Discriminating job applicants
through algorithmic decision-making

Miriam Kullmann*

As an increasing number of hiring(-related) decisions have been delegated to algorithms, the author explores the question to what extent EU non-discrimination laws, in combination with the General Data Protection Regulation (EU) 2016/679, provide (unsuccessful) job applicants with sufficient legal means to address any discriminatory or biased automated decision taken by an employer.

1 P.T. Kim, 'Data-driven Discrimination at Work', *William and Mary Law Review* 2017, Issue 3, p. 874.

2 M.A. Cherry, 'People Analytics and Invisible Labour', *Saint Louis University Law Journal* 2016, Issue 1, p. 1.

3 *www.textkernel.com*. Businesses using Textkernel's software, amongst which *Extract!4.0*: Uitvoeringsinstituut Werknemersverzekeringen (UWV), USG People, AkzoNobel, Amsterdam RAI, StepStone, AXA, Manpower, Adecco, Randstad, Rabobank, ING.

4 *www.businessinsider.de/ hirevue-ai-powered-job-interview-platform-2017-8* and *www.hirevue. com*. Interestingly, as of January 2019, also the Public Employment Service Austria (AMS) will use algorithms to rate the chances of unemployed persons regarding finding a new job. See, e.g., *https://derstandard. at/2000089095393/ AMS-bewertet-Arbeitslose-kuenftig-per-Algorithmus*.

1 Workforce analytics: the transparent job applicant

Finding the right employee for a job can be quite cumbersome for an employer. Not only can reading resumes and motivation letters be time-consuming, the employer also must make a decision on whom to invite for a job interview based on the information the job candidate discloses. Employers increasingly make use of *people analytics* (or workforce),[1] using data to quantify and analyse particular traits, experiences, and skills of employees to make and justify hiring decisions, to promote and fire employees, thereby – seemingly – replacing 'the unreliable and often biased "gut instinct" or anecdotal observation'.[2]

> Employers increasingly make use of people analytics to make and justify hiring decisions, to promote and fire employees, thereby – seemingly – replacing 'the unreliable and often biased "gut instinct" or anecdotal observation'

Particular software, such as Textkernel's *Extract!4.0*, allows for so-called curriculum vitae (CV) parsing, replacing manual sifting of CVs.[3] CVs will be imported into the parsing software, which then automatically can extract, store, analyse, sort and search the information provided, ultimately selecting which candidate(s) would be a possible fit for the job opening. It is said not only to benefit the job applicant, whose CV – and nowadays also social media profile(s) replacing or supplementing traditional CVs – are automatically converted and structured into the software's format, thereby also possibly eliminating manual data entry. Employers may also see benefits where the software replaces reading all CVs by humans and any manual conversion of information, while also assisting humans in finding suitable future employees. An even more sophisticated way allowing to 'screen' job applicants, as used by Unilever for instance, is to ask candidates to undertake a video interview using a particular software facilitating evaluation through artificial intelligence without human interference.[4]

Based on the job description and the requirements the candidate is expected to meet, the software will identify suitable candidates

for the job, taking into account *inter alia* personal data (name, contact information, social media profiles, profile picture, age), education (qualifications, training, courses), previous work experience, relevant (soft) skills, references and hobbies. Of course, the ability to quantify and analyse work(-related) data depends on the data being made available by the job candidate in a way that statistical metrics can accurately capture. Analytics are of limited usefulness in case the software is incapable of matching skills and characteristics with the job in question.[5]

Research has shown that selecting job applicants with Chinese, Indian or Pakistani-sounding names, without using any software, were 28% less likely to get an invitation to a job interview than the fake candidates that had English-sounding names, irrespective of having the same qualifications.[6] Although parsing software can be designed to anonymise names or to not filter on the applicants' names, ethnical background and/or other features, that same software can also reinforce existing discrimination or create new forms of bias, especially where some information may serve as a proxy and reflect correlations rather than causal relations.[7] Therefore, non-discrimination law should, as Kim argues, address so-called 'classification bias', referring to a situation where 'employers rely on classification schemes, such as data algorithms, to sort or score workers in ways that worsen inequality or disadvantage along the lines of race, sex, or other protected characteristics.'[8] This 'job-for-data trade-off', i.e. where the job applicant provides personal data to get the job, creates a particular challenge for job applicants who suspect that the recruitment process has not been successful because of alleged discrimination.

> The question is to what extent EU non-discrimination laws, in combination with the GDPR, provide (unsuccessful) job applicants with sufficient legal means to address any discriminatory or biased automated decision by an employee

As CV parsing software takes automated algorithmic decisions that have the potential to instantiate or introduce new biases, the question is to what extent EU non-discrimination laws, in combination with the General Data Protection Regulation (EU) 2016/679 (GDPR),[9] provide (unsuccessful) job applicants with sufficient legal means to address any discriminatory or biased automated decision by an employer. This contribution is structured as follows. First, attention will be paid to algorithmic parsing of CVs (section 2). Having looked at the technicalities of what algorithms are and how they can work, I will then analyse what it means when parsing software uses the job applicant's personal data and which role is played by the GDPR (section 3). The next step will be identifying the job applicant's position under EU non-discrimination laws, identifying the strengths and weaknesses of these laws when it comes to proving discrimination (section 4). This contribution ends with some final thoughts (section 5).

2 Algorithmic CV parsing: the fallacy of neutrality and objectivity

2.1 What is an algorithm?
Work-related decision-taking, such as hiring a new employee, can be 'outsourced' to parsing software, which imports and converts CVs and automatically extracts, stores, analyses, sorts and searches the information submitted, thereby possibly also drawing on other data sources, such as the applicant's social media accounts. Such software consists of algorithms, defined as formally specified sequences of logical operations providing step-by-step instructions for computers to act on data and thus automate decisions.[10] Algorithms are designed to solve a specific problem,[11] in our case: finding the right employee for the job in question.

Suppose a large Amsterdam-based law firm is looking for an experienced lawyer for their employment law practice. The firm's HR department uses CV parsing software. Possible requirements the software may look for could be that the job applicant has the necessary experience in employment law and has excellent (Dutch and English) drafting skills, can offer creative solutions which actually benefit the client, has an independent and entrepreneurial personality and strong communication and social skills at different levels, is ambitious and looking for new (international) challenges to further develop himself. In addition, the software may be designed in a way to search for candidates that are expected to be equally successful as currently employed employees.[12]

5 Cherry 2016, p. 1-2.
6 S. Thomson, 'Here's why you didn't get that job: your name', *World Economic Forum* 23 May 2017, *www.weforum. org/agenda/2017/05/ job-applications-resume-cv-name-discrimination*, with reference to R. Banerjee, J.G. Reitz & P. Oreopoulos, 'Do Large Employers Treat Racial Minorities More Fairly? A New Analysis of Canadian Field Experiment Data', University of Toronto/Munk School of Global Affairs 2017.
7 Kim 2017, p. 865.
8 Kim 2017, p. 865.
9 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 (*PbEU* 2018, L 119).
10 S. Barocas & A.D. Selbst, 'Big Data's Disparate Impact', *California Law Review* 2016, Issue 3, p. 674.
11 M. Kullmann, 'Platform Work, Algorithmic Decision-Making, and EU Gender Equality Law', *International Journal of Comparative Labour Law and Industrial Relations* 2018, Issue 1, p. 9.
12 The criteria mentioned are derived from a vacancy for an associate corporate (*advocaat-medewerker corporate*) at Allen & Overy, Amsterdam, *https:// perma.cc/MVV3-PTFB*.
13 Dutch courts ruling in an employment context have ruled that announcements made on a personal Facebook page do not belong to the private sphere of the
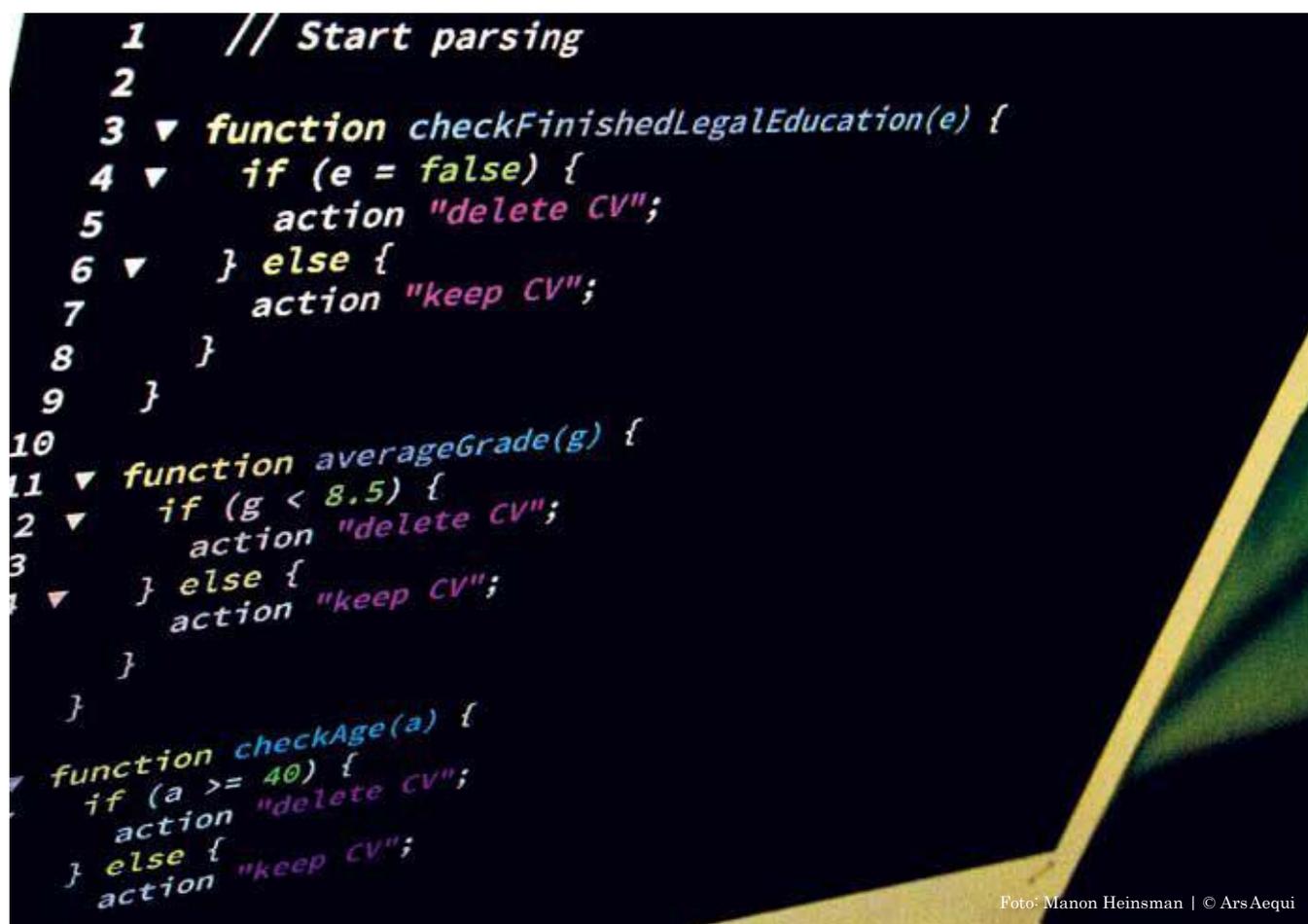
```
1    // Start parsing
2
3  ▼ function checkFinishedLegalEducation(e) {
4  ▼    if (e = false) {
5          action "delete CV";
6  ▼    } else {
7          action "keep CV";
8      }
9  }
10
11 ▼ function averageGrade(g) {
12 ▼    if (g < 8.5) {
         action "delete CV";
13 ▼    } else {
         action "keep CV";
      }
   }

   function checkAge(a) {
      if (a >= 40) {
         action "delete CV";
      } else {
         action "keep CV";
```

Foto: Manon Heinsman | © Ars Aequi

employee. See, e.g., Rb. Arnhem 19 maart 2012, ECLI:NL:RBARN:2012:BV9483 (*Blokker B.V./werknemer*).

14 Derived from: *www.guru99. com/introduction-to-test-management-for-curious-manager.html*. This is, however, only one way in which software can be developed.

15 Barocas & Selbst 2016, p. 677.

16 M.T. Bodie et al., 'The Law and Policy of PeopleAnalytics', *University of Colorado Law Review* 2017, Issue 4, p. 968.

17 C. Cain Miller, 'When Algorithms Discriminate', *New York Times* 9 July 2015, *www. nytimes.com/2015/07/10/upshot/when-algorithms-discriminate.html*; L. Kirchner, 'When Discrimination Is Baked Into Algorithms', *The Atlantic* 6 September 2015, *www.theatlantic.com/business/archive/2015/09/discrimination-algorithms-disparate-impact/403969*; W. Isaac & A. Dixon, 'Why big-data analysis of police activity is inherently biased', *The Conversation* 10 May 2017, *http://theconversation.com/why-big-data-analysis-of-police-activity-is-inherently-biased-72640*.

To submit his application, the candidate would usually need to register either with the law firm's website or, in case a recruitment or staffing agency is hired, a third party website and upload the requested documents. Then the law firm's parsing software will analyse and sort all kinds of data the applicant provided and, if the software is designed in that way, other data sources that should be considered to receive the most complete picture as possible. Thus, two questions arise: which data does the software need and is this 'private' or 'public' data (e.g., easily available data on Facebook, LinkedIn, and other online social media platforms).[13] Some of the more subjective skills might be difficult to be extracted and matched, by the software, from the data provided.

How an algorithmic model is designed to solve a certain problem depends on the type of algorithm used (e.g., deterministic models or 'supervised' or 'unsupervised' self-learning models) as well as the various interests that are to be represented. As far as deterministic models and supervised self-learning algorithms are concerned, there is always a conscious (human) decision that affects the algorithm and the data model it uses to solve the specified problem. Such algorithms can, if desired, be adjusted. This is different with unsupervised self-learning algorithms, where the algorithmic model can adapt itself and determine itself which data it will use. A deterministic algorithm, unlike a supervised or unsupervised self-learning algorithm, decides on the basis of predefined (computational) steps, parameters, and data sources embedded in the algorithm, which means that the algorithm decides on the basis of 'the same input, the same output'.

Which data does the software need and is this 'private' or 'public' data (e.g., easily available data on online social media platforms). Some of the more subjective skills might be difficult to be extracted and matched, by the software, from the data provided

An entire process (see Figure 1 below) precedes the development of an algorithm. It starts with business analysts determining the requirements that are decisive for the problems the software should solve. Software architects design a decision-taking model (the architecture), considering a customer's or company department's demand, based on which the problem the algorithm should solve will be defined. Software developers then will 'translate' these requirements into programming language. Not only does the software need to take into account different taxonomies depending on the client or industry that wants to use the parsing software, but it must also account for spelling mistakes in the CVs as well as be aware of the fact that a particular job may be known under different names. That actually would mean that, besides many other aspects, the algorithms would need to consider much more factors and hence adapt to create new rules that are more flexible to address ongoing changes in the world of job hunting.

The decision-making of deterministic and supervised self-learning algorithms can be adapted and improved. This is, of course, more complicated in the case of self-learning algorithms which, once designed by software architects and implemented by software developers, through 'learning' can independently adapt the model on which the algorithm is based as well as the sources from which the necessary information can be drawn. Ideally, the initial model of the software and the algorithms and all subsequent changes to it are documented in a detailed manner.

2.2 Biased algorithms and data models

Not only can algorithms identify useful patterns in datasets, they also decide based on these patterns, usually much faster and thus more efficiently than a human.[15] Algorithms endeavour to reduce the role of human subjectivity in perception by culling data from more objective means and subjecting that data to examination and statistical analysis.[16] However, it should be emphasised that algorithms are not free of – unconscious – prejudices.[17] At different levels in a company, the person or department that specifies the operational parameters of the algorithms may be biased and that bias may be adopted by software developers, or they may even implement new biases into the algorithm. An algorithm is usually designed in a way that the decision must lead to specified results, with some values and interests preceding over others.[18] As technology is not autonomous, humans have an influential role in the design of algorithmic models.[19]

Moreover, an algorithm can only be as good as the data it works with, meaning that the data model that analyses the algorithm to make decisions can be biased or discriminatory.[20] As Cherry writes: 'As the old saying about data goes, "garbage in, garbage out", meaning that poor quality input will always result in poor quality of result.'[21] Therefore, algorithms can (continue to) disadvantage historically disadvantaged groups[22] if based on negative and unfounded assumptions. An algorithm may harm individuals in four ways: (1) it may intentionally discriminate; (2) an individual's record errors may unfairly

Figure 1. Software development process (waterfall-model)[14]

requirements analysis & definition
business analyst/product management:
determining the needs/conditions to meet for a product

design
software architect:
developing a specification for the algorithm that meets requirements/needs

implementation
software engineers:
writing source code which implements the algorithm

testing
tester:
evaluating an attribute/capability of an algorithm

maintenance & improvement
software engineers:
post delivery service of software

18 B.D. Mittelstad et al., 'The ethics of algorithms: Mapping the debate', *Big Data & Society* 2016, Issue 2, p. 1.

19 F. Stalder, 'Algorithmen, die wir brauchen', *Netzpolitik.org* 15 January 2017, *https://netzpolitik.org/2017/algorithmen-die-wir-brauchen/*.

20 Barocas & Selbst 2016, p. 688; J.A. Kroll et al., 'Accountable Algorithms', *University of Pennsylvania Law Review* 2017, Issue 3, p. 680 ff.

21 Cherry 2016, p. 2.

22 Barocas & Selbst 2016, p. 673; M. Hardt, 'How big data is unfair: Understanding unintended sources of unfairness in data driven decision making', *Medium* 26 September 2014, *https://medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de*.

23 Kullmann 2018, p. 9, with reference to Kim 2017, p. 883-890.

24 Kullmann 2018, p. 12.

25 Kullmann 2018, p. 12, with reference to Kim 2017, p. 917 ff. On privacy rights see: B. van der Sloot, 'How to assess privacy violations in the age of Big Data? Analysing the three different tests developed by the ECtHR and adding for a fourth one', *Information & Communications Technology Law* 2015, Issue 1.

26 Art. 2 Regulation (EU) 2016/679. A 'filing system' refers to 'any structured set of personal data which are accessible according to specific criteria, whether centralised, decentralised or dispersed on a functional or geographical basis' (Art. 4 (6) Regulation (EU) 2016/679).

27 Art. 4 (4) Regulation (EU) 2016/679.

28 The 'controller' is 'the natural or legal person, public authority, agency or other body which, alone or jointly with others, *determines the purposes and means of the processing of personal data*; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law' (emphasis added) (Art. 4 (7) Regulation (EU) 2016/679).

29 The 'processor' is 'a natural or legal person, public
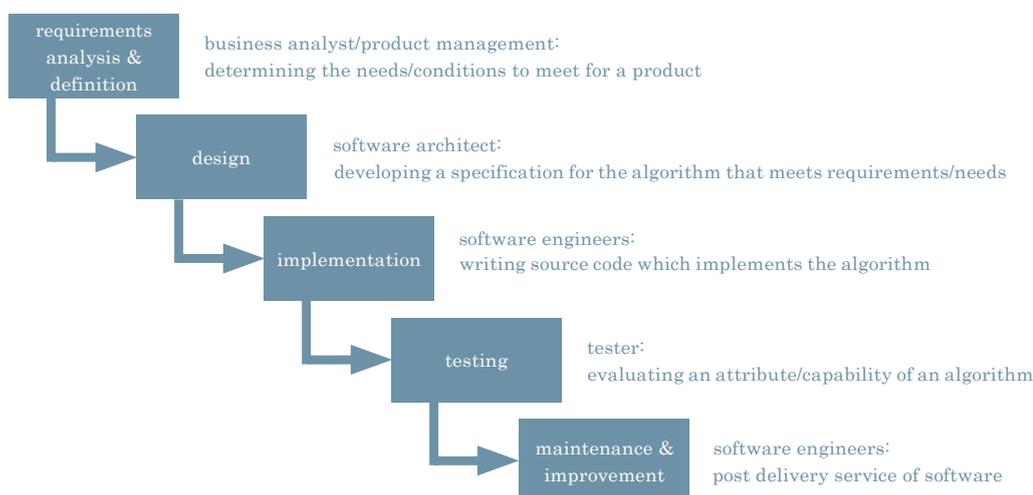
authority, agency or other body which processes personal data on behalf of the controller' (Art. 4 (8) Regulation (EU) 2016/679).

30 Art. 4 (1) Regulation (EU) 2016/679. The criterion of 'identifiability' separates data that is subject to the GDPR from those that fall outside. W. Gedefa Urgessa, 'The Protective Capacity of the Criterion of "Identifiability" under EU Data Protection Law', *European Data Protection Law Review* 2016, Issue 4, p. 521.

31 Recital 4 of the Preamble of Regulation (EU) 2016/679.

32 Art. 5 Regulation (EU) 2016/679.

33 Art. 6 (1) (a) in conjunction with Art. 4 (11) Regulation (EU) 2016/679.

34 Art. 7 (3) Regulation (EU) 2016/679 and Article 29 Data Protection Working Party Opinion 2/2017 on Data Processing at Work, p. 6.

35 Article 29 Data Protection Working Party Guidelines on Consent for the purposes of Regulation 2016/679, p. 18-19.

36 Art. 9 Regulation (EU) 2016/679.

37 Art. 22 Regulation (EU) 2016/679 and Article 29 Data Protection Working Party Guidelines on Automated Individual Decision-Making and Profiling for the purposes of Regulation 2016/679, p. 22.

38 The Article 29 Working Party (also referred to as WP29) anticipates the future work of the European Data Protection Board (EDPB) in issuing guidelines, recommendations and best practices in order to encourage a consistent application of the GDPR (Art. 70(1)(e) Regulation (EU) 2016/679).

39 Consent will not be considered to be freely given where the data subject: (1) has no real choice; (2) feels compelled to consent; (3) will endure negative consequences if he does not consent; (4) is subject to consent which bundled up as a non-negotiable part of terms and condition; or (5) is unable to refuse or withdraw his consent without detriment. See Article 29 Data Protection Working Party Guidelines on Consent for the purposes of Regulation 2016/679, p. 6.

deprive someone of a job opportunity; (3) a statistically biased data model may systemically disfavour a particular group because of the way it was created; and (4) a model may systematically operate to the detriment of members of a particular group, even if the model itself is not biased.[23]

An algorithm can only be as good as the data it works with, meaning that the data model that analyses the algorithm to make decisions can be biased or discriminatory

A not insignificant difference between a human and an algorithm when making decisions lies in the fact that an algorithm decides on the basis of predefined parameters and data models. A human being, on the other hand, possesses freedom of choice as to whether the information provided through workforce analytics will determine outcomes in the decision-making process or whether, and to what extent, other information will also be taken into account. So far, intuition is a unique human characteristic.[24] At the same time, however, it should be emphasised that humans are not free of prejudices.

3 The job applicant's personal data under the GDPR

3.1 The applicability of the GDPR

Living in a 'data society', algorithms can 'exploit the information in large datasets containing thousands of bits of information about individual attributes and behaviours', probably including information pertaining to a person's private sphere that might not be job- or work-related at all.[25] When it comes to the – wholly or partly – automated processing of personal data (by algorithms) and to the processing other than by automated means which (are intended to) form part of a filing system, the GDPR becomes relevant.[26] Processing includes 'profiling', i.e. 'any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, *in particular to analyse or predict aspects concerning that natural person's performance at work*, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements' (emphasis added).[27]

For the GDPR to be applicable, the processing of personal data must take place in the context of activities of an establishment of a controller[28] (i.e. (future) employer) or a processor[29] (i.e. the staffing or recruitment agency) in the EU, regardless of whether the processing takes place in the EU. Personal data is any information relating to a natural person who can be identified, directly or indirectly, in particular by reference to an identifier, e.g. a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.[30] Notably, the right to the protection of personal data is not absolute, especially where a job requires a certain healthy condition. That right must be considered in relation to its function in society and, in accordance with the proportionality principle, be balanced against other fundamental rights of the employer, for instance.[31]

3.2 The consent for lawful processing of personal data

Subject to a range of principles, the processing of personal data must be: lawful, fair and transparent with regard to the data subject; collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed; accurate and kept up to date; kept in a form which permits identification of data subjects for no longer than is necessary for the purposes; and processed in a manner that ensures appropriate security of the personal data.[32] Lawful data processing occurs where the job applicant has given his free, specific, informed and unambiguous 'consent' to the processing of his personal data for one or more specific purposes, i.e. the job application.[33] Consent might require the applicant to tick a box on the website before submitting the application. Job applicants should also be able to withdraw their consent.[34]

An 'explicit consent', for instance through an express written or electronic statement (via email),[35] is needed in two situations. First, where the job applicant's data is part of the 'special categories', i.e. racial or ethnic origin, political opinions, religious or philosophical beliefs or trade union membership, as well as the processing of genetic and biometric data for the purpose of the unique identification of a person, data on health, and data relating to a person's sexual conduct or

sexual orientation.[36] Second, where the job applicant is merely subject to automated processing, including profiling, which produces legal effects for him, e.g. the cancellation of a contract, or similarly significantly affects him, such as 'decisions that deny someone an employment opportunity or put them at a serious disadvantage'.[37]

The question, however, is whether the job applicant actually gives his free, specific, informed and unambiguous explicit consent. Assuming there is an 'imbalance of power' in the (future) employment relationship and the dependency on the (future) employer, the Article 29 Working Party, which is the EU Advisory Body on Data Protection and Privacy,[38] expresses its concerns that the data subject is not able to give (or deny) his consent 'to data processing without experiencing the fear or real risk of detrimental effects as a result of a refusal'. As the consent is unlikely to be given freely, i.e. real choice and control for data subjects,[39] it may be problematic that employers process personal data of (future) employees.[40]

> Regarding the recruitment process, the Working Party cautions employers for all too easily using data made available through social media, for the (private) information shared may have nothing to do with the job in question

Two ways exist in which the employer can, without consent by the job applicant, lawfully process data. First, with a view to 'specific categories' of data, data processing can be lawful where the data has been 'manifestly made public by the data subject'.[41] Regarding the recruitment process, the Working Party cautions employers for all too easily using data made available through social media, for the (private) information shared may have nothing to do with the job in question. Otherwise, this would be contrary to the requirement that data processing must be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed. In addition, the employer must inform the applicant in case he uses social media data.[42] Second, data subjects can be subjected to a decision based solely on automated data processing, including profil-

ing, if the decision 'is necessary for entering into, or performance of, a contract between the data subject and a data controller'. Here, it should be added that any – minimal – human intervention seems to allow 'circumventing' the general prohibition on being subject to automated decision-making, except where it concerns the above exemption.[43] Meaning that even if the software takes the ultimate decision on which candidate will get the job, if a human needs to push a button that makes sure the acceptance/rejection email is sent, this could interrupt the automated decision. This includes also the pre-contractual recruitment phase, where there is an exceptionally high number of job applicants and automated decision-making processes are used to manage the applications. According to the Working Party, relying on automated decision-making must be necessary, i.e. there may be no less intrusive method to achieve the objective.[44] Merely mentioning that profiling will be used would be insufficient to meet the necessity requirement.[45] Member States may, by law or collective agreements, adopt more specific rules in relation to the processing of employees' personal data in the employment context, and in particular for the purposes of the recruitment process and equality and diversity at the workplace.[46]

4 The position of the job applicant under EU non-discrimination laws

4.1 EU-non discrimination laws: the protected grounds
EU law protects (future) workers against discrimination on different grounds: sex (Art. 157 TFEU, Directive 2006/54/EC[47]), race and ethnic origin (Directive 2000/43/EC[48]), religion or belief, disability, age or sexual orientation (Directive 2000/78/EC[49]).[50] Although Article 51(2) EU Charter does not extend the EU's legislative competence, it is interesting, especially in relation to algorithms, that Article 21 lists additional characteristics, such as colour, social origin, genetic characteristics, language, political or other opinions, membership of a national minority, property and birth. Article 3 Directive 2000/43/EC and Directive 2000/78/EC apply to all persons, as regards both the public and private sectors, including public bodies, in relation to, *inter alia*, the conditions for access to employment, self-employment and occupation, including selection criteria and recruitment conditions, whatever the branch of activity and at all

40 Article 29 Data Protection Working Party Guidelines on Consent for the purposes of Regulation 2016/679, p. 8.

41 Art. 9 (1) (e) Regulation (EU) 2016/679.

42 Article 29 Data Protection Working Party Opinion 2/2017 on Data Processing at Work, p. 11.

43 See for an example: Article 29 Data Protection Working Party Guidelines on Automated Individual Decision-Making and Profiling for the purposes of Regulation 2016/679, p. 13.

44 Article 29 Data Protection Working Party Guidelines on Automated Individual Decision-Making and Profiling for the purposes of Regulation 2016/679, p. 23.

45 Article 29 Data Protection Working Party Guidelines on Automated Individual Decision-Making and Profiling for the purposes of Regulation 2016/679, p. 13.

46 Art. 88 Regulation (EU) 2016/679. The Netherlands has decided not to invoke the option provided by Art. 88 GDPR yet. See *Kamerstukken II* 2017/18, 34851, 3, p. 81.

47 Directive 2006/54/EC of 5 July 2006 on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation (recast) (*PbEG* 2006, L 204), implemented in the Equal Treatment in Employment (Men and Women) Act (*Wet gelijke behandeling van mannen en vrouwen*), de Equal Treatment Act (*Algemene wet gelijke behandeling*) and Art. 7:646 Civil Code (art. 7:646 Burgerlijk Wetboek).

48 Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (*PbEG* 2000, L 180), implemented in the Equal Treatment Act (*Algemene wet gelijke behandeling*).

49 Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation (*PbEG* 2000, L 303), implemented in the Equal Treatment Act (*Algemene wet gelijke behandeling*), Equal Treatment in Employment (Age Discrimination) Act (*Wet gelijke behandeling op grond van leeftijd bij de arbeid*), and Equal Treatment

(Disability and Chronic Illness) Act (*Wet gelijke behandeling op grond van handicap of chronische ziekte*).

50 See also Art. 19 TFEU.

51 See extensively N. Kountouris, 'The Concept of 'Worker' in European Labour Law: Fragmentation, Autonomy and Scope', *Industrial Law Journal* 2017, Issue 2. See ECLI: EU:C:2004:18, C-256/01 (*Allonby*).

52 ECLI:EU:C:1990:383, C-177/88 (*Dekker*) (sex/pregnancy), CJEU 4 October 2001, ECLI:EU:C: 2001:513, C-109/00 (*Tele Danmark / HK Danmark*) (sex/pregnancy), CJEU 10 juli 2008, ECLI:EU:C: 2008:397, C-54/07 (*Feryn*) (race/ethnic origin).

53 Critical on excluding self-employed persons from equal treatment laws: S. Fredman, 'Equality Law: Labour Law or an Autonomous Field?', in: A. Bogg, C. Costello, A. Davies & J. Prassl, *The Autonomy of Labour Law,* Oxford: Hart Publishing 2015, p. 260.

54 Fredman 2015, p. 260.

55 E.g., Recital 15 of the Preamble of Directive 2000/43/EC and Directive 2000/78/EC.

56 C. Barnard, *EU Employment Law,* Oxford: Oxford University Press 2012, p. 281, referring to sex discrimination, but this applies to discrimination based on other grounds as well.

57 Kim 2017, p. 877, with reference to Barocas & Selbst 2016, p. 691-962.

58 Art. 6 Directive 2000/78/EC.

59 A legitimate aim may, for example, consist in giving preference to a male candidate who did military service (CJEU 7 December 2000, ECLI:EU:C:2000: 676, C-79/99 (*Schnorbus*)) or ensuring high quality of service and continued competence justifying a fixed retirement age (CJEU 12 januari 2010, ECLI:EU: C:2010:4, C-341/08 (*Petersen*) and CJEU 21 juli 2011, ECLI:EU:C:2011: 508, C-159/10 and 160/10 (*Fuchs & Köhler*)).

60 Two recent cases: CJEU 14 maart 2017, ECLI: EU:C:2017:204, C-188/15 (*Asma Bougnaoui*) and CJEU 14 maart 2017, ECLI:EU:C:2017:203, C-157/15 (*Samira Achbita*).

levels of the professional hierarchy, including promotion. Similarly, Article 1 (a) Directive 2005/54/EC contains the principle of equal treatment in relation to access to employment, including promotion, and vocational training.

Whereas Article 157 TFEU on the right to equal pay for men and women applies to persons who meet the conditions applicable to a Union worker,[51] the other EU instruments apply to persons that find themselves in one of the circumstances defined in the directives. Self-employed persons are, with one exception, excluded from EU equal treatment law.[52] Nevertheless, they may, where they applied for a job to provide services, also experience the consequences of a discriminatory or biased algorithm.[53] It should be noted that discrimination is the exercise of power, which is clearly not limited to the relationship between an employer and an employee.[54] That means, also those working or wanting to work as self-employed persons may be subject to discriminatory decisions.

> The motive or intent of the discrimination does not matter: it is sufficient to show that the disadvantageous treatment is based on, or caused by, the application of a protected ground of discrimination

Both direct and indirect discrimination are protected. Direct discrimination refers to a situation in which one group is, has been, or would be treated less favourably, on grounds of sex or race for instance, than another group as regards access to employment. In order to prove direct discrimination, it is necessary to identify an actual or hypothetical comparator who is in a comparable position and who has been treated more favourably. The motive or intent of the discrimination does not matter: it is sufficient to show that the disadvantageous treatment is based on, or caused by, the application of a protected ground of discrimination. Indirect discrimination is aimed at provisions, criteria or practices which appear to be neutral but which *de facto* discriminate against, for instance, a considerably higher percentage of women than men, or job candidates of a particular age. Here, an actual or potential disparate impact must be shown either by statistical evidence or any other

means.[55] The pool from which the comparator must be taken would be all job applicants that applied for the same job.[56]

4.2 Objective justification for indirect discrimination

In the context of software designed to parse large amounts of CVs, it can be assumed that, unless of course the algorithms are designed in a way to select candidates based on protected grounds, there usually is a case of indirect discrimination. There is one exception, namely discrimination based on pregnancy which always is direct discrimination, as only women can be pregnant. The software and the data used, including the information the job applicant provides, may also be – unconsciously – discriminatory when neutral factors act as 'proxies' for sensitive characteristics like race or sex.[57] The address of the job applicant may be a proxy of race where it is located in an area where a large number of people with a certain ethnic origin live. If the job applicant provides information about previous jobs and these jobs are predominantly performed by women, this may be a proxy for sex. In the case of indirect discrimination, however, employers can objectively justify not hiring a particular job applicant. The only exception is made for direct discrimination based on age.[58] An objective justification exists where the employer can prove that the provision, criterion or practice is objectively justified by a legitimate aim,[59] and the means of achieving that aim are appropriate and necessary.

Where an employer looks for a candidate that is willing to work overtime, but female applicants having a family are rejected, this may be an indication of sex discrimination. Candidates indicating not being available for work on Sundays or on public holidays or other religious holidays may possibly be rejected based on religion or belief. This similarly applies to a candidate who submits a picture of herself wearing a headscarf.[60] Parsing CVs based on criteria that are identified with being young and dynamic may indicate the employer is looking for workers of a particular age group, based on which those that do not belong to that age group may be rejected for the job, implying possible age discrimination.

It can be argued that in cases where the job applicant can demonstrate facts that indirect discrimination exists, the employer should verify whether the algorithmic model and the data models used as such are non-discriminatory. Indeed, it is the employer that has access to information on how the decision

was made. Article 19(1) Directive 2006/54/EC, for instance, allows persons who consider themselves wronged because the principle of equal treatment has not been applied to them to establish, before a court or other competent authority, facts from which it may be presumed that there has been direct or indirect discrimination, with the result that it is for the respondent to prove that there has been no breach of the principle of equal treatment.[61]

A possible problem that may arise with proving discrimination is that an employer qualifies its algorithms as business secrets

4.3 Proving discrimination: business secrets and other practical and legal problems
A possible problem that may arise with proving discrimination is that an employer qualifies its algorithms as business secrets. In accordance with Directive 2016/943/EU on trade secrets, implemented (but not yet in force) in the Trade Secret Protection Act,[62] information that meets the following cumulative conditions is considered to be business secrets: a) is secret; b) it has commercial value because it is secret; and c) it is subject to reasonable measures to keep it secret by the person who lawfully holds it.[63] The purpose of the Directive is to protect against the unlawful acquisition, use and disclosure of business secrets. Nothing is regulated about the situation in which access to a business secret, such as an algorithm, is demanded because of a (allegedly) discriminatory decision taken. The question is whether this is problematic.

It could be argued that it is the employer who should justify why he prefers hiring a particular candidate over another, this should also apply to an employer who uses an algorithm that prefers the one over the other. The rejected job candidate should be given access to the algorithmic model and data model based on which the algorithm has decided, or the employer should at least

provide insight into why that decision can be objectively justified, in order to assess whether there indeed has been discrimination. It is suggested that where an employer is unable to explain why the software did decide in a particular way, and this might be even more difficult where unsupervised

machine-learning algorithms are involved, he should be able to justify the discriminatory treatment. Moreover, where decisions are based on correlations, it might be quite hard to identify any discrimination.

Under the GDPR, job applicants have the right to have access to information on where their data has been collected, what data has been collected and for what purpose. Although the latter might be clear when applying for a job, employers might also want to use the collected data for other (not specified) purposes that may even not be related to the particular job (vacancy) and the applicant. Job applicants may even request rectifying and erasing their data.[64]

Besides legal challenges, there may be practical problems here as well: as indicated above, software developers should ideally document the way in which the algorithm makes decisions and which data sources it uses for that purpose, including all subsequent changes to that model. This is important to evaluate whether the algorithm and the processed data have decided on a certain date that a male job applicant for instance is favoured over a female job applicant. If a subsequent change is made to the algorithmic model and/or other data models are used, it cannot be ascertained why the female candidate did indeed not receive the job. That documentation is crucial, is demonstrated by the fact that the data used by the algorithm may no longer be available, for example due to data protection rules such as the GDPR, which requires data to be available only for as long as is necessary and to delete data once there is no lack of necessity anymore. Employers must inform job candidates about the period for which the data will be stored.[65] At the same time, documentation would mean that a large amount of data would have to be stored, which can be rather costly for a company.

That documentation is crucial, is demonstrated by the fact that the data used by the algorithm may no longer be available, for example due to data protection rules such as the GDPR, which requires data to be available only for as long as is necessary and to delete data once there is no lack of necessity anymore

61 Also Art. 8 Directive 2000/43/EC and Art. 10 Directive 2000/78/EC.
62 Directive 2016/943/EU of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure (*PbEU* 2016, L 157/1); *Kamerstukken II* 2017/18, 34821. See on the Directive: P.J. van der Korst, 'Bedrijfsgeheimen: vanaf 2018 wettelijk geregeld', *Ondernemingsrecht* 2017/48, Issue 6; A.M. Helstone & T. Barkhuysen, 'Ter Visie – Bescherming van bedrijfsgeheimen', *Tijdschrift voor Arbeid & Onderneming* 2018, Issue 1; E. van Vliet, 'Het wetsvoorstel Wet bescherming bedrijfsgeheimen', *TRA* 2018/40, Issue 5.
63 Art. 2 Directive 2016/943/EU.
64 Art. 13, 16 and 17 ('the right to be forgotten') Regulation (EU) 2016/679. Only if the data subject knows which data are being processed, by whom and how, can he successfully exercise the rights guaranteed by the GDPR. R. Ducato, M. Kullmann & M. Rocca, '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*, 19-21 March 2018.
65 Art. 13 (2) (a) Regulation (EU) 2016/679.

Suppose access is provided and the algorithmic model and any subsequent changes to it including the data models it used are fully documented and thus traceable to the date on which the decision has been made, the outcome could be that indeed the decision has not been based on any protected ground, i.e. that there has been no case of discrimination. In such a situation, it would be hard for any human employed by the business using such software to deny the outcome of the analysis of the algorithmic model and its data sources, especially where it has been relied upon that decision in full without an individual interfering and perhaps disregarding the automated decision. If the decision on who will get the job would not be 'outsourced' to algorithms and would be taken by a human, there still would be some leeway in trying to go around the assumption that there has been some form of bias or discrimination involved; think of HR calling the applicant that he would not be a good fit within the team. Therefore, in an ideal situation, algorithmic decision-making could even be more honest and transparent than a human could be.[66]

5 Some final thoughts
Overconfidence in algorithms and their decision-making potential might be a huge danger.[67] Besides, it is common knowledge that software never is free of errors. Usually, a recruitment process consists of several steps, and CV parsing may only be one first, though important, step. While employers may experience organisational benefits in using specific software to select the right employees, job applicants might encounter difficulties in passing the scrutiny of algorithms, especially

where the information provided either is seen as not being relevant, or is not recognised as being relevant. Even though algorithms can, if not always perfectly, take over human decision-making, humans will play an important, if not new, role when it comes to identifying algorithmic decision-making patterns that just codify deep-seated biases or, by surfacing applicants who have certain attributes, making workplaces just as homogeneous as they were before.

> **If the algorithm's model and any changes that are being made to it are not documented at all or only sporadically, it might be impossible to prove the discrimination**

EU non-discrimination laws, in combination with the GDPR, provide (unsuccessful) job applicants with a few legal means to address discriminatory or biased automated decision taken by an employer. However, with regard to non-discrimination laws, the difficulty remains to prove any kind of discrimination. First, because the claimant needs to produce enough facts that substantiate the alleged discrimination and related to that, second, it might be difficult to get access to the algorithmic model. Moreover, even if access is provided, it is still necessary to know how the algorithm decided on the date the alleged discriminatory decision was taken. If the algorithm's model and any changes that are being made to it are not documented at all or only sporadically, it might be impossible to prove the discrimination.

66 C. Kuner et al., 'Machine learning with personal data: is data protection law smart enough to meet the challenge?', *International Data Privacy Law* 2017, Issue 1, p. 3.

67 C. Cain Miller, 'Can an Algorithm Hire Better Than a Human?', *New York Times* (25 June 2015) *www.nytimes.com/2015/06/26/upshot/can-an-algorithm-hire-better-than-a-human.html*.