ultimately made by an HR manager or not, the fact that it is based on automated data processing (e.g. the profiling or rating of workers by an algorithm) will increase the likelihood of that decision being discriminatory.[10]

The greater likelihood of discrimination arising from big data, algorithms and AI technology is not exclusive to the employment relationship. In fact, the European legislator (concerned about the impact that the automated processing of data may have on the lives of citizens and consumers) has included some specific protections (Article 22 on 'Automated individual decision-making, including profiling') in the General Data Protection Regulation (EU) 2016/679 (hereinafter GDPR). Thus, despite GDPR's apparent lack of specific provisions or protections for workers/employees, it applies to the employment relationship (Goodman and Flaxman, 2016: 83–88). Consequently, we will analyse the protections enshrined in that regulation, looking at the legal effects of automated processing used by an employer to profile workers or automate decision-making.

However, we should point out that these European-level protections are insufficient in view of the possibilities that today's technology offers to invade workers' private lives and make discriminatory decisions. For this reason, after analysing the current regulation, a call will be made for the intervention of the social partners to establish the necessary protections to prevent violations of workers' fundamental rights (what we call in this article the collective governance of data protection).

---

7 Bear in mind that in this case it is irrelevant whether the correlation is true or not, i.e., although statistically there is a correlation between the male sex and the success of running an Ibex 35 company, this correlation is socially and politically reprehensible, see Edwards and Veale (2017: 28).

8 The same happens if the decision is not made by the algorithm but when the algorithm simply classifies workers and the final decision is made by the human resources manager.

9 For instance, it is said that the Internet implies 'the risk of a multiplier effect of attacks against rights, goods and legal interests', see Perez (2006: 93); Garriga-Dominguez (2018: 109).

10 As pointed out by Ippolita (2012: 106), 'statistics know everything without proving anything, they are apparently scientific evidence of highly ideological assumptions'. Specifically, in the case of selection processes that use this technology, a number of studies have shown that historical stereotypes such as sex, race, religion, sexual orientation and even sexual attractiveness lower the likelihood of being called for a job interview. See Caers and Castelyns (2011: 439); WP29 (2018a: 28).

Specific protections against automated decision-making

Scope of the specific protections: decisions based solely on automated processes

Technological improvements to workforce-related decision-making are not limited to increasing the capacity to accumulate and process that information, but also allow some HR tasks previously done by an HR officer to be completely automated. Indeed, artificial intelligence, or machine learning, now allows fully automated decision-making without any human intervention (or only to a minimal extent). The European legislator, in view of this advance in technology, considers that special and specific safeguards are needed to protect citizens from automated decision-making. In this sense, it is understood that, as the degree of automation rises, not only is there a greater risk of discrimination or bias, but also that the more parts of the process are automated, the more control there will be (as it becomes cheaper).

Thus, when a company's decisions are automated, the protections set out in GDPR Article 22 will be applied. These protections will be set out in the following. However, first of all an analysis is necessary to determine in which cases a decision is made based solely on an automated process.

As technology improves, there will be more possibilities for artificial intelligence to take autonomous action. One can already envisage selection processes in which filters are automatically activated by grades found in academic records or by the university where a degree was obtained, with those applicants not meeting the requirements being screened out and never reaching the head of HR's 'desk' for further analysis. This kind of automatic filtering is in itself an automated decision (denial of employment). Analysing other hypothetical situations, however, may be more complicated. Think, for example, about the dismissal of an employee because his digital reputation is less than 4.6 out of 5 (for further debate about the digital reputation of workers, see Todolí-Signes, 2019). If in this case the dismissal is automated, it will come under the scope of GDPR Article 22. However, the dismissal may not have been enacted directly by the computer system, but by an HR officer receiving an 'alert' when a worker's rating drops below 4.6. Would this latter case be an automated decision?

A literal interpretation of Article 22 would lead us to think that it is not, as it establishes special protections for decisions based solely on automated processes. As some authors have posited (Wachter et al., 2017a: 92), any level of human intervention, however trivial, would render these protections inapplicable. However, the bulk of the literature seems to be inclined to understand that any type of human intervention would not be sufficient to impede the application of these safeguards (Selbst and Powles, 2017; Veale and Edwards, 2017). In fact, when human intervention is limited to *applying* the decision taken by the algorithm, we are still dealing with a decision taken solely by an automated system. Hence, the right to control decisions based on an automated process is upheld, provided that human intervention is limited to applying the decision taken by the computer system without any influence on the outcome.

The following – more philosophical – approach can be taken to support this statement. An algorithm does not have any real will; thus, an algorithm does not take decisions, but produces outcomes. In this regard, it is always a human taking the decisions: either the person programming the algorithm or the one applying the algorithḿs outcome. According to this interpretation, mere

human intervention cannot rule out the application of these protections since, given the necessary human intervention, these would be without content.

In my opinion, only when the head of HR has the authority to change the outcome produced by a computer system through assessing different aspects can it be understood that there is significant human intervention.[11] In this regard, in order to know whether the level of human intervention is 'significant', it will be necessary to assess how often an HR officer takes final decisions in a direction other than that produced by the algorithm or artificial intelligence.

Furthermore, it should be taken into consideration that partial decisions are also subject to these protections if they are automated; i.e., despite the fact that the final decision as to who is hired or who receives a bonus or promotion in a company is made by the head of HR with 'significant' intervention, if use has been made of automated systems that have discarded subjects in order to reduce the number of candidates, a denying automated decision will have been taken on them (WP29, 2017: 23). An automated decision is also understood as one sorting workers into categories or profiles, even though an HR officer subsequently takes the final decision (Edwards and Veale, 2017: 46; WP29, 2018a: 23). Indeed, if the computer system classifies workers into categories (e.g. A, B, C, D or assigns them scores from 1 to 5) and the head of HR subsequently decides to give promotion to those in a particular category, the decision, even if made by a human being, would be based on a previous decision made by a machine, thereby falling within the scope of the GDPR Article 22 on automated decisions.

Prohibiting automated business decision-making

The first form of protection in the case of decisions based solely on automated data processing is to prohibit them (right to object). Article 22 of the GDPR establishes the right for a person 'not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her'. The WP29 (2018a: 23) interprets this right as prohibiting (without any need actively to claim the right) data controllers from making decisions with this automated methodology. Nevertheless, this prohibition is only relative as there are exceptions in GDPR Article 22, namely: 'the decision i) is necessary for entering into, or performance of, a contract between the data subject and a data controller; ii) is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or iii) is based on the data subject's explicit consent.' Given that, in an employment contract, a

11 As stated in WP29 (2017: 21) the company (or data controller) cannot circumvent the restrictions of Article 22 by fabricating human intervention. For it to be understood that there is human intervention that excludes the application of these special restrictions, it is necessary for the person in charge to analyse all the relevant information and for this to be carried out by a subject with the authority and competence required to modify the decision. The WP29 (2017) emphasises in particular that the review of the decision by a human must be significant, otherwise human intervention cannot be understood as excluding the protection of Article 22 of the GDPR.

worker's consent will generally not be valid[12] that means that, in the absence of an internal regulation, the automated decision can only be considered valid if it is *necessary* for the conclusion or performance of the employment contract.

According to the WP29 (2017), in order to understand whether automated decision-making is 'necessary', the data controller must prove that it is the most appropriate way to fulfil the objective of the contract. In fact, the interpretation that can be drawn is that human intervention must be shown to be impractical or impossible because of the amount of data processed, i.e., the company must demonstrate that there are no other less intrusive ways to achieve the same aim. Specifically, the WP29 (2017: 23) establishes that it would be valid to use automated systems in the event that a company, in the case of a job offer, receives tens of thousands of applications. In that case, it would be impracticable for the company to carry out the selection process without first discarding some of the candidates by means of automated processes. However, once the list of candidates has been reduced to manageable numbers, automated decision-making would have to cease.

Accordingly, it would not seem possible to make automated decisions (or decisions based on automatically created profiles) regarding promotion or bonuses and, even less so, concerning dismissals without significant human intervention.[13]

'Right to an explanation'

If one of the exceptions to the prohibition applies (e.g. the need to fulfil the contract), Article 22(3) of the GDPR requires a company, as the 'data controller', to 'implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, 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'.

It appears, under this Article, that the data controller has the obligation to inform the data subject of the reasons that led to that decision, i.e., when a company automates decision-making, it must indicate that it has done so and provide details as to what parameters it has used to reach the decision in question (and what weighting it has assigned to each of them). This interpretation is supported by Article 5 of the GDPR, which requires the processing of personal data to be lawful, fair and transparent, and also by Articles 13(2(f) and 14(2)(g), requiring that when the subject is involved in automated decisions, including profiling, the data controller must provide the subject with 'meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject'.

Indeed, it seems that the doctrine unanimously interprets this as meaning that the data subject has the right to receive information on the automated processing of his or her data in three aspects:

---

12 Recital 43 of the GDPR states that it is not possible to accept the legality of processing data on the basis of consent in a relationship where there is a strong imbalance of power between the parties. Specifically, the Regulation refers to the relationship between Public Administrations and citizens, although such consideration seems perfectly applicable to the employment relationship. In fact, in its Guidance on consent under Regulation 2016/679, the WP29 (WP29, 2018b, 2018a: 8) understands that it is difficult for a worker's consent to be given in compliance with such requirements (specifically 'freely'). For this reason, the Working Group determines that, given the nature of the relationship between employer and employee, as a general rule, the granting of consent by the employee should not be understood as valid and ought only to be accepted as such under exceptional circumstances.

13 This conclusion is supported by the EESC (2018: 43), which has stated that the principle of 'Human in command', based on the idea that there should be a human being with sufficient autonomy and control over artificial intelligence, should be strictly followed at all times.

i) to be informed that he or she is involved in an automated decision-making process, i.e., inform the worker that the HR process will be fully or partially automated; ii) to be provided with meaningful information on the logic of the algorithm, i.e., among other things, indicating the parameters evaluated by the algorithm making the decision and their weighting; and iii) to be informed about the consequences of the process, i.e., what consequences the automatically taken decision will have for the worker, in one sense or another.[14]

Yet, the doctrine is not unanimous in interpreting the necessary extent of the explanation. While a few authors establish that the obligation is limited to requiring an *ex ante* (i.e., prior to the decision being made) and *general* description of the data supplied to the algorithm in order for it to make the decision[15], the majority considers that GDPR Article 22(3), together with Articles 13(2)(f) and 14(2)(g) and Recital 71, requires an *ex post* (i.e., after the decision has been taken) and *specific* explanation of how and why that decision has been taken with regard to that particular worker.[16]

The response given to the controversy is not trivial, as the minority interpretation would fulfil the mandate simply by providing generic information on the three aspects mentioned above, while in the second case it would be necessary to explain how the algorithm processes the data in order to reach its conclusions and how that decision has been reached for the specific data subject.

In my opinion, there are two arguments leading us to opt for the second stance: a literal interpretation and a finalistic interpretation of the precept.

Thus, Articles 13(2)(f) and 14(2)(g) literally require the data controller to provide a data subject with *specific* and easily accessible information on the automated decision-making concerning him or her, including profiling. Furthermore, Recital 71 expressly requires that automated decisions be 'subject to appropriate safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, *to obtain an explanation of the decision reached after such assessment* and to challenge the decision'.

This interpretation is also in line with the objective of this prescription[17], which is to avoid biases in an algorithm over which a data subject has no control and thus no knowledge about the reasons behind a decision. For this reason, it appears that the degree of 'explanation' given by a company regarding the reasons for taking such a decision should be sufficient to enable the data subject to exercise the corresponding rights to '*express his or her point of view and to challenge the decision*'. This implies that a generic explanation would be insufficient to avoid the defencelessness of a worker who is denied employment or dismissed on the basis of an automated decision. As the WP29 (2017: 25) says, the explanation must be sufficiently comprehensible for the data subject to understand the reasons for the decision.

In short, in my view, in those few cases where it is deemed necessary to use automated decision-making procedures, a worker should have the right to receive a *specific explanation after the*

---

14 See Wachter et al. (2017a: 79–90); Mendoza and Bygrave (2017); Edwards and Veale (2017: 18–82); Wan and Routledge (2017); Malgieri and Comandé (2017); Goodman and Flaxman (2017); Selbst and Powles (2017); Wachter et al. (2017b); Gil Gonzalez (2017: 165–179).

15 See Wachter et al. (2017a: 79–90); Wachter et al. (2017b); Gil Gonzalez (2017: 165–179).

16 See Mendoza and Bygrave (2017); Edwards and Veale (2017: 18–82); Wan and Routledge (2017); Malgieri and Comandé (2017); Goodman and Flaxman (2017); Selbst and Powles (2017).

17 The doctrine establishes three objectives that the 'explanation' must fulfil: i) to inform and help the subject understand why a particular decision was made; ii) to allow the subject to challenge an adverse decision; and iii) for the subject to understand what he or she should change in order to receive the desired response in the future, taking into account the current decision system; see Wachter et al. (2017b: 4).

decision had been taken. In addition, this explanation should be *sufficient* to understand the reasons leading to the decision.

The collective exercise of data protection rights

One of the biggest problems with data protection regulations in their application to labour relations is the uttermost lack of collective rights. Indeed, the European regulation has an individualistic character in which rights are granted exclusively to the person concerned without thinking about the possible existence of collective rights. While this may be more or less acceptable when the person concerned is a consumer, it makes little sense when the person concerned is a worker, for two reasons.

The first is the difference in bargaining power between an employer and an employee. This renders it illogical to make the processing of personal data dependent upon consent when it would be difficult for the worker to oppose such processing without the risk of losing his or her job. In other words, there is no real autonomy of will in individual labour relations (Ajunwa et al., 2017: 141).

The second is precisely the fact that, due to this difference in bargaining power, one peculiarity of labour relations is the existence of institutions representing workers and collectively upholding their interests: trade unions. However, although it is already known that the lack of a specific regulation on data protection in labour matters has given rise to a set of regulations that do not fit easily into the current context (Cardona-Rubert, 1994: 83; Fernandez-Villazón, 1994: 510) and that their interpretation and adaptation require a lot of hard work, the lack of consideration of collective rights in favour of trade unions in data protection matters is glaringly obvious.

However, European legislation has made some progress in collective matters compared to the preceding data protection directive. Specifically, Article 80 of the GDPR ('Representation of data subjects') allows the data subject to mandate a non-profit entity, organisation or association to submit a claim or to exercise a data protection right on his or her behalf. This means that any trade union may, on behalf of its members, exercise any of the rights (access, rectification, deletion, opposition, etc.) granted to workers by the Regulation – not only in court, but also before the employer or a national data protection agency.

Towards collective governance of data protection rights by workers' representatives?

While the progress made in this area is welcomed, there is still a lack of real collective 'governance' of workers' data protection rights. While the GDPR lays down a number of very strict obligations on transparency and data protection safeguards – also for workers' data –, such safeguards may always be decided unilaterally by the employer without the GDPR giving trade unions the power collectively to monitor the use of such data by the company within employment relationships. Indeed, it only provides for the exercise of individual rights by a third party. Nonetheless, national implementations of the GDPR stipulations could require that such safeguarding instruments for the protection of workers' data be negotiated with workers' representatives, or that the latter are at least consulted, a provision which could be called the collective governance of data protection. In fact, Article 88 calls for national regulations to establish more rights-based safeguards for the protection of workers' data or for such protection to be provided by collective agreements.

With regard to the protection of consumer data, some authors have stressed the importance of empowering agencies, non-governmental organisations and civil society to ensure that there is no

discrimination or bias in decisions made on the basis of big data (Edwards and Veale, 2017: 23). In other words, the GDPR should not simply allow such decisions to be monitored externally on behalf of individuals, but it should also be possible to analyse the legality and legitimacy of the actions undertaken by companies on citizens' data from the outset. Translated into the framework of labour relations, this, in my opinion, entails the need for workers' representatives and trade unions to have the power not only to exercise certain rights on behalf of workers, but also to be able to verify from within the use made of workers' information and to ensure that, in selection, evaluation and dismissal processes, the information used has been obtained and processed lawfully (what we have called 'data protection governance').

One of the cross-cutting issues in the new data protection legislation is the obligation to establish *ex officio* safeguards to protect data subjects against automated data processing. It imposes an obligation of result (the protection of the right to data protection), leaving it up to the data controller to choose the formula (the methods and the safeguards) to fulfil that objective. This formulation – the unilateral power on the part of the data controller to choose the safeguards – makes sense in the consumer field because of the lack of 'representativeness' of consumer associations. Precisely the lack of interlocutors with consumers seems to prevent the collective governance of consumer data protection.[18] In the world of labour relations, however, it would perhaps make more sense if the methods and safeguards to protect the fundamental right to data protection were agreed with the trade unions. Indeed, as set out in Article 88, the establishment of safeguards for the protection of workers' data may be the subject of collective bargaining.

On the other hand, if nothing is said in collective bargaining about the methods and mechanisms of workers' protection with regard to data protection, the rights to information and access in favour of workers' representatives will continue to exist (Article 27 of the Charter of Fundamental Rights of the European Union).[19] The works council will therefore have the right to be informed and consulted on any business decision affecting the processing of data or any kind of monitoring of workers.[20]

The rights to 'information and access' granted by The Charter of Fundamental Rights of the European Union (CFREU) Article 27 have the precise aim to restrain corporate powers. Thus, in the face of a regulation (GDPR) obliging a company to establish limits to its corporate powers (safeguards for the benefit of workers), it would be pointless to prevent workers' representatives from participating in the establishment of such limits – even in their weakest form of participation, as is the case of 'information and access'. This right to receive information should include access to all information on technical procedures and company use of a worker's data, including algorithm parameters and its consequences, as described above in the 'right to an explanation' section.

However, we consider this form of workers' participation to be insufficient. The possibilities for the surveillance and monitoring of workers have multiplied as a result of the technological

18 As has already been mentioned, some doctrinal sectors advocate empowering NGOs or associations for such defence.

19 These rights to information and access are laid down in the Charter of Fundamental Rights of the European Union (Article 27). However, the CJEU judgment of 15 January 2014 pronounces on the binding force of the rights to information and access as recognised for workers or workers' representatives in Article 27 of the Charter of Fundamental Rights of the European Union, concluding that Article 27 of the CFREU alone cannot be invoked directly in a dispute. A critical review of this judgment can be found in García-Murcia (2014: 91–116). Also in Sciarra (2014: 127–130).

20 Birgillito and Birgillito (2018: 41) argue that this right to be informed comes from ILO Conventions 87 and 98 as the right to receive previous information and consultation to open negotiations for collective bargaining.

capabilities of information collection. For this reason, faced with the implementation or revision of systems for the organisation and monitoring of work in the future, these legal powers of information and access seem far from sufficient to achieve the desired collective governance of data protection.

A proposal for a joint data protection committee

Hence, in view of this technological change, a regulation is proposed to govern all that information about the worker's behaviour, performance, attitudes, personality traits, etc. and the decisions taken by a company based on these vast amounts of information collected and subsequently processed.

In regard to our proposal, we find that, in order to allow the collective governance of data protection, it is necessary to introduce, among a union's statutory powers, certain rights that already exist in the field of occupational risk prevention, such as the right of proposal (Article 11(1) of the 'Framework' Council Directive 89/391/EEC)[21], or the existence of a genuine joint data protection committee (analogous to the existing health and safety committee on occupational risk prevention). Prerogatives that exist to govern workplace health and safety should be implemented in the field of data protection at work too.

Finally, in the absence of a specific regulation in law (Ajunwa, 2017: 102), it will be up to collective bargaining to establish instruments that ensure the participation of workers in the choice of measures and safeguards to be imposed to protect their data.[22] Regulation in this field must remain flexible and highly adaptable to technological change. That is why, in my opinion, it would be appropriate, *de lege ferenda*, to have a basic legislation making it compulsory to negotiate these extremes (the collective governance of these instruments for processing workers' information), thereby allowing for collective negotiation to specify the necessary safeguards and limits to protect workers' fundamental rights (De Stefano, 2018).

Conclusions

1. With technology enabling more and more forms of data and information processing, the GDPR provides a number of protections. Experts warn of the possible ways in which big data and artificial intelligence can discriminate against the subjects concerned when such technologies are used to make decisions that produce legal effects. The primary objective of previous legislation in this area was to protect individuals' privacy. The new regulations not only protect the capacity of data subjects to control the extent to which they want their personal information to be made known, but also add, with greater intensity, the protection of the right to equality and non-discrimination. As technology now allows not only the processing of data, but also the capacity to create profiles of individuals and even for decisions to be made by algorithms instead of human beings, there is growing concern that such profiling and automated decision-making may affect citizens' fundamental rights.
2. This concern appears to be sufficiently justified by the possibilities of technology to infer certain sensitive information. For instance, algorithms are capable of deriving restricted information (trade union membership, political opinions) from other information until now

21 Council Directive of 12 June 1989 on the introduction of measures to encourage improvements in the safety and health of workers at work (89/391/EEC).

22 For some examples in the international spectrum, see Moore (2018a): 34.

deemed harmless. The risks for fundamental rights are obvious. Yet, the GDPR does not choose to prohibit this technology but to allow its use in exchange for obliging those who use it to establish sufficient safeguards to ensure that there is no kind of discrimination. The GDPR therefore places greater emphasis not only on data protection, but also on ensuring that what is done with this data (the profiles created and the decisions taken) is fair and non-discriminatory.

3. The existence of a 'right to an explanation' is crucial to achieving this protection. Given that today's technology is capable of deriving sensitive (discriminatory) information from other harmless information and of making automatic decisions on the basis of such, it is necessary for an employer to explain how a certain decision has been reached and why. The aim is clear: increased decision-making transparency is necessary in a world where technology affords a wider range of cases in which discrimination may occur. In addition, a full explanation of the grounds used by technology to take a particular decision is deemed necessary in order to avoid the person concerned being defenceless and to enable him or her to oppose that decision (or plead whatever he or she considers appropriate).
4. Especially in the world of work where business decisions affect the very physical and physiological health of workers (Jahoda, 1982), such decisions need to be transparent in order to prevent arbitrariness and discrimination. In fact, the importance of this transparency is such that it could be argued that individual rights are not sufficient. This article proposes the collective governance of data protection within a company. Trade unions are in a privileged position to prevent technology from being used to discriminate or to introduce unwanted bias into our society. This fact needs to be highlighted. The GDPR requires a data controller – i.e., in our case, the employer – to ensure, by establishing safeguards, that fair, transparent and non-discriminatory use is made of the information. At the same time, it allows those involved to lodge a complaint *ex post* if the data controller does not comply. However, in my opinion, it would make more sense, in the field of labour relations, if those safeguards were not chosen unilaterally by the employer but jointly – through negotiations – with the unions; in this way, not only would there be an *ex post* monitoring, but also data protection would be established from the outset.
5. A proposal is put forward, *de lege ferenda*, for a specific regulation of the protection of workers' data (Article 88 GDPR) – not the generic one for all citizens that exists now – where, among other things, joint data protection committees would be set up, as well as a right of proposal from trade union representatives to improve existing company data protection safeguards. Whatever the case, as long as the legal regulations do not require it, collective bargaining has wide scope for the collective governance of data protection and the surveillance and monitoring of workers.

Funding

This research has received funding from the Generalitat Valenciana [Valencian Regional Government under the Research project GV/2019/164 "Mi jefe es un algoritmo: Los efectos de la reputación online y la inteligencia artificial en el trabajo en plataformas digitales"].

References

Ajunwa I (2019) Algorithms at work: Productivity monitoring applications and wearable technology as the new data-centric research agenda for employment and labor law. *63 St. Louis University Law Journal 21*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3247286 (accessed 24 July 2019).

Ajunwa I, Crawford K and Schultz J (2017) Limitless worker surveillance. *California Law Review* 105(3): 102–142.

Birgillito G and Birgillito M (2018) Algorithms and ratings: Tools to manage labour relations. Proposals to renegotiate labour conditions for platform drivers. *Labour & Law Issues* 4(2): 25–50.

Bodie MT, Cherry MA, McCormick ML et al. (2016) The Law and policy of People Analytics. Sant Louis U. Legal studies Research Paper 2016-6. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2769980 (accessed 24 July 2019).

Caers R and Castelyns V (2011) Linkedin and Facebook in Belgium: The influences and biases of social network sites in recruitment and selection procedures. *Social Science Computer Review* 29(4): 437–448.

Cardona Rubert MB (1994) Tratamiento automatizado de datos personales del trabajador. *Revista de Trabajo y Seguridad Social* 16/1994: 83–118.

Crawford K and Schultz J (2014) Big data and due process: Towards a framework to redress predictive privacy harms. *Boston College Law Review* 55(1): 93–128.

Daws R (2016) Adopting fitness trackers in businesses saves $1000 per employee. Available at: http://www.wearabletechnology-news.com/news/2016/oct/19/adopting-fitness-trackers-businesses-saves-1000-employee/ (accessed 25 March 2019).

De Stefano V (2018) "Negotiating the Algorithm": Automation, Artificial intelligence and labour protection. *Comparative Labor Law & Policy Journal*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3178233 (accessed 24 July 2019).

Edwards L and Veale M (2017) Slave to the Algorithm? Why a 'Right to an Explanation' Is Probably Not the Remedy You Are Looking For. *Duke Law & Technology Review* 16(1): Available at: https://scholarship.law.duke.edu/cgi/viewcontent.cgi?article=1315&context=dltr (accessed 24 July 2019).

EESC (2018) *Opinion of the European Economic and Social Committee on 'Artificial Intelligence – The consequences of artificial intelligence on the (digital) single market, production, consumption, employment and society'. JO C 288, 31 May 2018*. Brussels: European Economic and Social Committee.

Fernandez Villazón LA (1994) Tratamiento automatizado de datos personales en los procesos de selección de trabajadores. *Relaciones Laborales* 1/1994: 510–538.

García Murcia J (2014) Naturaleza y fuerza vinculante de los Derechos Fundamentales en la Unión Europea: A propósito de las previsiones sobre información y consulta en materia laboral. *Civitas* 52/2014: 91–116.

Garriga-Dominguez A (2018) La elaboración de perfiles y su impacto en los DDFF. Una primera aproximación a su regulación en el RGUE. *Derechos y Libertades* 38/2018: 107–139.

Gil Gonzalez E (2017) Aproximación al estudio de las decisiones automatizadas en el seno del Reglamento General Europeo de Protección de Datos a la luz de las tecnologías big data y de aprendizaje computacional. *Revista española de la Trasparencia* 5/2017: 165–179.

Goñi-Sein JL (2017) Nuevas tecnologías digitales, poderes empresariales y derechos de los trabajadores: Análisis desde la perspectiva del Reglamento Europeo de Protección de Datos de 2016. *Revista de derecho social (RDS)* 78/2017: 15–42.

Goodman B and Flaxman S (2017) European Union regulations on algorithmic decision-making and a "right to explanation". *AI Magazine* 38(3): 50–57. Available at: https://arxiv.org/abs/1606.08813 (accessed 24 July 2019).

Grensing-Pophal L (2009) Recruiting 2.0. *Credit Union Management Magazine*, 1 September 2009.

Hardt M (2014) How big data is unfair. Medium 2014. Available at: https://medium.com/@mrtz/how-big-data-isunfair-9aa544d739de (accessed 24 July 2019).

Hildebrandt M (2012) The dawn of critical transparency right for the profiling era. In: Bus J, Crompton M, Hildebrandt M et al. (eds) *Digital Enlightenment Yearbook*. Amsterdam: IOS Press, pp. 41–56.

Ippolita (2012) *En el acuario de Facebook. El irresistible ascenso del anarco-capitalismo*. Madrid: Enclave de libros.

Jackson Lewis PC (2016) Employee monitoring and workplace privacy law. *American Bar Association, National Symposium on Technology in Labour & Employment Law*. Available at: https://www.americanbar.org/content/dam/aba/events/labor_law/2016/04/tech/papers/monitoring_ella.authcheckdam.pdf (accessed 24 July 2019).

Jahoda M (1982) *Employment and Unemployment: A Social-Psychological Analysis*. Cambridge: Cambridge University Press.

Malgieri G and Comandé G (2017) Why a right to legibility of automated decision-making exists in the general data protection regulation. *International Data Privacy Law* 7(4): 243–265.

Mendoza I and Bygrave LA (2017) The Right Not to Be Subject to Automated Decisions Based on Profiling. In: Synodinou TE, Jougleux P, Markou C et al. (eds) *EU Internet Law: Regulation and Enforcement*. Springer International Publishing, p. 418. Available at: https://papers.ssrn.com/abstract=2964855[https://perma.cc/XV3T-G98W] (accessed 24 July 2019).

Mittelstandt BD, Allo P, Taddeo M et al. (2016) The ethics of algorithms: Mapping the debate. *Big data & Society* 3(2): 1–21.

Moore P (2018a) Digitalisation of work and resistance. In: Moore P, Upchurch M and Whittaker X (eds) *Humans and Machines at Work: Monitoring, Surveillance and Automation in Contemporary Capitalism*. Basingstoke: Palgrave Macmillan, pp. 17–44.

Moore P, Upchurch M and Whittaker X (eds) (2018b) *Humans and Machines at Work: Monitoring, Surveillance and Automation in Contemporary Capitalism*. Basingstoke: Palgrave Macmillan.

Ouridi ME, Ouridi AE, Segers J et al. (2016) Technology adoption in employee recruitment: The case of social media in Central and Eastern Europe. *Computers in Human Behavior* 57(April 2016): 240–249.

Perez (2006) *La tercera generación de derechos humanos*. Navarra: Thomson Aranzadi.

Sameen S and Cornelius S (2013) Social networking sites and hiring: How social media profiles influence hiring decisions. *Journal of Business Studies Quarterly* 1(2): 1–7.

Sciarra S (2014) Association de médiation sociale. El controvertido papel de los principios fundamentales de la UE y el punto de vista de la legislación laboral. *Revista de derecho Social* 68/2014: 127–142.

Selbst AD and Powles J (2017) Meaningful information and the right to explanation. *International Data Privacy Law* 7(4): 233–242.

The Week Staff (2015) *The Rise of Workplace Spying*. 5 July 2015. Available at: http://theweek.com/articles/564263/rise-workplacespying (accessed 23 May 2018).

Thierer A, Koopman C, Hobson A and Kuiper C (2015) How the internet, the sharing economy, and reputational feedback mechanisms solve the "Lemons Problem". Mercatus Working Paper. Available at: https://www.mercatus.org/system/files/Thierer-Lemons-Problem.pdf (accessed 17 April 2018).

Todolí-Signes A (2019) The Evaluation of Workers by Customers as a Method of Control and Monitoring in the Firm: Digital Reputation and Data Protection. *International Labour Review*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3333093 (accessed 24 July 2019).

Veale M and Edwards L (2017) Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling. *Computer Law & Security Review* 34(2): 398–404.

Wachter S, Mittelstadt B and Floridi L (2017a) Why a right to explanation of automated decision-making does not exist in the general data protection regulation. *International Data Privacy Law* 7(2): 76–99.

Wachter S, Mittelstadt B and Russell C (2017b) *Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR*. Available at: https://arxiv.org/ftp/arxiv/papers/1711/1711.00399.pdf (accessed 24 July 2019).

Wan TW and Routledge B (2017) Algorithmic Transparency, a Right to Explanation, and Placing Trust. Squarespace. Available at: https://static1.squarespace.com/static/592ee286d482e908d35b8494/t/59552415579fb30c014cd06c/1498752022120/Algorithmic+transparency%2C+a+right+to+explanation+and+trust+%28TWK%26BR%29.pdf. (accessed 24 July 2019).

Wolf MV, Sims J and Ynag H (2014) Social media utilization in human resource management. *Web based communities and social Media 2014 conference*, Lisbon. Available at: https://www.researchgate.net/publication/269093272_Social_Media_utilization_in_Human_Resource_Management (accessed 24 July 2019).

WP29 (Article 29 Data Protection Working Party) (2017) *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*. Adopted on 3 October 2017. Brussels: European Commission.

WP29 (Article 29 Data Protection Working Party) (2018a) *Guidelines on transparency under Regulation 2016/679*. Revised and Adopted 11 April 2018. Brussels: European Commission.

WP29 (Article 29 Data Protection Working Party) (2018b) *Guidance on consent under the Regulation 2016/679*. Adopted 10 April 2018. Brussels: European Commission.