

FIGHTING ALGO-DISCRIMINATION

Relying on AI's learning skills and embedding fair information principles within machines' heart-design-specifications

Abstract

After having built private networks of knowledge, firms are aware of users' personal data and are capable of creating systems to sort people into groups. The potential of powerful and opaque algorithms to create discriminatory biases has been widely acknowledged, and there is no explanation provided with regard to their decision-making. People ignore ways, in which their information is created or modified, and there is a need for user-centric systems to implement trust and transparency principles into machines' design specifications. This paper studies practices that firms conduct to algorithmically reach a perfect audience. The European regime is examined to prove the discriminatory nature of these practices and support that it cannot be justified by law. Taking into account machines' potential, but also their ability to learn and their inability to forget, proposals are submitted to avoid discrimination. Finally, conclusions are drawn to support that data-scientists, the ones capable of making data speak, could play an important role in embedding fair information principles within machines' life cycle.

Keywords: personal data, algorithms, discrimination, decision-making

1. Introduction

In the Big Data environment¹, firms control personal data and can hence very well build private networks of knowledge². While power is accumulated in private

¹ The number of the traditional "3Vs" characteristics of Big Data (Volume, Variety and Velocity) is increasing; Variability, Veracity, Visualization, and –most importantly– Value have been added to the list. See Tim Chartier, Vertigo Over the Seven V's of Big Data, *The Journal of Corporate Accounting & Finance*, Vol. 27, Issue 3, March/April 2016, pp. 81-82, at p. 81 (<https://doi.org/10.1002/jcaf.22145>), who also offers another "V": Vertigo. See also Samuel Fosso Wamba, Shahriar Akter, Andrew Edwards, Geoffrey Chopin, Denis Gnanzou, How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study, *International Journal of Production Economics*, Vol. 165, July 2015, pp. 234-246, at p. 235, who define Big Data as a "holistic approach to manage, process and analyze 5 Vs (i.e., volume, variety, velocity, veracity and value) in order to create actionable insights for sustained value delivery, measuring performance and establishing competitive advantages". Available at <https://doi.org/10.1016/j.ijpe.2014.12.031>.

hands, the fundamental right to non-discrimination may be threatened³. For example, during online advertising personal data is collected⁴ to profile⁵ users or create target groups, to which personalized ads are addressed⁶. And since firms can be aware of a user's sensitive information⁷, e.g. her gender or ethnicity, they are capable of discriminating against her⁸ on the grounds of gender, ethnicity, or any other information collected⁹. So, systems can be created to sort people into groups, exclude most, and reach a perfect audience to target ads¹⁰. But sorting and profiling

² See Julia Powles & Hal Hodson, *Google DeepMind and healthcare in an age of algorithms*, Health and Technology, Springer, 2017, Vol. 7(4), pp. 351-367, doi: 10.1007/s12553-017-0179-1. Available at <https://www.ncbi.nlm.nih.gov/pubmed/29308344>.

³ Under Article 21(1-2) of the Charter of Fundamental Rights of the European Union “[...] Any discrimination based on any ground such as sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation shall be prohibited [...] Within the scope of application of the Treaties and without prejudice to any of their specific provisions, any discrimination on grounds of nationality shall be prohibited [...]”.

⁴ Neil M. Richards & Jonathan H. King, *Big Data and The Future For Privacy*, Handbook of Research on Digital Transformations, Elgar, 2016, pp. 10-13 (with further references). Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2512069.

⁵ Under Article 4(4) of the Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), hereinafter referred to as the “GDPR”, “[...] ‘profiling’ means 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 [...]”.

⁶ Kati Förster & Ulrike Weish, *Advertising Critique: Themes, Actors and Challenges in a Digital Age*, in Gabriele Siegert, M. Bjørn von Rimscha, Stephanie Grubenmann (eds), *Commercial Communication in the Digital Age, Information or Disinformation?*, de Gruyter GmbH, 2017, at p. 19.

⁷ Danah Boyd & Kate Crawford, *Six Provocations for Big Data, A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society*, Oxford Internet Institute, 2011. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431.

⁸ Alessandro Mantelero, *Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection*, Computer Law & Security Review, Volume 32, Issue 2, April 2016, pp. 238-255, at pp. 239-240; Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, Boston College Law Review, Vol. 55, Issue 1, 2014, pp. 93-128, at pp. 94-95, 98.

⁹ Ira S. Rubinstein, *Big Data: The End of Privacy or a New Beginning?*, International Data Privacy Law, Vol. 3, No. 2, 2013, at p. 76; Cathy O'Neil, *Weapons of Math Destruction, How Big Data Increases Inequality and Threatens Democracy*, 2016/2017, Broadway Books, New York, at pp. 3-5, 130-134, 151.

¹⁰ Under Article 22(1) of the GDPR, “[...] The data subject shall have the right 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 [...]”. Interestingly, under Article 22(2)(c) of the GDPR, the above paragraph shall not apply if the decision is based

are undertaken not only to project the perfect ad¹¹ but also to promote the appropriate good at the appropriate price¹², predict criminal behaviors¹³ or evaluate the accused before sentencing courts¹⁴.

Indeed, the potential of the above systems to create discriminatory biases, unfair results, or, in general, exacerbate inequality has been widely acknowledged¹⁵. While individuals enjoy the right to obtain details of any personal data used for profiling¹⁶, there is no right to an explanation of a particular decision¹⁷. So, one may ignore ways in which the sum of information existing about her, i.e. her information identity, is created or modified¹⁸.

on the data subject's "*explicit*" –but, perhaps, given by a single "mouse-click"– consent. See Recital 32 of the GDPR, mentioning that consent may be validly given by ticking a box in a website. It should also be noted that the above Article 22 of the GDPR is a general prohibition, rather than a right to opt-out. See Article 29 Data Protection Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, Adopted on 3 October 2017, at p. 9; Michael Veale & Lilian Edwards, Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling, in *Computer Law and Security Review*, Vol. 34, 2018, pp. 398-404, at p. 400.

¹¹ See Article 29 Data Protection Working Party, Opinion 03/2013 on purpose limitation, 00569/13/EN, WP 203, adopted on 2 April 2013, at p. 46. Available at http://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf.

¹² Joseph Turow & Lee McGuigan, Retailing and Social Discrimination: The New Normal?, in Seeta Peña Gangadharan (Ed.), *Data and Discrimination: Collected Essays*, Open Technology Institute, New America, 2014, pp. 27-30, at pp. 28-29 (available at <https://na-production.s3.amazonaws.com/documents/data-and-discrimination.pdf>); European Data Protection Supervisor, Opinion 7/2015, Meeting the challenges of big data - A call for transparency, user control, data protection by design and accountability, at p. 19 (available at https://edps.europa.eu/sites/edp/files/publication/15-11-19_big_data_en.pdf).

¹³ Anupam Chander, The Racist Algorithm?, *Michigan Law Review*, Vol. 115, Issue 6, 2017, pp. 1023-1045, at p. 1026. Available at <https://repository.law.umich.edu/mlr/vol115/iss6/13/>.

¹⁴ See *State v. Loomis*, 881 N.W. 2d 749 (Wis. 2016), available at <https://harvardlawreview.org/2017/03/state-v-loomis/>.

¹⁵ See Alex Campolo, Madelyn Sanfilippo, Meredith Whittaker, Kate Crawford, *AI Now 2017 Report*, Andrew Selbst, Solon Barocas (eds), *AI Now Institute*, at p. 13. Available at https://ainowinstitute.org/AI_Now_2017_Report.pdf.

¹⁶ See Article 15 of the GDPR.

¹⁷ See Article 29 Data Protection Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, *id.*, at p. 24, where it is also mentioned that simple ways should be found to tell the data subject about the rationale behind or the criteria relied on in reaching the decision. A mathematical explanation about how algorithms work should also be provided to allow experts to verify how the decision-making process works. Article 29 Data Protection Working Party, *id.*, at pp. 14, 29.

¹⁸ Sandra Wachter, Normative challenges of identification in the Internet of Things: Privacy, profiling, discrimination, and the GDPR, *Computer Law & Security Review*, Volume 34, Issue 3, June 2018, pp. 436-449, at p. 439. As Wachter argues, when a person is unaware that her devices generate information about her, she lacks the ability to incorporate this information into her self-

By linking datasets¹⁹, even seemingly neutral data, such as postcodes, can lead to discrimination²⁰ based on sensitive information, like sexual preferences²¹. This reveals the need for user-centric systems²² to establish and maintain trust²³ and grant users oversight and choice over ways in which machines communicate and act on people's behalf²⁴. While it is obvious that discrimination issues are raised, this paper examines ways in which machines can be used to reach the perfect audience and influence people's opportunities. European non-discrimination laws are studied to support the above practices' discriminatory nature, which cannot be justified under the EU regime. Thereafter, proposals, based on machines' capacities and their inability to forget, are submitted to avoid negative effects. Finally, further discussion is conducted to draw final conclusions and highlight data-scientists' capacity to implement fair information principles into machines' design specifications and their entire life cycle.

2. Algorithms to reach the perfect audience

As reported some years ago²⁵, business models and strategies built around collection and use of persons' data raise very important issues with regard to

constructed identity, and to view herself as others view her. See also Luciano Floridi, *The Informational Nature of Personal Identity*, 2011, *Minds and Machines*, Vol. 21, p. 549.

¹⁹ Salvatore Ruggieri, Dino Pedreschi, Franco Turini, *Data Mining for Discrimination Discovery*, *ACM Transactions on Knowledge Discovery from Data*, Vol. 4, No. 2, Article 9, May 2010.

²⁰ Solon Barocas & Andrew Selbst, *Big Data's Disparate Impact*, 2016, *California Law Review*, Vol. 104, pp. 671-732, available at SSRN: <https://ssrn.com/abstract=2477899> or <http://dx.doi.org/10.2139/ssrn.2477899>.

²¹ See Bart W. Schermer, *The limits of privacy in automated profiling and data mining*, *Computer Law & Security Review*, Vol. 27, Issue 1, February 2011, pp. 45-52, at p. 47.

²² Omer Tene and Jules Polonetsky, *Big Data for All: Privacy and User Control in the Age of Analytics*, 2013, *Nw. J. Tech. & Intell. Prop.*, Volume 11, Issue 5, pp. 239-273, at pp. 242-243, arguing that user-side applications and services are needed to ensure "featurization" of big data, which will, in turn, allow individuals to declare their own policies in ways that can be automated for them and for the firms they engage.

²³ Charith Perera, Arkady Zaslavsky, Peter Christen, and Dimitrios Georgakopoulos, *Context Aware Computing for The Internet of Things: A Survey*, *IEEE Communications Surveys & Tutorials*, Vol. 16, No. 1, 1st Quarter 2014, pp. 414-454, at pp. 424, 443, 445.

²⁴ Rodrigo Roman, Pablo Najera, Javier Lopez, *Securing the Internet of Things*, *Computer*, Vol. 44, Issue 9, Sept. 2011, pp. 51-58, at p. 55, mentioning that users must have "[...] *tools that accurately describe all their interactions, so that they can form an accurate mental map of their virtual surroundings* [...]".

²⁵ John Podesta, Penny Pritzker, Ernest J. Moniz, John Holdren, Jeffrey Zients, *Big Data: Seizing Opportunities, Preserving Values*, Executive Office of the President, May 2014, The White House, Washington, p. 45. Available at

discrimination. Powerful and opaque algorithms²⁶ make automated decisions and score individuals to statistically characterize everything²⁷; from one's ability to pay to whether a prisoner is eligible for parole²⁸. Such scores may be used to influence a person's opportunities to, for example, find housing or a job, or estimate health²⁹.

With regard to advertising techniques, innumerable targeting settings are set to exclude more and more individuals and reach a perfect audience³⁰. For instance, job seekers may be excluded from seeing a job advertisement³¹, and this threatens equality, whose principle demands that every individual should have the same opportunities –including access to employment³².

https://obamawhitehouse.archives.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf.

²⁶ To put it simply, an algorithm is defined by a sequence of steps and instructions that can be applied to data. Algorithms generate categories for filtering information, operate on data, look for patterns and relationships, or, generally, assist in analysis of information. See John Podesta, Penny Pritzker, Ernest J. Moniz, John Holdren, Jeffrey Zients, *Big Data: Seizing Opportunities, Preserving Values*, id, at p. 46.

²⁷ See, in general, Frank Pasquale, *The Black Box Society: The Secret Algorithms that Control Money and Information*, 2015, Harvard University Press, 2015, Cambridge, Massachusetts, London, England, at pp. 8, 30, 34; Cathy O'Neil, *Weapons of Math Destruction*, id.

²⁸ Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, *Machine Bias, There's software used across the country to predict future criminals. And it's biased against blacks*, ProPublica, May 23, 2016. Available at <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.

²⁹ Pam Dixon and Robert Gellman, *The Scoring of America: How Secret Consumer Scores Threaten Your Privacy and Your Future*, World Privacy Forum, April 2, 2014, at pp. 13-15. Available at http://www.worldprivacyforum.org/wp-content/uploads/2014/04/WPF_Scoring_of_America_April2014_fs.pdf.

³⁰ Amit Datta, Michael Carl Tschantz, Anupam Datta, *Automated Experiments on Ad Privacy Settings, A Tale of Opacity, Choice, and Discrimination*, in *Proceedings on Privacy Enhancing Technologies (PoPETs)*, 2015, De Gruyter Open. Extended version available at <https://arxiv.org/pdf/1408.6491.pdf>. The above authors found that, in case of Google Ads system, setting the gender to female resulted in getting fewer instances of an ad related to high paying jobs.

³¹ Paul Post and Rikki Holtmaat, *A False Start: Discrimination in Job Advertisements*, *European Gender Equality Law Review* – No. 1/2014, at p. 12. Available at <https://openaccess.leidenuniv.nl/bitstream/handle/1887/35024/EGELR%202014-2%20-%20PP%20%26%20RH%20-%20A%20False%20Start.%20Discrimination%20in%20Job%20Advertisements.pdf?sequence=1>.

³² See, in general, Richard J. Arneson, *Equality of Opportunity: Derivative Not Fundamental*, *Journal of Social Philosophy*, Wiley Periodicals, Inc., Vol. 44 No. 4, Winter 2013, pp. 316-330, available at <https://onlinelibrary.wiley.com/doi/pdf/10.1111/josp.12036>; Richard Arneson, *Egalitarianism*, *The Stanford Encyclopedia of Philosophy*, Summer 2013 Edition, Edward N. Zalta (ed.), available at <https://plato.stanford.edu/entries/egalitarianism/>.

Online advertising, such as Facebook Ads³³, differs from traditional techniques, since, today, human input is used to run an ad campaign and, thus, the audience can be manipulated precisely, rather than generally. Hence, the effectiveness of advertising increases because personalized ads reach the right individual at the right time³⁴. So, attention is drawn to reaching the right people and showing them what they wish to see, i.e. content appealing enough to click on.

In fact, a firm may target its ads based on location (e.g. zip code, city, county, or country), age, gender, demographics (e.g. income, job title, employer name, language, relationship status, education, financial or parental status), interests, behaviors (such as online shopping or travel habits), or connections (meaning people connected to one's page, app, etc)³⁵. The above targeting may be conducted by humans or by an automated process. And when Artificial Intelligence³⁶ (AI) is used to tailor automated ads, it makes targeting decisions and examines results based on pre-programmed and self-learned strategies.

And how do machines learn how to make decisions?

Machine learning³⁷ makes AI better capable of making the right decision; the longer it learns the more efficient and higher the quality of its decisions. For

³³ See <https://www.facebook.com/business/products/ads>.

³⁴ Alexander Bleier, Maik Eisenbeiss, Personalized Online Advertising Effectiveness: The Interplay of What, When, and Where, *Marketing Science*, 2015, Vol. 34 No. 5, pp. 669-688, at p. 669.

³⁵ See Kane Jamison, The Big Damn Guide To Facebook Ad Targeting, available at <https://www.contentharmony.com/blog/facebook-ad-targeting/#facebook-ad-targeting-overview>.

³⁶ Artificial Intelligence (AI) could be defined as any system or device that perceives its environment and undertakes actions that maximize its chance of success at some goal. See definitions of AI in Stuart Russell, Peter Norvig, *Artificial Intelligence, A Modern Approach*, 2nd Ed., 2003, 1995 by Pearson Education, Inc. Pearson Education, Inc., at pp. 2-5. Available at [http://www.eng.uerj.br/~fariasol/disciplinas/Topicos_B/AGENTS/books/Stuart%20Russell,%20Peter%20Norvig-Artificial%20Intelligence_%20A%20Modern%20Approach-Prentice%20Hall%20\(2002\)-2nd-ed.pdf](http://www.eng.uerj.br/~fariasol/disciplinas/Topicos_B/AGENTS/books/Stuart%20Russell,%20Peter%20Norvig-Artificial%20Intelligence_%20A%20Modern%20Approach-Prentice%20Hall%20(2002)-2nd-ed.pdf). See also Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, 2014.

³⁷ Machine learning focuses on ways to construct a system that automatically improves through experience and on finding the fundamental statistical, computational, information theoretic laws that govern all learning systems. See M. I. Jordan, T. M. Mitchell, Machine learning: Trends, perspectives, and prospects, *Science*, 2015, Vol. 349, Issue 6245, pp. 255-260, DOI: 10.1126/science.aaa8415, at p. 255. Available at <http://science.sciencemag.org/content/349/6245/255.full>. See also proposals on machine learning systems designing to limit discrimination in Michael Veale & Reuben Binns, Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data, *Big Data & Society*, 2017, Vol. 4, Issue 2 (available at <http://journals.sagepub.com/doi/abs/10.1177/2053951717743530>). For machine learning systems, see, in general, Trevor Hastie, Robert Tibshirani & Jerome Friedman, *The Elements of Statistical*

instance, with regard to job ads, the system will learn the targeting settings that are the most effective for a job. Since the algorithm looks for statistical correlations of data (rather than understanding cause-and-effect relationships)³⁸ patterns may seem random to humans, to whom settings would never have been occurred. Namely, AI may find that female individuals perform worse than males when undertaking a specific task. So, it can exclude females and show the (e.g. job) ad to a male audience.

Is this not discrimination?

3. EU non-discrimination regime

Discrimination describes a situation where an individual is disadvantaged in some way because of a protected characteristic. Under the European non-discrimination law³⁹, individuals who are in similar situations should receive

Learning: Data Mining, Inference, and Prediction, Springer Series in Statistics, 2nd ed., 2009. Available at <https://web.stanford.edu/~hastie/Papers/ESLII.pdf>.

³⁸ Allan G. King & Marko Mrkonich, “Big Data” and the Risk of Employment Discrimination, *Oklahoma Law Review*, 2016, Vol. 68, Issue 3, pp. 555-584, at p. 555, available at <https://digitalcommons.law.ou.edu/olr/vol68/iss3/3/>. See also Viktor Mayer-Schonberger & Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, And Think*, Eamon Dolan ed., Mariner Books, 2014, Boston, New York, at p. 68 mentioning that “[...] *Causality won’t be discarded, but it is being knocked off its pedestal as the primary fountain of meaning. Big data turbocharges non-causal analyses, often replacing causal investigations [...]*”.

³⁹ Under the EU regime, three directives prohibit discrimination. The Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation (hereinafter referred to as the “Employment Equality Directive”) prohibits discrimination on the basis of sexual orientation, religious belief, age and disability in the area of employment. The Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin (hereinafter referred to as the “Racial Equality Directive”) prohibits discrimination on the basis of race or ethnicity in the context of employment and in accessing the welfare system and social security, and goods and services. The Directive 2006/54/EC of the European Parliament and of the Council 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), hereinafter referred to as the “Gender Equality Directive”, prohibits discrimination based on sex in the context of employment and social security. It should be noted that all European Member States are party to several United Nations human rights treaties that contain prohibitions on discrimination. See the International Covenant on Civil and Political Rights (ICCPR, <http://www.ohchr.org/en/professionalinterest/pages/ccpr.aspx>); the International Covenant on Economic Social and Cultural Rights (ICESCR, <http://www.ohchr.org/EN/ProfessionalInterest/Pages/CESCR.aspx>); the Convention on the Elimination of All Forms of Racial Discrimination (ICERD, <http://www.ohchr.org/EN/ProfessionalInterest/Pages/CERD.aspx>); the Convention on the Elimination of Discrimination Against Women (CEDAW,

similar treatment and not be treated less favorably because of a particular protected characteristic that they possess. This is called direct discrimination⁴⁰. EU law also stipulates that individuals, who are in different situations, should receive different treatment –to the extent that this is needed to allow them to enjoy particular opportunities on the same basis as others. These same “protected grounds” should be taken into consideration when undertaking specific practices or providing particular rules. This is known as indirect discrimination⁴¹.

The difference of treatment, to which an individual is subject, is at the heart of direct discrimination, the main element of which is evidence of such unfavorable treatment⁴². So, an example of direct discrimination can be to receive lower pay, albeit a complaint about lower pay can be a claim of discrimination, only if it can be shown that such pay is lower than that of someone employed to perform a similar task –and by the same employer. This means that a comparator is needed, meaning a person in materially similar circumstances, with the main difference

<http://www.ohchr.org/Documents/ProfessionalInterest/cedaw.pdf>); the Convention Against Torture (<http://www.ohchr.org/EN/ProfessionalInterest/Pages/CAT.aspx>); the Convention on the Rights of the Child (CRC, <http://www.ohchr.org/EN/ProfessionalInterest/Pages/CRC.aspx>).

⁴⁰ Direct discrimination occurs when an individual is treated unfavorably by comparison to how others, who are in a similar situation, have been or would be treated and the reason for this is a particular characteristic they hold, which falls under a protected ground. See Article 2(2) of the Racial Equality Directive; Article 2(2)(a) of the Employment Equality Directive; Article 2(1)(a) of the Gender Equality Directive (Recast); Article 2(a) of the Council Directive 2004/113/EC of 13 December 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services (hereinafter referred to as the “Gender Goods and Services Directive”); European Union Agency for Fundamental Rights (FRA), Council of Europe, Handbook on European non-discrimination law, 2010, at p. 22. The European Court of Human Rights (hereinafter referred to as “ECtHR”) has held that there must be a “*difference in the treatment of persons in analogous, or relevantly similar, situations*”, which is “*based on an identifiable characteristic*”. See ECtHR, *Carson and Others v. UK* (Grand Chamber), No. 42184/05, 16 March 2010, par. 61; ECtHR, *D.H. and Others v. the Czech Republic* (Grand Chamber), No. 57325/00, 13 November 2007, par. 175.

⁴¹ See European Union Agency for Fundamental Rights (FRA), Council of Europe, 2010, Handbook on European non-discrimination law, id, at p. 21.

⁴² Direct discrimination is characterized by differential treatment, meaning it must be shown that the alleged victim has been treated less favorably based on the possession of a characteristic falling under a protected ground. Less favorable treatment is determined through a comparison between the alleged victim and another person in a similar situation who does not possess the protected characteristic. See ECtHR, *Burden v. The United Kingdom* (Grand Chamber), No. 13378/05, 29 April 2008, par. 60, mentioning that “[...] *a difference in the treatment of persons in relevantly similar situations [...] is discriminatory if it has no objective and reasonable justification; in other words, if it does not pursue a legitimate aim or if there is not a reasonable relationship of proportionality between the means employed and the aim sought to be realized [...]*”.

between the two persons being the protected ground⁴³. With regard to the protected ground, one should ask whether the person would have been treated less favorably if he or she had been of a different sex, race, age etc. If the answer is yes, less favorable treatment is being caused by the ground in question⁴⁴.

On the other hand, the elements of indirect discrimination are a neutral rule, criterion or practice, which affects a group defined by a protected ground in a significantly more negative way by comparison to others in a similar situation⁴⁵. The most important requirement is an apparently neutral rule, criterion or practice, meaning that there must be some form of requirement that is applied to everybody⁴⁶. This apparently neutral provision, criterion or practice is required to

⁴³ For instance, in case *Richards v. Secretary of State for Work and Pensions*, the complainant, who had undergone ‘male to female’ gender reassignment surgery, wished to claim her pension on her 60th birthday, which was the age that women were entitled to pensions in the UK. However, the government refused to grant the pension, arguing that the complainant had not received unfavorable treatment by comparison to those in a similar situation (as alleged by the government, the correct comparator in this case was “men”, as the complainant had lived as a man). But as the ECJ held, since national laws allow a person to change his/her gender, the correct comparator was “women” and, thus, the complainant was being treated less favorably than other women (by having a higher retirement age imposed on her). See European Court of Justice (now Court of Justice of the European Union), *Richards v. Secretary of State for Work and Pensions*, Case C-423/04, ECR I-3585, 27 April 2006.

⁴⁴ See, for example, European Court of Justice, *P. v. S. and Cornwall County Council*, Case C-13/94, ECR I-2143, 30 April 1996: the complainant was in the process of undergoing “male to female” gender reassignment when she was dismissed by her employer. The ECJ found that such dismissal constituted unfavorable treatment. Concerning the relevant comparator, the ECJ stated that “[...] where a person is dismissed on the ground that he or she intends to undergo, or has undergone, gender reassignment, he or she is treated unfavourably by comparison with persons of the sex to which he or she was deemed to belong before undergoing gender reassignment [...]” (*P. v. S. and Cornwall County Council*, par. 21). With regard to the grounds, although it could not be shown that the complainant was treated differently, because she was a man or a woman, albeit it could be shown that the differential treatment was based around the concept of her sex.

⁴⁵ European Union Agency for Fundamental Rights (FRA), Council of Europe, 2010, Handbook on European non-discrimination law, id, at p. 29; Article 2(2)(b) of the Racial Equality Directive; Article 2(2)(b) of the Employment Equality Directive; Article 2(1)(b) of the Gender Equality Directive (Recast); ECtHR, *D.H. and Others v. the Czech Republic* (Grand Chamber), No. 57325/00, 13 November 2007, par. 184; ECtHR, *Opuz v. Turkey*, No. 33401/02, 9 June 2009, par. 183; ECtHR, *Zarb Adami v. Malta*, No. 17209/02, 20 June 2006, par. 80, mentioning that “[...] a difference in treatment may take the form of disproportionately prejudicial effects of a general policy or measure which, though couched in neutral terms, discriminates against a group [...]”.

⁴⁶ For example, in case *D.H. and Others v. the Czech Republic*, some tests were used to establish the intelligence and suitability of pupils to determine whether they should be moved out of mainstream education and into special schools (which were designed for those with intellectual disabilities or other sources of learning difficulty). Same tests were applied to all pupils, who were to be placed in special schools. But, in practice, the test had been designed around the mainstream Czech population with the consequence that Roma students were inherently more likely to perform badly. As they did so, between 50% and 90% of Roma children were educated

place a protected group at a particular disadvantage, and this is exactly where indirect discrimination differs from direct; it moves the focus away from differential treatment and looks at differential effects⁴⁷.

Coming back to direct discrimination, the differential treatment may be justified in a limited set of circumstances. For example, the Genuine Occupational Requirement allows employers to differentiate against individuals on the basis of a protected ground, where the latter has an inherent link with the capacity to perform or the qualifications required for a particular job⁴⁸. Namely, some professions may require specific attributes: a female singer to fit with a taste in performance style; a young actor to play a role, or; women for particular types of fashion modeling.

It is important to mention that some issues of facts, such as the existence of prejudice or the intention to discriminate, are not of relevance to determining whether the legal test for discrimination has been satisfied. What matters and what must be proven is simply the existence and the very fact of differential treatment, on a basis of a prohibited ground that is not justified. This means that there is no need to prove that the perpetrator is motivated by prejudice, e.g. that he or she has racist views, to prove race discrimination⁴⁹. Besides, law cannot regulate individuals' entirely internal attitudes –but it may regulate actions through which these attitudes manifest themselves⁵⁰. Moreover, there is no need to show that the practice in question is intended to result in a differential treatment. For instance, if a public authority points to a well-intentioned and good-faith practice, but the very

outside the mainstream education system. Thus, the court held that this was a case of indirect discrimination. See ECtHR, *D.H. and Others v. the Czech Republic* [GC] (No. 57325/00), 13 November 2007, par. 79.

⁴⁷ In order to be presumed discriminatory, the measure must affect “[...] a far greater number of women than men [...]” (ECJ, *Rinner-Kühn v. FWW Spezial-Gebäudereinigung*, Case C-171/88, ECR 2743, 13 July 1989) or “[...] a considerably lower percentage of men than women [...]” (ECJ, *Nimz v. Freie und Hansestadt Hamburg*, Case C-184/89, ECR I-297, 7 February 1991) or “[...] far more women than men [...]” (ECJ, *De Weerd, née Roks, and Others v. Bestuur van de Bedrijfsvereniging voor de Gezondheid, Geestelijke en Maatschappelijke Belangen and Others*, Case C-343/92, ECR I-571, 24 February 1994).

⁴⁸ See Article 14(2) of the Gender Equality Directive (Recast); Article 4 of the Racial Equality Directive; Article 4(1) of the Employment Equality Directive.

⁴⁹ See European Union Agency for Fundamental Rights (FRA), Council of Europe, Handbook on European non-discrimination law, *id.*, at pp. 127, 128.

⁵⁰ See ECJ, Judgment of the Court (Second Chamber), 10 July 2008, in Case C-54/07, *Centrum voor gelijkheid van kansen en voor racismebestrijding v. Firma Feryn NV*, ECLI:EU:C:2008:397.

effect is to disadvantage a specific group, this will amount to discrimination⁵¹. Finally, there is no need to prove that there is actually an identifiable victim⁵².

Given that the motive is irrelevant and that there is no need to establish an identifiable victim, let us examine whether algorithms violate EU non-discrimination laws.

4. Algo-discrimination: Justifiable?

Online ads are shown based on data that is collected about users. The more accurate the data collected, the more effective the targeted advertising efforts, which are to be fueled. For example, Facebook creates ad profiles based on data provided by users, which may concern people's behavior and innumerable other factors⁵³. Such data may not be true, albeit this is irrelevant, since discrimination can be based upon an assumed ground, whether true or not⁵⁴. Moreover, as noted above, there is no need to be an identifiable victim for sanctions to apply⁵⁵. This is

⁵¹ See ECtHR, *D.H. and Others v. the Czech Republic*, id.

⁵² See ECJ, *Centrum voor gelijkheid van kansen en voor racismebestrijding v. Firma Feryn NV*, id.

⁵³ See <https://www.facebook.com/about/privacy/update#what-kinds-of-information-do-we-collect>.

⁵⁴ See Lilla Farkas, *How to Present a Discrimination Claim*, Handbook on seeking remedies under the EU Non-discrimination Directives, European Network of Legal Experts in the non-discrimination field, European Commission, Directorate-General for Justice, 2011, at p. 25, mentioning that “[...] *If an employee is assumed to be gay, Jewish, Muslim or pregnant and is therefore not employed, not promoted or laid off, she can seek protection under European law. It does not matter whether or not she is really gay, Jewish, Muslim, or pregnant [...]*”. Available at https://www.ab.gov.tr/files/ardb/evt/1_avrupa_birligi/1_6_raporlar/1_3_diger/European_Commission_How_to_Present_a_Discrimination_Claim_Handbook_onseeking_remedies_under_the_EU_Nondiscrimination_Directives.pdf. See also ECJ, Judgment of the Court (Grand Chamber), 17 July 2008, in case C-303/06, *S. Coleman v. Attridge Law and Steve Law*, ECLI:EU:C:2008:415, at par. 50, mentioning that “[...] *Although, in a situation such as that in the present case, the person who is subject to direct discrimination on grounds of disability is not herself disabled, the fact remains that it is the disability which, according to Ms Coleman, is the ground for the less favourable treatment which she claims to have suffered. As is apparent from paragraph 38 of this judgment, Directive 2000/78, which seeks to combat all forms of discrimination on grounds of disability in the field of employment and occupation, applies not to a particular category of person but by reference to the grounds mentioned in Article 1 [...]*”.

⁵⁵ See ECJ, Judgment of the Court (Second Chamber), 10 July 2008, in Case C-54/07, *Centrum voor gelijkheid van kansen en voor racismebestrijding v. Firma Feryn NV*, id, at par. 25, mentioning that “[...] *The fact that an employer declares publicly that it will not recruit employees of a certain ethnic or racial origin, something which is clearly likely to strongly dissuade certain candidates from submitting their candidature and, accordingly, to hinder their access to the labour market, constitutes direct discrimination in respect of recruitment within the*

important, since a system may exclude people from seeing an online ad, albeit those excluded will most probably be unaware of the very fact that they are discriminated against (as they will never see the above ad).

So, the answer to our questions on discrimination seems to be straightforward. Given EU non-discrimination law, one could fairly argue that targeting settings, narrowing down an audience, by excluding people based on the protected grounds, amount to direct discrimination. This means that if for example a job ad campaign reaches only straight people, discrimination is a fact.

But could it be justified?

Direct discrimination can be justified when it is in pursuit of aims expressly set out in legal provisions⁵⁶. To focus on the Genuine Occupational Requirement⁵⁷, a difference of treatment which is based on a characteristic related to e.g. sex shall not constitute discrimination where, by reason of the nature of the particular occupational activities concerned or of the context in which they are carried out, such a characteristic constitutes a genuine and determining occupational requirement, provided that its objective is legitimate and the requirement is proportionate. So, as mentioned above, this exception demands that the characteristic is directly related to the capacity to perform or the qualifications required for a particular job. Thus, employing an individual of Indian ethnicity in an Indian restaurant –to maintain authenticity– could be justified⁵⁸.

But the above exception can apply in very limited circumstances that are met when the objective of the Genuine Occupational Requirement is legitimate and proportionate⁵⁹. For example, the aim of setting the age limit for recruitment to intermediate career posts in the fire service at 30 years may be to ensure the operational capacity and proper functioning of the professional fire service, and

meaning of Directive 2000/43. The existence of such direct discrimination is not dependant on the identification of a complainant who claims to have been the victim [...]”.

⁵⁶ See the Employment Equality Directive, the Racial Equality Directive, and the Gender Equality Directive. See also European Union Agency for Fundamental Rights (FRA), Council of Europe, Handbook on European non-discrimination law, id, at p. 43.

⁵⁷ See Article 14(2) of the Gender Equality Directive (Recast); Article 4 of the Racial Equality Directive; Article 4(1) of the Employment Equality Directive.

⁵⁸ See European Union Agency for Fundamental Rights (FRA), Council of Europe, Handbook on European non-discrimination law, id, at p. 47.

⁵⁹ See Recital 23 of the Employment Equality Directive; Recital 18 of the Racial Equality Directive; Recital 19 of the Gender Equality Directive. Similarly, as provided by its very definition, indirect discrimination does not occur when the practice can be objectively justified by a legitimate aim, and the means of achieving that aim are appropriate and necessary. See Article 2(1)(b) of the Gender Equality Directive; Article 2(2)(b) of the Racial Equality Directive; Article 2(2)(b) of the Employment Equality Directive.

can, thus, constitute a legitimate objective⁶⁰. The principle of proportionality requires that the derogation is not only appropriate but also necessary to achieve the aim⁶¹, meaning there are no other measures that could achieve the same result while being less harmful. Proportionality also requires the exception to be transparent in nature⁶².

If all of the above criteria are met, the exception applies.

But given opaque procedures of online ads, one cannot know whether objectives are legitimate and proportionate, or whether derogation is both appropriate and necessary to achieve such aims. When firms massively collect data to feed their algorithms that operate in an automated process, who can tell whether there could be other measures that could achieve the same result while being less harmful? In any case, the “transparency criterion” demanded by the principle of proportionality is definitely not met.

So, after having found direct discrimination, incapable of being justified, let us examine whether AI’s capabilities could be used to eliminate negative effects.

5. Recognizing protected grounds

To avoid direct discrimination use of the protected grounds as targeting settings has to be prevented. However, use of grounds that are e.g. Genuine Occupational Requirements could be allowed. This means that if the protected ground is not explicitly required for a job, it should not be used as a setting. Thereafter, the AI has to recognize such characteristics to prevent them from being used as targeting settings.

A proposal to exclude such grounds from the Big Data environment could be an option. But in such case AI would not be able to pick them up⁶³ nor could it

⁶⁰ See ECJ, Judgment of the Court (Grand Chamber), 12 January 2010, in Case C-229/08, *Colin Wolf v. Stadt Frankfurt am Main*, at par. 33, 39, ECLI:EU:C:2010:3.

⁶¹ See ECJ, Judgment of the Court, 15 May 1986, in Case 222/84, *Marguerite Johnston v. Chief Constable of the Royal Ulster Constabulary*, at par. 38, mentioning that “[...] That principle requires that derogations remain within the limits of what is appropriate and necessary for achieving the aim in view and requires the principle of equal treatment to be reconciled as far as possible with the requirements of public safety which constitute the decisive factor as regards the context of the activity in question [...]”.

⁶² See ECJ, Judgment of the Court, 30 June 1988, in Case 318/86, *Commission of the European Communities v. French Republic*, at par. 25-27.

⁶³ Sara Hajian and Josep Domingo-Ferrer, A Methodology for Direct and Indirect Discrimination Prevention in Data Mining, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 25, No. 7, 2013, pp. 1445-1459, at pp. 1445, 1446. Available at <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6175897>.

recognize them. Instead of excluding them from Big Data, one could use AI systems that do not contain such characteristics. However, in this case, AI models would still be able to use settings that would be apparently neutral, albeit would affect or would be able to affect those who would possess this characteristic⁶⁴. So, manipulating Big Data or AI systems does not seem a sufficient solution.

However, Big Data could be used to determine which neutral grounds would be inextricably linked to the relevant protected grounds. For instance, finding that 100% of pregnancies relate to women would prove the “inextricable link” in every case. This way, a list containing all such grounds could be introduced as a filter between correlations and AI’s output. Thereafter, an AI system could recognize discriminatory targeting settings and would not use those that would be directly discriminatory.

The exception of Genuine Occupational Requirement could also be used. This would let a system know when it could allow use of otherwise directly discriminating targeting settings. For instance, the machine could learn to recognize real life situations and predict their outcomes based on previous events⁶⁵. To put it simply, a machine could make job profiles, realize that some jobs, like fashion models, include both males and females, and, thus, allow targeting based on gender.

Another argument to support that we need to focus on AI systems’ capacity to learn, rather than preventing them from being aware of the above characteristics, is based on machines’ inability to “forget”.

6. Demanding oblivion: A failure?

One could claim that EU laws cannot handle AI, in fields of which data deletion is impossible. The hotly debated Right to Be Forgotten⁶⁶ demands information –made public– to be deleted –and, thus, made private. But when one makes –previously– public information private, she “requests” that others forget it. Does this work in AI?

⁶⁴ See Opinion of Advocate General Kokott, delivered on 12 March 2015, *CHEZ Razpredelenie Bulgaria AD v Komisia za zashtita ot diskriminatsia*, Case C-83/14, ECLI:EU:C:2015:170, at par. 82, mentioning that “[...] *Direct discrimination must also be taken to exist where a measure is apparently neutral, but actually affects or is capable of affecting only persons possessing a certain characteristic mentioned in Article 21 of the Charter of Fundamental Rights and in the anti-discrimination directives* [...]”.

⁶⁵ See Pat Langley, Herbert A. Simon, *Applications of Machine Learning and Rule Induction*, in *Communications of the ACM*, 1995, Vol. 38, No. 11, pp. 55-64. Available at <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.294.4673&rep=rep1&type=pdf>.

⁶⁶ See Article 17 of the GDPR.

Human memory is different from AI's. The right to be forgotten seems to require deletion⁶⁷ of data as if it were a physical thing that could be destroyed. With regard to human mind, scientists have not yet made clear what may get stored in long-term memory⁶⁸. But this is not the case in AI; scientists know very well how machines' memory works and ways these systems treat data storage and deletion.

The right to be “forgotten” is made for humans and it is hard to apply to machines, which are incapable of forgetting the way we, humans, do. For instance, a transaction in a Data Base Management System (DBMS) must maintain Atomicity, Consistency, Isolation, and Durability⁶⁹. The above properties – commonly known as ACID properties– ensure accuracy, completeness, and data integrity. In particular, Atomicity states that operations are done as a whole or not at all. For instance, insert of a data input has to be done for the whole record or not at all⁷⁰. Consistency demands that when an operation is finished, the system has to remain in a consistent state, meaning all relations are unambiguous and the system is normalized⁷¹. Isolation requires that, in case of parallel transactions, the system must guarantee that no transaction affects the existence of any other⁷². Finally and most importantly, Durability means that data must be permanently stored: the database should be durable enough to hold all its latest updates even if the system

⁶⁷ Interestingly, the GDPR does not mention the term “deletion” in its text, while the term “remove” appears twice, albeit does not refer to the right to be forgotten. See Recitals 10 and 65 of the GDPR. Instead, the GDPR uses the term “erasure”, which is not explained nor is it defined.

⁶⁸ See Julian De Freitas, *Why Is Memory So Good and So Bad? Explaining the memory paradox*, May 29, 2012, Scientific American, available at <https://www.scientificamerican.com/article/why-memory-so-good-bad/>; Edward K. Vogel, *Why Do We Forget Things? The brain can store a vast number of memories, so why can't we find these memories when we need to? A new study provides insights into this question*. November 4, 2008, Scientific American, available at <https://www.scientificamerican.com/article/why-do-we-forget-things/>.

⁶⁹ See, in general, Raghu Ramakrishnan, Johannes Gehrke, *Database Management Systems*, 3rd Ed., McGraw-Hill, Inc. New York, USA, 2003. See also DBMS – Transaction, Tutorials Point, available at https://www.tutorialspoint.com/dbms/dbms_transaction.htm.

⁷⁰ See DBMS – Transaction, Tutorials Point, id, mentioning that “[...] *This property states that a transaction must be treated as an atomic unit, that is, either all of its operations are executed or none [...]*”.

⁷¹ See DBMS – Transaction, Tutorials Point, id, mentioning that “[...] *The database must remain in a consistent state after any transaction. No transaction should have any adverse effect on the data residing in the database. If the database was in a consistent state before the execution of a transaction, it must remain consistent after the execution of the transaction as well [...]*”.

⁷² See DBMS – Transaction, Tutorials Point, id, mentioning that “[...] *In a database system where more than one transaction are being executed simultaneously and in parallel, the property of isolation states that all the transactions will be carried out and executed as if it is the only transaction in the system. No transaction will affect the existence of any other transaction [...]*”.

fails or restarts⁷³. So, deletion cannot be achieved and if it were, this would contrast the very properties with which a DBMS should comply.

It is worth mentioning that other systems do not adopt ACID. Most NoSQL systems are BASE (Basically Available, Soft state, Eventual consistency) compliant to increase availability and performance⁷⁴. This means the system guarantees data accessibility through data mirroring⁷⁵, i.e. real-time operation of copying data, as an exact copy, from one location to a local or remote storage medium. So, whether BASE or ACID compliant, the system does not “forget” data, albeit it permanently stores it, as it has been collected or as an exact copy –of the data collected⁷⁶.

Besides, users do expect several features, such as efficient and fast operation, enough history to roll back, capacity to control when and which data was changed, or back-ups, to enjoy a pleasant experience in a usable environment. Thus, it may be reasonable to claim that “*you can’t go out and remove content from everybody’s computer just because you want the world to forget about something*”⁷⁷.

7. Conclusions

We need to accept that Big Data is here to stay. And, perhaps, “data-scientist” should be regarded as the “sexiest job” of the 21st century⁷⁸. Indeed, these scientists

⁷³ See DBMS – Transaction, Tutorials Point, id.

⁷⁴ See Veda C. Storey, Il-Yeol Song, Big data technologies and Management: What conceptual modeling can do, in Data & Knowledge Engineering, Vol. 108, March 2017, pp. 50-67 (available at <https://doi.org/10.1016/j.datak.2017.01.001>); Abdullah Talha Kabakus, Resul Kara, A performance evaluation of in-memory databases, Journal of King Saud University, Computer and Information Sciences, Vol. 29, Issue 4, October 2017, pp. 520-525 (<https://doi.org/10.1016/j.jksuci.2016.06.007>).

⁷⁵ See Meikel Poess, Tilmann Rabl, Hans-Arno Jacobsen, Analysis of TPC-DS - the First Standard Benchmark for SQL-Based Big Data Systems, SoCC’17, September 25-27, Santa Clara, California USA, pp. 573-585, at p. 574, mentioning that a more relaxed version of durability is required, i.e. data accessibility. And to satisfy this “*a system must continue executing queries and data integration functions with full data access during and after a permanent irrecoverable failure of any single durable medium containing any database objects*”.

⁷⁶ See also Junliang Shu, Yuanyuan Zhang, Juanru Li, Bodong Li, Dawu Gu, Why Data Deletion Fails? A Study on Deletion Flaws and Data Remanence in Android Systems, ACM Transactions on Embedded Computing Systems, January 2015, Vol. 1, No. 1, Article 1, pp. 1-22, discussing data deletion failure with regard to Android Systems.

⁷⁷ See Matt Warman, Vint Cerf attacks European internet policy, 29 Mar. 2012, The Telegraph, available at <https://www.telegraph.co.uk/technology/news/9173449/Vint-Cerf-attacks-European-internet-policy.html>.

⁷⁸ See Kevin Daniel, André Carillo, Let’s stop trying to be “sexy” – preparing managers for the (big) data-driven business era, Business Process Management Journal, 2017, Vol. 23 Issue: 3, pp. 598-622, at p. 599, available at <https://doi.org/10.1108/BPMJ-09-2016-0188>; Thomas H. Davenport, D.J. Patil, Data Scientist: The Sexiest Job of the 21st Century, Harvard Business

are the ones able to make data speak in a business world that aims to complete its data-mutation and incorporate “analytics”⁷⁹ as an essential gene of all organizations’ DNA⁸⁰.

The so-called “datafication” phenomenon may cover all aspects of life and create new ways of communicating and cooperating⁸¹. And in data-driven businesses it is not about success⁸²; the goal is experimental failures that eventually lead to success⁸³. So, algorithms suffer from myopia⁸⁴, while an “algorithmic governmentality” aims to bring order to the chaos of the online world⁸⁵; make information easily manageable for the everyday user, the one who will be targeted and profiled. New ways of creating knowledge about future preferences, attitudes, or behaviors emerge in a data-behaviorism context, where systems, aiming at

Review, October 2012, Vol. 90 No. 10, pp. 70-76, at pp. 72, 73, mentioning that “[...] *More than anything, what data scientists do is make discoveries while swimming in data* [...]”, available at <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>.

⁷⁹ “Analytics” refer to the analysis techniques and tools that allow transformation of big data into actionable insights. See Steve LaValle, Eric Lesser, Rebecca Shockley, Michael S. Hopkins, Nina Kruschwitz, Big Data, Analytics and the Path From Insights to Value, MIT Sloan Management Review, Winter 2011, Vol. 52 No. 2, pp. 21-32, at p. 21, available at <http://www.ttvanguard.com/realtime/bigdata.pdf>.

⁸⁰ Kevin Daniel, André Carillo, Let’s stop trying to be “sexy” – preparing managers for the (big) data-driven business era, id, at p. 613.

⁸¹ Claudia Loebbecke, Arnold Picot, Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda, The Journal of Strategic Information Systems, Vol. 24, No. 3, Sep. 2015, pp. 149-157, at p. 149, available at <https://doi.org/10.1016/j.jsis.2015.08.002>.

⁸² Algorithms are useful marketing tools as they allow categorization, but there is no need for these profiles to be entirely accurate. Antoinette Rouvroy, The end(s) of critique: data-behaviourism vs. due-process, in Privacy, Due Process and the Computational Turn, Mireille Hildebrandt, Ekatarina De Vries (eds), Routledge, 2012, pp. 143-167, at p. 151, available at: http://works.bepress.com/antoinette_rouvroy/44 - https://www.academia.edu/7754445/The_end_s_of_critique_data-behaviourism_vs._due-process; Adam Harkens, The ghost in the legal machine: algorithmic governmentality, economy, and the practice of law, Journal of Information, Communication and Ethics in Society, 2018, Vol. 16, Issue 1, pp. 16-31, at p. 18, <https://doi.org/10.1108/JICES-09-2016-0038>.

⁸³ See, in general, Kevin Daniel, André Carillo, How to Engrain a Big Data Mindset into Our Managers’ DNA: Insights from a Big Data Initiative in a French Business School, 2015, DOI10.4018/978-1-4666-9649-5.ch005, in Anil Aggarwal (ed.), Managing Big Data Integration in the Public Sector, Chapter: 5, IGI Global. Available at https://www.researchgate.net/publication/284655215_How_to_Engrain_a_Big_Data_Mindset_into_Our_Managers'_DNA_Insights_from_a_Big_Data_Initiative_in_a_French_Business_School.

⁸⁴ Michael Luca, Jon Kleinberg, Sendhil Mullainathan, Algorithms Need Managers, Too, Harvard Business Review, No. 1, January–February 2016, pp. 96-101. Available at <https://hbr.org/2016/01/algorithms-need-managers-too>.

⁸⁵ Rouvroy uses the term “algorithmic governmentality” to refer to new regimes of power brought by the computational turn and the prevalence of algorithms in daily life. See Antoinette Rouvroy, The end(s) of critique: data-behaviourism vs. due-process, in Privacy, id.

“n=all”⁸⁶, rely on data, rather than users’ internal motivations⁸⁷. Conclusions are invisible to humans⁸⁸, whose data are valuable to firms: without it they would become bankrupt⁸⁹.

At the same time, individuals treat their data as wastage, as they waive their rights to access free digital services in websites, whose terms of service are never being read⁹⁰. This way, people are uninformed and ill-equipped to manage information and make innumerable decisions required by privacy self-management⁹¹.

So, we definitely need solutions of groundbreaking nature to deal with the above uncertainties. Instead of aiming at making machines forget our information, we could rely on data scientists, the ones able to design AI systems in ways that would benefit humanity and protect people’s fundamental rights. In fact, implementation of several principles into the very design specifications of systems could be promoted; trust and transparency rules could be embedded within data processing and analyzing procedures⁹², while tools could be used to keep systems

⁸⁶ Rob Kitchin, *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*, 2014, Sage Publications, London, at p. 72, mentioning that big data is “[...] *exhaustive in scope, striving to capture entire populations or systems (n=all)* [...]”.

⁸⁷ See Antoinette Rouvroy, *The end(s) of critique: data-behaviourism vs. due-process*, id, at p. 143.

⁸⁸ Mireille Hildebrandt, Bert-Jaap Koops, *The Challenges of Ambient Law and Legal Protection in the Profiling Era*, *The Modern Law Review*, Vol. 73, Issue 3, May 2010, pp. 428-460, at pp. 429-432. Available at <https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1468-2230.2010.00806.x>.

⁸⁹ Finn Brunton, Helen Nissenbaum, *Obfuscation: A User's Guide for Privacy and Protest*, MIT Press Scholarship Online: September 2016, at p. 50. DOI:10.7551/mitpress/9780262029735.001.0001. Available at <http://mitpress.universitypressscholarship.com/view/10.7551/mitpress/9780262029735.001.0001/upso-9780262029735>.

⁹⁰ See, amongst others, Simon Chesterman, *Privacy and Our Digital Selves*, September 7, 2017, *The Straits Times*, September 2, 2017, available at SSRN: <https://ssrn.com/abstract=3033449> (at p. 3, mentioning that “[...] *The British retailer GameStation gave us memorable proof of this one April Fool’s Day, when more than 7,000 people clicked “I accept” to terms and conditions that included the surrender of their immortal souls to the company. (The company later rescinded all claims, temporal and spiritual)* [...]”). See also Alexis Madrigal, *Reading the Privacy Policies You Encounter in a Year Would Take 76 Work Days*, *The Atlantic* (Mar. 1, 2012), available at <https://www.theatlantic.com/technology/archive/2012/03/reading-the-privacy-policies-you-encounter-in-a-year-would-take-76-work-days/253851/>.

⁹¹ Scott R. Peppet, *Regulating the Internet of Things: First Steps Toward Managing Discrimination, Privacy, Security, and Consent*, *Texas Law Review*, 2014, Vol. 93, pp. 85-176, at p. 159, available at <https://texaslawreview.org/wp-content/uploads/2015/08/Peppet-93-1.pdf>.

⁹² Similarly, the principle of Privacy by Design (PbD) has been established to guarantee that privacy and data protection are embedded within the “*entire life cycle of the technology, from the very early stage, right through to their ultimate deployment, use and ultimate disposal*”. Article

user-centric⁹³. In this context, lessons could be learnt from fields, where similar approaches have been conducted to implement fair information principles into technologies⁹⁴: in fields of biometrics⁹⁵, encryption has been proposed to put control over biometric data on the hands of the individuals and thus enhance their confidence towards the system⁹⁶; in fields of e-Health⁹⁷, emerging devices are

25 of the GDPR; Opinion of the European Data Protection Supervisor on Promoting Trust in the Information Society by Fostering Data Protection and Privacy, 2010/C 280/01, at par. 19. See also the seven principles proposed by Cavoukian in Ann Cavoukian, *Privacy by Design, The 7 Foundational Principles - Implementation and Mapping of Fair Information Practices*, available at <http://www.ontla.on.ca/library/repository/mon/24005/301946.pdf>.

⁹³ Daley speaks of a “crisis of cybersecurity” and argues that “permissionless innovation” cannot be a synonym for unaccountability. As Daley puts it, “*software development is hard [...] secure software development is very hard*”. John Daley, *Insecure Software Is Eating The World: Promoting Cybersecurity in an Age of Ubiquitous Software-Embedded Systems*, *Stanford Technology Law Review*, Spring, 2016, Vol. 19, No. 3, at p. 533. Available at <https://law.stanford.edu/publications/insecure-software-is-eating-the-world-promoting-cybersecurity-in-an-age-of-ubiquitous-software-embedded-systems/>.

⁹⁴ Anna Romanou, *The necessity of the implementation of Privacy by Design in sectors where data protection concerns arise*, *Computer Law & Security*, 2018, Vol. 34, pp. 99-110, at pp. 104-108.

⁹⁵ Biometrics can be defined as the technology that uses automatic personal recognition based on psychological or behavioral characteristics. See, in general, Ruud Bolle, Sharath Pankanti, *Biometrics, Personal Identification in Networked Society: Personal Identification in Networked Society*, Anil K. Jain (ed.), Kluwer Academic Publishers Norwell, MA, USA, 1998. Any human trait may be regarded as biometrics as long as it is universal (meaning every human being possesses it), distinctive (i.e. unique), permanent (i.e. it remains invariant for some period of time), and quantitatively measurable (meaning certain quantity is needed to be measureable). S. Prabhakar, S. Pankanti, A.K. Jain, *Biometric recognition: Security and privacy concerns*, *IEEE Security & Privacy*, Vol. 99, Issue: 2, Mar-Apr 2003, pp. 33-42. See also Article 4(14) of the GDPR (“[...] ‘*biometric data*’ means personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data [...]”).

⁹⁶ See European Commission, *Putting privacy at the heart of biometric systems*, 18 August 2011, available at <https://ec.europa.eu/digital-single-market/en/news/putting-privacy-heart-biometric-systems>; A. Cavoukian, A. Stoianov, *Privacy by Design Solutions for Biometric One-to-Many Identification Systems*, *IPC Technical Report*, June 2014, available at <https://www.ipc.on.ca/wp-content/uploads/2014/06/pbd-solutions-biometric.pdf>.

⁹⁷ E-Health can be defined as an emerging field in the intersection of medical informatics, public health, and business that refers to services and information delivered and enhanced through the Internet and similar technologies. Eysenbach regards e-Health as a commitment for global thinking to improve healthcare locally and worldwide. G. Eysenbach, *What is e-health?*, *J Med Internet Res* 2001, Vol. 3, No. (2), doi:10.2196/jmir.3.2.e20, available at <http://www.jmir.org/2001/2/e20/>.

designed to capture the minimum data required⁹⁸; in fields of video-surveillance techniques, smart cameras have been proposed to avoid discriminatory targeting⁹⁹.

So, by implementing fair information and non-discrimination principles into the very heart-design-specifications and the entire life cycle of AI, negative effects could indeed be avoided. To put it simply, instead of letting AI –find that e.g. female individuals perform worse than males and thus– exclude females and show the relevant ad to a male audience, why not let the system learn to recognize discriminatory targeting settings and not use those that are or can be directly discriminatory, while, at the same time, let AI learn when it could allow use of otherwise directly discriminating targeting settings?

To put it more simply, why not embed trust and transparency within “*sexy data-science*”?

⁹⁸ See, for instance, Intel’s device to better connect clinicians with patients at <https://www.intel.com/pressroom/archive/releases/2008/20081110corp.htm>; <http://web.nchu.edu.tw/pweb/users/arborfish/lesson/7736.pdf>; <http://thefutureofthings.com/6259-intel-health-guide/>.

⁹⁹ Andrea Cavallaro, Privacy in Video Surveillance, IEEE Signal Processing Magazine, March 2007, at p. 166, available at <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4117949>.