LUX
BARCELONA

METROPOLITAN TRANSPARENCY REVIEW

04 | *English edition* |

AMB

LUX BARCELONA

METROPOLITAN TRANSPARENCY REVIEW

AMB

LUX BARCELONA
Metropolitan Transparency Review, n. 04

Lux Barcelona is a journal promoted by the Transparency Agency of the Barcelona Metropolitan Area (AMB) that is published in three versions (Catalan, Spanish and English).

Advisory Board for Metropolitan Transparency
Joaquim Brugué Torruella
Manel Camós Grau
Victoria Camps Cervera
Karina Gibert Oliveras
Oriol Mir Puigpelat
Joaquim Triadú Vila-Abadal

The opinions expressed in the published works are the sole responsibility of their authors.

Director
Gemma Calvet, Director of the Transparency Agency

Coordination of contents
AMB, Transparency Agency

Publishing coordination
AMB, Editing Office

Graphic design and layout
AMB, Communication Service
Cecília M. Clos

Printing
Vanguard Gràfic, SA

DL: B 30269-2017
ISSN 2604-0956 (print)
ISSN 2604-0980 (electronic)

© photographs: Kim Manresa
© graphic illustrations: Jordi Duró
© texts: their authors
© Àrea Metropolitana de Barcelona (Barcelona Metropolitan Area)

Barcelona, November 2022

LEADERSHIP IN EQUALITY: THE DRIVING FORCE OF GOOD GOVERNANCE

The theme of the fourth volume of the magazine Lux Barcelona is feminism and gender equality policies, which are a fundamental principle of human rights and increasingly a key part of advanced governance models.

Despite full legal recognition of formal equality between women and men, there are still many cultural, social and economic barriers that prevent true equality and maintain situations of discrimination against women. For this reason, numerous public authorities are promoting gender equality policies for the fair distribution of rights, obligations, resources and opportunities, the eradication of existing inequalities, based on recognition and respect for the differences between women and men. In this context, Goal 5 of the Sustainable Development Goals (SDGs) of the United Nations 2030 Agenda seeks to achieve gender equality and empower all women and girls as the foundation for building a peaceful, prosperous and sustainable world.

With the aim of contributing to raising the visibility of on-going situations of inequality between women and men in many areas and taking measures to combat them, in 2017 the Transparency Agency promoted the Demeter Programme for transparency in gender equality, as an instrument to incorporate the gender perspective in open government policies and thus contribute to achieving co-responsibility and a 50-50 balance between men and women in companies, public authorities and civil society.

Since the launch of the Demeter Programme, 22 metropolitan city councils, five organisations associated with the AMB and one company providing metropolitan public services have joined. Within the framework of this programme, several studies and workshops have been conducted to obtain a more thorough diagnosis of certain aspects of gender inequality in the metropolitan area, leading to the drawing up of recommendations that include several specific measures to promote effective equality between men and women. These recommendations, which propose several courses of action to strengthen equality and the feminisation of metropolitan leadership, are one of the central pillars included in this fourth volume of the Lux Barcelona magazine.

As part of the Demeter Programme, in October 2020, the Transparency Agency set up a working group of public and private sector professionals committed to gender equality to develop a diagnosis of the resistance and obstacles to female leadership in organisations and identify good practices for the promotion of equality. The results of the work of these experts were set out in a summary

document entitled "The Role of Leadership for Egalitarian and Ethical Organisations", which is published in this fourth volume and is authored by Sara Berbel, Gemma Calvet, Àngel Castiñeira, Oriol Estela, Vittorio Galletto, Jaume Garcia, Maria Eugènia Gay, Anna Gener, Pilar Molina, Carme Poveda and Juliana Vilert.

Along the same lines, in October and November 2020, the Transparency Agency commissioned the 50a50 Association to organise a series of workshops to reflect on various aspects affecting gender equality in organisations: the legal framework, implementation of equality plans, industrial relations and political and managerial leadership. As a result of this series of workshops, the 50a50 Association produced a final report entitled 'The Challenge of Equality in Organisations', which presents an analysis of the weaknesses, strengths, threats and opportunities of the feminisation of organisations and concludes with a number of specific recommendations to improve the awareness and training of society and managers in order to move towards full equality. Given the interest of this analysis, this report has also been included as a central pillar in this volume.

However, gender equality has many other dimensions beyond gender balance in public institutions, business and civil society. Francesc Torralba, Núria Tria, Mar Rosàs and Guillem Martí publish the study 'Algorithmic Ethics and Gender Perspective: from Opacity to Transparency' as a central pillar. It is an extremely interesting research paper on gender bias in artificial intelligence algorithms, which includes several proposals regarding advertising, transparency and accountability to ensure that ethical principles are met and gender discrimination in algorithms is prevented. This research project was funded by the Transparency Agency's line of grants to promote research in the field of transparency, public information and good governance.

The publication of these four articles as central pillars for this fourth volume of Lux Barcelona is complemented, as always, by three short articles under the heading 'Lighthouse', which focus reflections on ideas related to good governance of the metropolis. In the first one, the municipal manager of Barcelona City Council, Sara Berbel, presents the organisational measures promoted by the city government to foster gender equality in municipal administration. In the second one, Maria Teixidor, lawyer and former member of the Board of

Directors of Futbol Club Barcelona, addresses the feminisation of a sector as male-dominated as football, looking at the case of the professionalisation of the Barça women's team. In the third lighthouse, the journalist Natza Farré takes a real case of gender-based wage discrimination to argue that the wage gap, economic violence, job insecurity and the unequal distribution of domestic tasks are manifestations of women's lack of freedom and raises the need to bring about true equality in economic power between men and women as a driver for peace and prosperity.

Following the line of the previous volumes of Lux Barcelona, this fourth volume intersperses six snapshots of the metropolitan landscape, taken by the renowned photographer Kim Manresa, and three illustrations by the graphic designer Jordi Duró, addressing gender equality from his particular perspective.

All this is material that the AMB Transparency Agency makes available to the public to provide new ideas and knowledge related to feminism and gender equality, to promote critical reflection among the public and raise awareness of the challenges posed by achieving effective equality between women and men.

The Demeter Programme is therefore a useful instrument for aligning the Transparency Agency, the AMB and all the member city councils and bodies with Goal 5 of the United Nations 2030 Agenda. Unfortunately, there is still a long way to go in achieving equality in power leadership, a victory that still eludes women.

Gemma Calvet Barot

Director of the Transparency Agency
of the Barcelona Metropolitan Area

Badia del Vallès. Time Banking Festival
© KIM MANRESA

EQUALITY AND INNOVATION IN BARCELONA CITY COUNCIL

It has been over 15 years since Organic Law 3/2007 on effective equality between men and women was passed, and almost 7 since the passing of Catalan Law 17/2015, yet figures show that equality is still a long way off. The International Monetary Fund notes that, at the current rate, it will take 200 years to achieve equality in Western countries. What can we do to drive this necessary and just movement forward?

The first requirement is ensuring regulations on equality are in place. Barcelona City Council, which employees 16,000 people, has implemented some of the very latest instruments for driving equality policy in municipalities. But this is not enough. Establishing stable foundations for gender equality in organisations requires profound institutional change, over and above regulations and procedures. It requires political will, good executive governance, a balanced presence of men and women, resources and innovation to work effectively, together with staff awareness and constant training.

In this sense, through the Municipal Manager's Office, the council has promoted a number of organisation-wide actions based on the 3rd Equal Opportunities Plan for Men and Women (2020-2023), agreed on with union representatives, and through the leadership of the Deputy Mayor's Department of Social Rights, Global Justice, Feminism and LGBTI; the Gender Justice Plan (2021-2025) and the Gender Equality Regulation.

Economic and structural actions have been essential in ensuring funding for gender equality promotion policies, which already receive 4.52% of the current expenditure budget (we are now close to the target of 5%). And 1% of the budget will eventually be spent on creating and strengthe-

ning structures to drive these policies. In addition, municipal taxes, levies and public prices are analysed from the gender perspective.

Mention should be made of the inclusion of equality officers assigned to the Council areas and districts and to municipal companies (already established in five public-private partnerships and two autonomous organisations) and specialists to tackle gender-based harassment and violence.

Despite all this, in the City Council there is still a significant glass ceiling for women (those invisible, but nevertheless real, barriers that stop professional advancement), and the so-called career quicksand: domestic and family responsibilities culturally taken on by women forcing them into more precarious and poorly paid sectors. Combating these phenomena is the key to ending the pay gap. This is why we have implemented equal pay for the top executive positions, the managers, established quotas in male-dominated areas such as the local police force and the Fire Brigade and created specific jobs to provide career opportunities for women in more feminised and precarious sectors.

Training and awareness raising are constantly provided, as they are essential to ensuring all the staff fully understand of the measures. Other projects, such as the new Directorate for Values and work on creating new leadership models, also contribute to creating a more horizontal, less hierarchical and more equal organisation.

In doing this, we never lose sight of our goal, which can be summed up perfectly in the words of our great teacher Simone de Beauvoir: 'Let nothing limit us. Let nothing define us. Let nothing hold us. Let freedom be our own substance.'

Sara Berbel Sánchez

Cerdanyola del Vallès. Alba Synchrotron
© KIM MANRESA

WOMEN'S BARÇA, A TRIUMPH ON THE ROAD TO EQUALITY IN SPORT

The professionalisation of the Barça women's first team marked a clear turning point for the Club at which I was a director for five years, before resigning in April 2020 for matters of transparency. This commitment provided further substance to the Club's motto *més que un club* (more than a club), making it a powerful voice for the changes demanded by today's society. I would like to provide a brief description of how we managed to achieve our goals of winning the Champions League in our 6-year term of office and becoming world leaders, as it may provide a source of inspiration for other projects.

Firstly, leadership. The Board of Directors grasped the moment and included the professionalisation of the women's team in its strategic plan for 2015-2021. Such unconditional support for the project from the governing body was essential to its development and, in the case of Barça, it was forged not only with the professionalisation of the women's football team but also through the work of the Edelmira Calvetó Group, where I was president, and reinforced government action in restoring the historical memory of Barça women, creating synergies among members and opening up the female side of the Club to civil society.

Secondly, the sports team. The transition from amateur to professional meant increasing staff numbers from the initial 3 to over 20: a manager, a second coach, analysts, a goalkeeping coach, physiotherapists and a doctor, a psychologist, a delegate and a communications officer. It also meant selecting the best footballing talent and setting up all the youth teams to ensure a generation of home-grown talent steeped in the Club's methodology from the start.

Thirdly, resources, besides funding: remuneration for the football players higher than the minimum interprofessional salary, ensuring they no longer required second jobs to get by; specific sponsors; morning training sessions; lunch and dinner at La Masia training school; training camps on grass or good-quality artificial turf pitches; performance tracking; launching a new app to track menstrual cycles and adapt training to them; and building the Johan Cruyff stadium, with a specific women's dressing room (along with another for men and a third for the away team).

And fourthly, collateral supporting actions. Themed conferences for 8 March aimed to breaking down prejudices regarding sport, women, brands and image among journalists and civil society; restoring the historical memory of women at the Club, such as the first female Barça member (Edelmira Calvetó) and the first female director (Anna Maria Martínez Sagi), and recognition of female Barça supporters and women from the world of sport who have helped make the Club great (Edelmira Calvetó awards); and organisation of the first mixed summer tour in the history of football (United States, summer 2018).

Strength of purpose, vision and teamwork are required to ensure true women's empowerment" is achieved. The feminisation of FC Barcelona shows us how cultural changes are possible with the right tools and leadership, responsibility and commitment, even in such masculine and disputed environments as football.

Football has over 4 billion followers worldwide and incredible potential to penetrate society and change dynamics, beliefs and culture. Men's football, with its constant controversies over transfers (with million-dollar commissions paid to third parties), state-owned clubs and unsustainable finances, faces numerous challenges

that could be turned into opportunities to demonstrate how transparency, honest management and hard work can produce more human, sustainable and prosperous systems. Governments, supranational institutions and non-governmental organisations with power in the world of sport need to commit to bringing back the values still maintained and represented in the female game to global football. In making this demand we also seek to continue along the path toward the inclusive, balanced, sustainable and healthy society that we and future generations deserve.

Maria Teixidor Jufresa

Corbera de Llobregat. The Bandit Market
© KIM MANRESA

A NEW ERA FOR HUMANITY

In 2011, Jill Abramson was appointed editor of *The New York Times*, the first woman ever to achieve this. Three years later, in 2014, she was abruptly dismissed. The newspaper's owner, Arthur Sulzberger, justified doing so by saying Abramson had a nasty character. I know nothing about her temperament, but am convinced all her predecessors were affable and mild-mannered, like most executives. The unofficial reason for her sacking is much more convincing. They say that Abramson found out that the previous editor earned more than her so she wanted to renegotiate her salary. Her fight for equal pay was sufficiently unwelcome that she was kicked out immediately. Her audacity was harshly punished. Abramson has long recommended that women being promoted should always ask what their salary will be and how much the person they are replacing was earning. This is advice we should all follow and pass on.

The risk of receiving less pay for doing the same work as a man is not unsubstantiated. It is as real as the stubbornly persistent wage gap that turns women into victims of economic violence that cannot be disassociated from physical, psychological or institutional violence. They are all part of the same whole. The relevance of economic power in a society such as ours is so clear, one inevitably concludes that if women were equal participants in that economic power, the world would be a radically different place. Indisputably better. I am certain of this.

The unfortunate experience of a global pandemic has highlighted what we already knew: women's work is essential in all circumstances. But this indispensable work is poorly paid and in many cases simply not paid at all. How long have we been like this? Forever. How much longer do we want to stay like this? Not another minute. Are we on the right track? No. The forecasts are dire, but must always be read with the idea that they are reversible and, with the conviction that they do not necessarily need to be fulfilled. Such indicators are often designed to discourage us from trying to turn

them around, acting as a brake on motivation and providing stability for those who resist change. Nevertheless, the World Bank estimates that at the current rate it will take 257 years to close the gap, a figure that increases with every year that passes. As we've already mentioned. It is a way of saying the pay gap will NEVER close. By contrast, estimates on what would happen if it disappeared suggest global GDP would increase by between $12 trillion and $28 trillion over four years. What makes people want to avoid improvement? What is this age-old fear of equality? The paradigm shift would be so dramatic it is bound to generate anxiety, especially if you happen to live on the privileged side of town. Fortunately, there are antidotes to this anxiety. And unfortunately, the side of town without such privilege is far bigger.

In times of crisis, women get paid less and lose their jobs sooner because most of the burden of precarity falls on them and because work-life balance is a matter for mothers, not fathers. Now, well into the 21st century, the dichotomy between family and career continues to tip the scales toward women in the vast majority of cases. They are the ones who ask for leave and shorter working hours, which translates into less money for women and zero money for care. The work that women do, which is essential for the survival of humanity, is not included in paid work. Men, on the other hand, can increase their wages through overtime because they have someone to cover for them at home. In other words, they can proudly enjoy their family, while it takes up much less of their time. The glass ceiling remains unbroken because of historical mistrust of women's abilities, women's own lack of self-esteem and the painful choice imposed on them between their public and private lives. Women are asked to do much more just because they are women, often wondering if the effort is worth it. Moreover, at the very moment when feminism has become a mass political movement, demanding positive discrimination to break the dynamics of systematic discrimination against women, the most reactionary sectors rush to defend the meritocracy in order to kill off a quota system that is essential to overcoming the inevitable impasse. I can think of a very

long list of men lacking all merit who nevertheless have held, hold and will hold positions they are not and never will be qualified for, but who have not been, are not, nor never will be questioned in their position, because they give each other mutual slaps on the back that we women neither give nor receive. Feminism shows in whose hands privilege lies, and those hands run to grab the advantages knowing they are unfair to their mothers, sisters, wives and daughters, i.e., the women closest to them, and who they use to defend themselves against accusations of sexism, as if having a mother, sister, wife or daughter provided them with some sort of exemption. It is a weak argument that collapses at the slightest touch.

Economic violence against women is relentless. Women's inability to access the cash box directly without going through men has generated a dependency that for centuries has subjected us to the absolute control of men. That is why it is so important to break out of this dynamic and start exercising control over ourselves and our own possibilities. This means we must generate a stable network that capitalises on women's knowledge to produce a new economy where we decide how to manage it and who benefits. Women's economic power is intimately linked to their freedom, just as today inequality is linked to their lack of freedom. 'Money makes an ugly woman beautiful', 'if you want a woman to love you, give her all the money, let her do what she wants and never say anything to her' are just a couple of examples of traditional ideas on women's relationship with money. As always, discrimination is accompanied by popular culture, which is responsible for transmitting the values on which our Western societies are built. These competitive values are what generate unhappiness in these societies and, as Linda Scott, author of *The Double X Economy*, says, 'If the global community chose to dissolve the economic obstacles facing women, an unprecedented era of peace and prosperity would follow'. This era must begin now.

Natza Farré

El Papiol. Cherry Festival
© KIM MANRESA

The ethics of algorithms and the gender perspective: from opacity to transparency

Francesc Torralba, Núria Tria, Mar Rosàs i Guillem Martí
(Ethos Chair – Ramon Llull University)

1. Introduction

In personnel recruitment processes, organisations are increasingly using artificial intelligence (AI) solutions based on the programming and implementation of algorithms: logical systems that automatically analyse documents such as CVs and other datasets to identify, in the blink of an eye, the most relevant candidate profiles for a specific job.

Initially, if we put ourselves in the position of the human resources departments (or companies) responsible for carrying out these recruitment processes, this would appear to entail two significant advantages for both the organisations and people looking for a job.

The first advantage, which is ethical in nature, is that the use of said technology would seem to guarantee that all applicants will be subject to the same criteria. Thus, processes that until recently ran the risk of being biased by subjective factors are now characterised by their objectivity and standardisation.

The second advantage is that the tasks of pre-screening applicants or job postings – which, until recently, took up immense amounts of time – is now automated, albeit not completely, at a number of important stages. At the very least, algorithms provide an initial filter. For organisations, the filter screens applicants passing on to a second phase of the recruitment process, guaranteeing that only those best-suited for the position progress, without having to spend time scrutinising their CVs. In the case of jobseekers, they screen the number of situations vacant they engage with, supposedly ensuring that they only deal with those for which they are best qualified.

Note 1: This article was first published in Catalan in 2021 in issue 27 of the journal *Ars Brevis. Anuari de la Càtedra Ramon Llull Blanquerna*, pp. 211-282.

Note 2: The research for this article obtained a competitive grant from the '3rd Call for applications for grants for research in the field of transparency, access to public information and good governance 2021' from the Barcelona Metropolitan Area (AMB) Transparency Agency, under the research sub-line 'Personnel recruitment and transparency'.

Nonetheless, the use of algorithms in recruitment processes may end up being discriminatory. This is because algorithms are never axiologically neutral, but are in fact programmed in accordance with a series of criteria that someone has decided upon. Thus it is that these criteria can replicate human biases. Added to this 'external' source of potential biases ('external' in that they are introduced in the design processes by those responsible for programming) is the creation of biases – or, at least, the amplification of their effects – by the very self-reproductive logic of these technologies, which are based on training and learning.

The consequences of this are especially serious in the case of gender, particularly if we bear in mind that personnel recruitment processes are already likely to be affected by other forms of gender discrimination. For example, it is common to see women needing to be far better qualified to receive the same rating as men in CV assessments[1] or, in interviews, interviewers tending to perceive men as more competent than women.

The goal of this study is not to argue against the use of artificial intelligence solutions. However, it is important to highlight those regards in which, for example, the automation of recruitment processes could give rise to certain discriminatory logics that may run counter to the Universal Declaration of Human Rights, which forms the basis for contemporary democratic societies' ethical and legal frameworks. So, the goal of this work is to reduce the opacity that frequently surrounds the social application of algorithm-based technologies and contribute, from an ethical standpoint, to the debate on how to ensure that algorithms do not foster discriminatory decisions and gender biases.

So it is that the object of study of our research is the potential gender biases arising from recruitment and selection algorithms. Additionally, our work on this has meant we have been able to formulate a series of suggestions for preventing potentially discriminatory effects of the use of algorithms in staff recruitment processes.

In Part I of our study, we identify and summarise the main theoretical challenges of the ethics of algorithms, particularly with regard to gender, and provide some real-life case studies of both bad and good practices, by way of example and to aid understanding of the issue.

Firstly, we'll provide an introduction to the idea of artificial intelligence and to the ethics of algorithms. One of the main challenges we find in tackling this issue is the inherent difficulty in understanding the subject, added to which is a degree of terminological confusion resulting from the sensationalist and simplistic treatment given to it by the media. There is therefore a need for a conceptual approach that smooths the way for the subsequent tackling of the ethical challenges.

1 In some cases, two and half times better qualified than men (*Libro Blanco. Situación de las mujeres en la ciencia española*, Unidad de Mujeres y Ciencia, p. 14).

Next, we shall consider how a standpoint based on some of the main theoretical contributions made by feminist theory, queer theory and gender studies can help us to detect gender bias (and, indeed, other kinds of bias) in algorithms in general. These theories will be useful in detecting and assessing the different types of discrimination in algorithms and in conceiving of and building ways and strategies to try to prevent them or, as the case may be, mitigate the harm that, as shown by the literature we have consulted, the increasingly widespread use of AI is causing many people in a number of areas of life (and particularly in the field we are examining here).

Lastly, we shall take an in-depth look at the particular contexts of the world of work and, more specifically, personnel recruitment and selection processes, with the wish and aim of understanding what happens there when they begin to implement automated processes and make decisions based (in whole or in part) on AI.

To complement this general explanation, we shall be providing real-life case studies illustrating some of the ethical challenges faced by companies and other organisations that have implemented different kinds of automation in staff recruitment processes. A number of these cases have caused quite a stir in the media. Others – less well-known, perhaps, yet nonetheless with a real impact on people's lives – have been the subject of scrutiny by conferences, working groups and academic journals. We will be including each of these at the exact point where they can be most of the greatest help in clarifying the issue at hand.

All this information will fuel Part II of this study, in which we set forth a series of measures to help detect/prevent/mitigate gender biases and their effects upon staff recruitment and selection processes based on or supported by algorithmic technologies.

This second part is based on the idea that it is not, from an ethical viewpoint, permissible to assume that the growing trend of automation in staff recruitment and selection processes through AI-based technology must inevitably entail opacity in said processes, a kind of 'black box' effect that hinders transparency and traceability and that therefore negatively impacts people's rights.

To prevent this from happening, there is a need for some guiding principles and a set of measures, both general and specific, that can be utilised to prevent or, if this is not practicable, at least counter or mitigate these undesirable effects of the use of AI-based software.

While the biased (discriminatory) practices of the traditional (human) recruitment system may be clear, at least to a certain degree, and combatted (again traditionally) with the help of trades union, appeals to regulatory bodies or legal action, the danger and inherent problem with algorithms is that these biases are 'buried', and thus go unnoticed, unquestioned and unchallenged.

There is therefore a need for legal frameworks, many of which are still under construction, to govern these practices, particularly those associated with oversight of the private sector: that is, the tech industry.

However, and despite the apparent complexity of the issue, a significant amount of the challenges under review here can be appropriately addressed with some suitable ethical standards that can help us avoid an exclusively or excessively technological response.

To make this ethical governance of the digital environment both possible and effective, there is a need to provide measures at different levels. Such an approach will help provide coverage for the wide range of situations that may be found by those responsible for an organisation's recruitment and selection processes.

What's more, implementing the principles and measures we propose here could help avoid the potential 'reputational risk' arising from an undertaking involuntarily or negligently committing discrimination in their AI-based staff recruitment processes. This is an ancillary benefit arising from our suggested measures and one which we believe is not entirely without interest.

1.1. Methodology

To prepare for this study, we have carried out a critical bibliographic review of academic literature around the issue sourced from the usual databases: Scopus, Web of Science and Google Scholar. We have used the following inclusion criteria: i) the applicability and relevance of the topic, ii) that it be a peer-reviewed article or a paper given at an international conference, iii) that it date from the last 5 years (i.e. from between 2016 and 2021), iv) that it be written in English, Spanish, French, Italian or Portuguese, and v) that the full text be available. After eliminating duplicates, we collated a total of 51 documents (11 from Scopus, 17 from the Web of Science and 25 from Google Scholar), which we then analysed. Worth noting is the fact that, although the period under consideration was five years, the bulk of the texts appeared in the last three. In this latter period, an exponential increase in the academic output on the topic can be seen, in terms of both articles and conference proceedings. We believe that this is due, firstly, to the fact that this is a relatively new technology only now being introduced into society and, secondly, potentially due to the media impact of the so-called 'racist algorithm' case seen in Amazon, which we shall examine later on and which, if nothing else, has helped focus public interest and academic debate around this issue of the ethics of algorithms.

With regard to the sources selected, we felt that it was important not to limit ourselves exclusively to strictly academic texts (i.e. peer-reviewed academic articles), so as to also give a voice to the output from different stakeholders. A systemic focus, as justified later on, makes this essential. This is why we have also included papers from conferences, which feature a far greater input from the tech sector itself. Although the proceedings review process tends to be a little more forgiving, we believe that this possible lack of thoroughness is more than offset by the

increased diversity of viewpoints thereby included in the sample, an important aspect that also fits neatly with the goals we have set ourselves here.

The selection from the databases was carried out using the keywords and search strategies set forth in Figure 1. The search strategy is shown by means of generic formulations, in which we ignore strict syntax and some of the Boolean operators to make things more accessible and understandable. The terms between brackets are the synonyms. By related terms, we mean terms more specific than the general search term, with which they are related, which are relevant to one or more of the sub-fields of AI, and which we have used as alternative forms to the main keywords.

Figure 1. Keywords and generic search strategy

Artificial intelligence (AI, algorithm*) Related terms: algorithmic decision-making, automation computers decisions, decision-making automation, Information Retrieval Systems, Machine Learning, Search Engines, Word Embeddings, Deep Learning, Natural Language Processing, Text Analytics, AI-based interview, Virtual Work Environment, Aumented Intelligence
AND **Algorithm* bias** (bias in AI, fairness bias, algorithmic fairness, bias in digital ratings, discriminat*) Related terms: implicit bias, indirect discrimination, implicit association, prejudice, «invisible cage», embodiment, fairness, trust, perceived fair, emigrants and immigrants, ethnic groups, minority groups, stereotyping, racism, social justice, inequity
OR **Gender bias** (gender bias, gender inequality, sexism, gender neutral, implicit gender bias) Related terms: gender role, gender stereotypes, women's rights, male, female, gender-non-binary, LGTBI*
AND **Recruitment** (AI recruitment, automation in HR, hiring algorithm, automated hiring, algorithmic hiring, HR analytics, talent search, talent acquisition, recommendation systems, ranking algorithms) Related terms: online recruiting, online job advertisement, job postings, employment, occupations, Personnel Selection, Human Resources, recruiting, selecting, hiring, management, job opportunities, STEM careers, IT workers, platform workers
AND/OR **Mitigate** (algorithm auditing) Related terms: data protection, evaluation, control, opaque, transparency, impact assesment, accountability, ethics, law, discrimination law

PART I. The ethics of algorithms and the gender perspective

2. An introduction to artificial intelligence and the ethics of algorithms

2.1. Some conceptual clarification

We define the ethics of algorithms as that branch of ethics, applied to technology, which explores and seeks to resolve all those issues arising from the programming of an algorithm.

One paradigmatic example of this would be the programming of self-driving vehicles to deal with an unavoidable accident. How should a vehicle react when confronted with the dilemma of choosing between *x* and *y*? How should it weigh up the risks faced by the different parties involved?

Self-driving vehicles are just one example, but the majority of algorithms are subject to parameters and default options that are of great moral importance.

Philosophers are great fans of categories, definitions and, in short, concepts. Indeed, one field of applied ethics consists in creating categories that help us understand and, potentially, transform reality. The notion of the ethics of algorithms helps us embark upon this conceptual analysis. To begin with, the expression gives us a scope of application: the set of moral issues posited by programming.

This area can be placed within a sub-area of AI ethics-related issues associated with what is good, fair or virtuous to do in AI systems. The ethics of AI are, in turn, a branch of the ethics of technology (or technoethics) or, more precisely, of the information technologies.

However, the ethics of algorithms are very close to what has been dubbed robot ethics or machine ethics. The benchmark for this branch of ethics is provided by Wallach and Allen's book *Moral Machines*. In this essay, published in 2008, the authors laid the foundations for a question that had previously not been considered in practical philosophy: how can we develop computational morality?

The ethics of algorithms can be distinguished from those of robots because they are more all-encompassing. An algorithm does not necessarily require a robot host any other specific form of embodiment to raise moral questions. Any research, recommendation or translation application can, in fact, be subject to a moral assessment. The question is always the same: How to act within the program and the algorithm to meet moral standards?

Indeed, the ethics of algorithms are separated from machine ethics by just a few connotations. They break with the unique vision we have of machines.

Machines spontaneously appear as individualities, artefacts, moral patients. Nevertheless, from an ontological viewpoint, it may perhaps be better to talk in plural and free ourselves of this unified concept of machines. It may be better to regard machines equipped with AI as specialised entities that remain cognitively opaque.

At the heart of the problem lies not the artefact, nor its form, but rather the programming installed in it. From a moral philosophy standpoint, robots and machines are nothing more than the wrappings surrounding algorithms.

The ethics of algorithms encourage philosophers to examine internal systems and understand them in detail and with transparency. To dedicate oneself to the ethics of algorithms is to concentrate on a very particular field of knowledge but also at a very special, highly uncommon scale. How does this shift in scale manifest itself? What makes the ethics of algorithm different from the ethics of AI that encompasses it?

Let's delve deeper with some practical questions.

— Should self-driving cars be allowed in cities? The ethics of AI.

— How should a self-driving vehicle be programmed to deal with unavoidable accidents? The ethics of algorithms.

— Should sexbots be allowed? The ethics of AI.

— Should such robots provide the option of simulating resistance to the other's advances? The ethics of algorithms.

— Under what conditions should a robot have rights? The ethics of AI.

— Should a dating app automate or reduce certain kinds of discrimination? The ethics of algorithms.

While the ethics of AI holds that AI systems should not foster discrimination, the ethics of algorithms asks how we should translate these general principles into computer code, since what they deal with is exactly that: how to encode morals. Algorithms require that extremely precise and transparently clear decisions be taken. This means that, in the ethics of algorithms, there is no room for doubt or hesitation.

This kind of study requires us to think about the models of justice available and apply one of them. *Technology Review* published a good example of the issue, referencing the case of COMPAS, a criminal reoffence prediction system criticised for unfairly dealing with black people.

The ethics of algorithms cover more than just programming. They have a top-down perspective. While those of AI favour a bottom-up, more macro perspective, close to technoethics, the ethics of algorithms put the focus on the details, on the microcosmic scale.

It is no wonder that analytic philosophers express their affinity with the ethics of algorithms requirements for transparency and precision, while continental philosophers feel more comfortable with the broad strokes of the ethics of AI.

Both approaches are necessary. While the ethics of algorithms explores the internal workings of machines, the ethics of AI is interested, broadly speaking, in humankind, its history, its surrounding environment and its relationships: in short, in everything that the implementation of artificial intelligence may alter and transform.

It's easy to see how the dividing line between these two fields of study is blurred. If it is possible to have satisfactory moral programming for robots, a subject studied by the ethics of algorithms, then there is a good reason for building them, an issue that is the field of the ethics of AI. However, what happens if this is not the case?

An application that may appear to be completely necessary may exert a pressure on decision-taking in the ethics of algorithms, thus giving rise, for example, to an anti-pandemic app that prioritises security over respect for privacy. The moral application of its design belongs to the field of the ethics of algorithms, while discussions around the tension between the right to security and the right to privacy are the purview of the ethics of artificial intelligence.

The dividing line is by no means clear. Indeed, the ethics of AI and those of algorithms represent two poles linked together by a continuum rather than two distinct fields.

Take, for example, the voices employed by personal assistants. Should they be, by default, female, male or neutral? Should they seem young, mature or elderly? Should they imitate the human voice or should they have a metallic, robotic sound that is distinctly different? What tone of voice should they have?

All these issues impact the design of a personal assistant and how it interacts with human interlocutors. They are relatively clearly delineated. However, they also depend upon a set of very general considerations, such as what type of robots we wish to live alongside.

Put another way, a decision that may appear limited solely to the design of a social robot's voice presupposes a global thought process, at the scale of the ethics of AI. Programmers understand algorithms much better than the ethics of algorithms. The latter area lies outside of their bailiwick, and hence the need for the interdisciplinary work of philosophers, engineers, mathematicians and sociologists.

Machines equipped with artificial intelligence have to take decisions. They generally have to establish a default option. The fact is that most users do not change the factory settings. So, it is worth asking whether the female voices of Siri or Google Home simply reinforce the stereotype associating women with domestic service. Similarly, we should ask ourselves what happens in people's brains when they associate a non-biological entity with a gender. It might lead to the conclusion that it is not a good idea to reproduce the male/female duality in inanimate devices.

Should certain norms be automated?

AI-equipped personal assistants and other robots have not appeared in a morally untouched and pristine world. They exist in a world full of ethical challenges and (let's be honest) beset with injustice. This means it is also full of social hierarchies, stereotypes and all manner of implicit biases. A significant part of the work of the ethics of algorithms entails identifying these and asking whether they should really be automated and reproduced in the smart machines we use.

It appears difficult for programmers to shun this responsibility. Just as with the default option, the choice of menu is – already and in itself – a morally-charged decision.

In France, for example, voice assistants have a French accent and, in Quebec, a Quebecois one. But what actually is a "French" accent? That of Toulouse, Marseille or Poitiers? No, it is that of capital rather than any provincial one. It is not a working-class accent, but that of a television presenter. However, who is to say that it is not a good idea, in France, for Alexa to answer in a Quebecois accent and for Siri to have Haitian one? Who knows whether this would make life easier for Quebecois immigrants in France and for Haitians in Quebec? At the very least, users could be offered the choice. This could be one option among a dozen or so accents, and is a menu-related challenge.

Choices need to be made when programming, and it is by no means clear that AI systems do so on the basis of the rules and categories at play in the real world.

Another specific area of the ethics of algorithms is associated with autonomous devices. This often entails much more than programming a simple reaction to a signal, such as smoke activating a fire alarm system. It involves equipping the system with the ability to take decisions by integrating a number of sources of information.

The artificial intelligence that played the game Go adapted to its adversary, and was able to react with surprising moves. But what to do with more complex and morally impacting behaviour? How should an autonomous or partially autonomous entity be programmed?

This means that one cannot simply implement a behavioural rule. A whole set of metarules also needs to be incorporated. Self-driving vehicles can help us illustrate this issue: they obviously need to be programmed to stop at a red traffic light, but the system can also be smarter than that. The vehicle will need to jump the red light if this means preventing an accident. It is here that the notion of applied autonomy arises in artificial intelligence. Morally programming a self-driving vehicle means giving it the ability to decide for itself. To an extent, it entails delegating the decision-making process to it.

However, this delegation raises many questions. If a robot is programmed, how it is autonomous? Can one really speak of autonomy in such a case? Is it not a contradiction in terms?

This latter question may also be valid for biological entities like human beings. Are we programmed? And, if so, are we really autonomous (i.e. free-willed) or do we just appear to be? The old debate between determinism and indeterminism sneaks back onto centre stage.

A car's AI does not really need to be free – in the philosophical meaning of the term – to be in a position to decide whether or not to respect a red light. All it needs is to be installed with some metarules and an ability to predict what will happen.

Can we say that programming an autonomous entity is, more or less, like educating it? Is this the right word or should we keep the word 'educate' solely for referring to human beings? Can you educate animals? Or machines? Is educate the same as train? This semantic back and forth is inevitable when we analyse these kinds of issues philosophically.

Can self-driving vehicles be taught to carefully negotiate curves when the going is icy or to avoid running an animal over as it crosses the street? Can it be taught to resolve tragic dilemmas such as choosing between saving a small child or a senior citizen?

A number of experts believe that the ethics of algorithms are, in a way, a form of artificial moral education. For some, educating a self-driving vehicle and educating a child are not so very different, even if they are not identical. In both cases, the aim is to prepare them to think for themselves, without the help of their engineers/parents, so that when one is on the road and the other has to face important decisions in the future, they will know what to do.

A human's life is full of crossroads and unforeseen events. Children are equipped with values and criteria, but it is they themselves who, faced with a given situation, must assess what is fair, right and sensible to do. The same is true, despite all the differences, with a self-driving vehicle. Once it has been fed with its general criteria or metarules, the device will move under its own steam and will have to take decisions in situations that have not been contemplated by its engineers on the basis of this set of criteria. As Michel Montaigne put it, 'vaut mieux une tête bien faite qu'une tête bien pleine'.

When it comes to educating a young child, however, we understand that there is a need for them to develop critical thought and a sense of fair play. But can the same be done with a robot? Can a robot be critical of the very program installed in it? Will it be capable of distancing itself from itself as an actor and become an observer of its very self? In short: can a robot's moral programming be compared with a child's moral education?

This drawing of parallels obviously has its limits. A young child is subject to a weakness of will (acratia): a robot is not. The latter can receive, in seconds, new moral programming within an AI system, while a child's moral education takes a long time and a large number of properly coordinated actors all sharing the same aim.

Algorithms can be controlled perfectly, as they do no more than what they have been ordered to do. This is not the case with a child, nor with the human condition in general. There's room for transgression, disruption, infringement of legal norms. As Georges Bataille put it, being seduced by the prohibited is part of human nature itself.

The archetypal tale in Genesis provides a good example of this. God creates Adam, but does not program him to simply follow orders, the rules the Creator has given him. He makes him free, including the possibility of disobeying Him, of transgressing the law and eating the forbidden fruit.

Can machine learning be linked with the teaching of these ethical criteria? This AI technique permits the completion of a task, not by following pre-established rules, but by providing other rules, often incomprehensible to us, based on examples or trial and error.

According to some experts, it is possible to subject machine learning to the ethics of algorithms. Nick Bostrom, for example, argues that the controlling a superintelligence is no different than doing the same for a deaf and dumb child who 'divorces' their parents.

How to educate it? How to guide it towards what is good? How can we make it possible for it to act fairly?

2.2. The importance of algorithms in the construction of today's world

Day and night, we have to deal with algorithms. Their influence on our political opinions, our humour and our decisions has been clearly demonstrated.

Far from being neutral, they are chock-full of the value judgements of their programmers, and we often use them without even being aware of the fact. So, we need to question their ethics and find real solutions that are applicable to the biases that their users unconsciously follow.

What is Facebook's goal? And Twitter's? What is the purpose of a social media network? The simple (but nevertheless right) answer is to select the information presented on our feed so that we spend as much time as possible on the platform in question. Hiding behind the newsfeed is a curated selection of content (with or without advertising) optimised for each given user and bolstered by algorithms.

Thanks to them, social media sites establish what is most interesting to each of us. Without questioning the usefulness of such sites, their workings nevertheless raise great, thorny ethical questions.

According to Christine Balagué, researcher at the Institut Mines-Télécom Business School and former vice-president of France's Conseil national du numérique (National Digital Council), the issue of collecting personal data is

well-known, but less familiar is the processing of this data by algorithms. Even though users are becoming increasingly careful of what they share on social media, this does not mean – far from it, in fact – that they ask themselves how the services they use actually work. Only Facebook or Twitter themselves are privy to this knowledge.

Algorithms are everywhere in our lives, in our mobile apps and the web services we use each and every day. We have to take decisions from the early morning until late at night, but we do not do so in a vacuum: instead, we are confronted with all manner of suggestions, exposed to information processed by algorithms.

Netflix, Citymapper, Waze, Google, Uber, TripAdvisor, Airbnb… all take on a life of their own. A growing number of articles by researchers stress algorithms' power over the public, their preferences, likes and choices.

In 2015, Robert Epstein, a researcher at the American Institute for Behavioral Research and Technology, showed how a search engine could influence the results of an election. His study, based on a sample of more than four thousand people, allowed him to establish that the candidate's social class in search results influenced at least twenty per cent of undecided voters.

Research carried out in 2012 by Facebook on 700,000 users showed that those exposed to negative feed content made mainly negative posts, while those subject to positive feeds had basically positive posts. This shows that algorithms are capable of manipulating people's emotions without them being aware or informed of the fact.

This opacity is one of the main ethical problems of algorithms. Using a search engine like Google, two different users making the same search will not end up with the same result. The explanation offered by the service is that the responses are personalised to better serve each individual, but the mechanisms selecting the results remain unclear.

Of the parameters used to establish which pages will feature in the search result, more than one hundred are associated with the user making the search. Under the guise of trade secrecy, the exact nature of these personal parameters and how they are taken into account by Google's algorithms currently remain unknown. It is almost impossible to ascertain how the company categorises us, how it determines what we are interested in or how it predicts our behaviour.

Once we have been pigeonholed into such a category, can we escape it? How can we control the perception the algorithm has created of us? Given the aforementioned opacity, it is just as impossible to know any bias the processing of our data may be subject to.

Studies by Grazia Cecere, professor of Economics at the Institut Mines-Télécom Business School, have revealed the presence of gender discrimination in a leading social media network's recommendation algorithms. The myth of

the evil artificial intelligence notwithstanding, the source of this type of bias is to be found in human actions. All too often forgotten, we should be mindful of the presence of the programmers responsible for every line of their code.

Algorithms exist, first and foremost, to suggest services, most often commercial in nature. They therefore form part of a business strategy. They are created in response to economic demands.

Data scientists work on algorithm optimisations projects without necessarily considering the ethical issues such programs entail. We human beings have perceptions of society that we incorporate – more or less consciously – into the logical frameworks we develop. We project onto them what we are and what we think without realising the fact.

An algorithm's value judgement is often that of its creators, as the works of Grazia Cecere show. An algorithm learns what it is made to learn and robotically reproduces the stereotypes held by its creators.

One illustrative example of this phenomenon is to be found in diagnostic imaging. An algorithm that classifies a cell as ill or healthy needs to be configured to find a balance between the number of false positives and that of false negatives. Its programmers have to decide the extent to which it is acceptable to have positive tests in healthy people so as not fail to take account of sick ones that would have negative results.

Doctors would prioritise being made aware of false positives over false negatives. On the other hand, the scientists developing the algorithm would prefer to be aware of false negatives rather than false positives, because scientific knowledge is cumulative. Depending upon the values prioritised by the programmers, one or other profession will be given preference.

As already noted, the only way to combat opacity in algorithms is to make them more transparent. In France, since October 2016, Axelle Lemaire's Law for a Digital Republic has required transparency in all public algorithms. And, slowly but surely, companies are also complying with this law.

Since 17 May 2017, Twitter has allowed its users to see what interests it has associated with them. Despite the good intentions, this transparency is scarcely enough to guarantee an ethical dimension.

The intelligibility of coding is often ignored. Algorithms are produced in formats that make them hard to read and understand, even by professionals. Transparency may, therefore, end up being a fabrication.

For example, Twitter has not disclosed how it associates interests with users. How do its algorithms carry out the weighting between 'Science news' and 'Business and finance' to add contents to a user's feed?

What's more, there is a need to assess how transparent algorithms are. This is the mission of TransAlgo, a platform that gauges transparency taking into ac-

count the data employed and the data produced, and that is the first of its kind in all of Europe. Similarly, mention should be made of DataIA, a convergence AI institute bringing together the expertise of the Université Paris-Saclay for a period of ten years.

The latter is a unique interdisciplinary programme encompassing research on AI algorithms, their transparency and ethical challenges. It brings together multidisciplinary teams of scientists with the mission of studying the mechanisms for developing algorithms.

Human sciences can contribute a great deal to the analysis of the values and decisions hidden in coding. It is becoming increasingly necessary to deconstruct algorithmic methods, to carry out reverse engineering, to measure their potential biases and discriminations and to make them more transparent. There is a need for ethnographic research on their programmers, examining their intentions and studying the socio-technical fusion of algorithms.

As digital services become increasingly important in our lives, it is essential that we be able to define the risks algorithms represent to users. On the basis of study topics as diverse as the environment, health, robotics and nanotechnologies, we need to raise tech designers' awareness of ethical issues.

2.3. Against opacity in algorithms

The word 'algorithm' is becoming quite fashionable. And not just the noun itself but also its derivatives. Countless words are being penned on algorithmic culture, algorithmic love, the ethics of algorithms, the power of algorithms, the algorithmic society and even algorithmic governance.

This has been noted by S. Habitable and G. Donem when they say that algorithms are transforming science, industry and society, and are changing our perceptions of work, property, government, privacy and humanity.

On the one hand, algorithms make life easier but, on the other, they raise many serious doubts and questions. Algorithms transmit all manner of fears, but it is good that programmers themselves help us to remove them from society.

We are completely surrounded by algorithms. Some react to the financial markets, others work in the field of insurance or in the media or in law and order. Others still guide our choices as consumers or drive social media sites.

Algorithms are now part of society's life, and have become so because, to a degree, they have been increasingly trusted with essential operations. Decisions on how to advise a student, whether they will receive a grant or not, forecasting whether a prisoner is likely to reoffend, anticipating the outcome of a trial, all these are actions that algorithms carry out every single day, but this does not mean they do so in compliance with a set of ethical principles, such as respect for individual rights, equality and non-discrimination.

As things currently stand, algorithms raise a number of particularly thorny issues, as they come wrapped in an aura of scientific objectivity, as if any decision made on their advice were unquestionable because it is purely mechanical and lacking any form of prejudice. The presupposition is that they are neutral.

If that were is the case, it would be easy to conclude that it would be better to accept the verdict of an algorithm rather than a human judge, likely to take decisions that vary over the course of a day due to the effects of tiredness and mood.

The algorithm, with its mathematical functioning, is thus presented as a solution to mitigate human fallibility. Nevertheless, it is worth taking an ethical stand and question this alleged objectivity.

As Dominique Cardon says, algorithms are not neutral. They reinforce a view of society given to them by their programmers or those who pay the latter. These technical artefacts mirror the principles, interests and values of their designers. The operational implementation of these values depends on technical decisions, statistical variables and calculation methods.

Given that algorithms classify, operate, categorise and recommend, among other operations, they fall fully within the field of ethics. So, it is essential that we consider how they do so, how they reach their conclusions, but this is only possible if we pierce their opacity. In light of said capacity, we need to call for transparency and make clearer than ever the need to open up these black boxes.

The reason for this demand is obvious. In a democratic society, we need to publicly decide which criteria and which principles we want algorithms to work with and the extent to which we wish to delegate to them our decisions and, at the same time, what kind of control we want to exercise over them.

The examples provided over the course of this article show that we cannot base ourselves on the assumption that algorithms are neutral. In that they are given decision-making criteria, algorithms are by no means so. They are automated systems that can easily contain biases.

Discussion of the issue of bias has deep roots. As far back as 1996, Batya Friedman and Helen Nissenbaum noted the possible biases of a computer system. They stated that these biases could come in three forms: pre-existing, i.e. contained in the attitudes or practices and in the institutions predating the system and that could impact decision-making either explicitly and consciously or implicitly and unconsciously. Secondly, there are technical ones stemming from technical considerations and, thirdly, emergent ones arising within the context of a system's users.

Although ethical questions also arise around classic procedural algorithms, the issue of opacity in algorithms has once again become trending news with the

successes of *big data* and artificial intelligence, which lead us to focus on statistical learning algorithms.

Until recently, the main topic of discussion was supervised learning algorithms. As Yann Lecun reminds us, the basic principle of supervised learning is always the same: it consists in adjusting the system's parameters to minimise the cost function, which measures the average error between real and the desired system outcomes, calculated with regard to a set of learning examples. Reducing this cost function and training the system are actually one and the same thing.

Ethical considerations need to take into account two factors: firstly, the data and, secondly, the algorithm. Data can, obviously, contain biases due to their being unrepresentative or because they reflect a reality that is in itself discriminatory.

Let us not forget the controversy caused by Google Photos when it classified black people as gorillas. In this case, the problem was due to the fact that the training dataset was not representative of the actual overall population. The system had undoubtedly been trained based on clichés formulated, in the main, by white people.

To illustrate this with a second example, imagine an algorithm that has been trained on the basis of past decisions. Here, it is obvious that it will reproduce certain discriminatory biases, for example, with regard to women.

This is why attention is paid to a certain number of ethical principles with regard to the data processed, such as the absence of bias in the data, that the data respects the diversity of cultures or groups, that they do not entail any risk of discrimination, that the designers ask themselves whether the variables are capable of being socially discriminate.

While the issue of data is undoubtedly very important, it is also worth stressing that of the processing and ways of exploiting these data. Algorithms work on these data to make them talk.

The problem becomes particularly serious when it is difficult to say, exactly, which criteria have been taken into account when implementing a classification: and it is precisely this situation that arises with some artificial intelligence algorithms. This makes it hard to trace the route taken by the machine to arrive at a certain decision or to identify the criteria it has taken into account to reach it.

The ability to learn makes this explanation considerably more difficult and means that designers are not in a position to understand a system's behaviour. While classic algorithms provide a model that is explainable, because it is made by analysts, learning creates a model by adjusting its parameters, and works with millions of datapoints.

This issue leads naturally to other algorithm-related ethical challenges, such as our ability to understand them to be able to govern them. After all, we wish to govern them, not be governed by them (don't we?).

So, what is it we are asking of algorithms? Simple: we're asking them to be accountable, we demand of them transparency, trustworthiness and explanations for the decisions they take.

But what do we mean when we talk of transparency? When is an algorithm transparent? When its workings are clearly explained and the data it processes are accurate. Transparency in an algorithm allows us to check the decisions it takes and the choices it makes. Thus it is that the principle of transparency in algorithms is closely linked to that of trustworthiness and that of fairness.

While a principle of this nature is obviously desirable from an ethical standpoint, its real-life implementation can be problematic.

Algorithms can be subject to three kinds of opacity:

i. The first is the result of an intentional strategy of keeping trade secrets.
ii. The second arises from the code itself not being understandable to all.
iii. The third – and most problematic – stems from the conflict between the demand for optimisation of higher-dimensional mathematics and the semantic demands of human explanation-seeking reasons.

The first type of opacity can have considerable impact on any demand for transparency. Such transparency would require the divulgation of the algorithmic system's code, and this could clash with trade secrets and intellectual property rights. And, even if it were to be disclosed, the algorithm would still remain unintelligible to the majority of the general public.

These reasons would limit any attempts to implement a national algorithm audit platform, be this in the form of a public body of experts or by means of type-approved private audits.

The issue of whether we really wish to govern algorithms and not be governed by them is a pressing one, but can only be addressed if we have the power to understand the general logic of an algorithm's workings, the criteria governing its decision-making processes and if all this is explainable and interpretable. Members of the public are entitled to know the bases upon which they are processed and classified.

As noted above, these demands are problematic with regard to the new artificial intelligence algorithms. There are a growing number of technical initiatives designed to make them more understandable. This is good news, but means that the ethical debate must be supported by this intensification of technical research.

Fate has nothing to do with the workings of algorithms. Neither does random chance. Quite the opposite: they are programmed. They can be properly audited and we can demand that the criteria that govern them be ethically noble, such as those promoting justice and fairness.

The public must be enabled to understand computer systems. Or, put another way, there is a need to bolster autonomy and considered thought to mitigate any situations of asymmetry that may be created by algorithms. People need to be trained to make informed, lucid decisions.

The ethical requirements are therefore clear: bolster algorithms' democratic matrix and strengthen the freedom of users so that they do not give up algorithmic governance.

Aside from explaining algorithms, some academics are concerned with what has been dubbed 'algorithmic governance'. But what does this mean?

This expression is used to describe a kind of (*a*) normative or (*b*) political rationality based on the collection, aggregation and automatic analysis of massive amounts of data to model, anticipate and affect in advance possible behaviour.

The underlying idea is that people can be governed simply by gathering raw, supposedly neutral and objective data, and that, by algorithmically exploiting this, their behaviour can be accurately predicted without making hypotheses. The revolutionary nature of this concept cannot be stressed strongly enough, and nor can the changes that it may bring about in society.

The dizzying pace of development of *big data* and AI may also bring about comprehensive mutations in the scientific method. With enough data, figures speak for themselves. Hypotheses could be produced from data. It is thought that the inductive logic governing statistical learning methods is also heading in this direction.

As A. Rouvroy notes, the problem is that surprises are found along this road not to improve the world in which we live, but quite the opposite, to leave it just as we have found it, as noted above, with the possibility of inequalities being enshrined and reinforced, while at the same time neutralising any chance of criticism. The goal of objectivity is, in fact, to consign political choice to oblivion. This means that opening up these black boxes plays a crucial and critical role.

However, we should not forget that there is an enormous difference between creating a recommendation system to improve (for example) suggestions made to users by platforms like Netflix and modelling society as a whole, in all its complexity.

Good scientific practices and the ethics of scientific research must prevail. We cannot process data without first questioning the reality they represent.

We are entitled to question the implementation of models in which prediction is a priority over understanding or explaining the phenomena. We need to assess the extent to which such models may entail the destruction of our freedoms.

2.4. Prejudice embodied in algorithms

Some corporations are working on developing ethical solutions to the problems raised by the development of smart machines and their implementation in society, in an attempt to find the right mechanisms to prevent the appearance of AI systems that spawn or bolster racial or sexist prejudices.

Constructed in accordance with the machine learning model, artificial intelligence algorithms are able to optimise their calculations on an ongoing basis as they process data.

Through its operational criteria or its training data, an algorithm may – despite everything – be biased and reflect society's discriminations.

There are two types of algorithmic bias: those that reproduce the discriminations existing in society and those that manufacture them, because they are constructed upon training datasets that are not representative of society.

In 2015, researchers pointed out the sexist bias in AdSense, Google's automated advertising platform: women were being provided with advertisements for jobs that were worse paid than those aimed at similarly-qualified men.

In 2018, researchers noted that an Amazon recruitment platform discriminated against women, even though gender was not a key data point. Women, it seemed, used fewer relevant keywords in their applications than their male counterparts.

The algorithm's sexist bias faithfully reproduced a pre-existing societal bias that the engineers designing it failed to give enough critical thought to.

Ensuring that an algorithm does not discriminate against certain groups in society is one of the most important challenges facing programmers, who we make responsible for decoding inequalities to encode equality. To prevent the introduction of learning biases, though, data scientists need to ensure that datasets are sufficiently diverse. Codes need to be crosschecked and other engineers allowed to contribute different viewpoints.

The ethical challenges surrounding AI consist in avoiding reinforcing existing stereotypes and not creating new ones. According to Marie Crappe, CTO at StaffMe, a French on-demand assignment platform, this calls for permanent oversight. This platform is dedicated to eliminating all irrelevant dimensions from the recruitment process: gender, place of birth, nationality, etc., to avoid introducing sexist biases.

What's more, there is a need to diversify the profiles of those coming up with algorithms. As with politics, an unrepresentative decision-making body may have negative consequences for society.

In this regard, companies should be encouraged to permanently ask questions around the possible discriminatory impact of algorithms by implementing eth-

ics by design from the very conception stage itself, in preference to carrying out ex-post-facto impact assessments of discrimination.

All the links in the algorithmic chain, from companies creating devices to web surfers, need to be trained in ethics. Those who cannot understand the abilities of AI are denied an element of free choice, and it is information that provides the basis for this understanding. Additionally, companies should be advised to provide training seminars and awareness-raising programme for all the players in the digital world.

Implementing ethics provides an opening for bolstering trust in the digital economy. Some businesses have created a position – the head of digital ethics – to act as a key interlocutor, responsible for the coherence and consistency of the ethical and digital policies, beyond conformity with the law. This is a first step.

Linking equality with innovation is a golden opportunity that must not be wasted by large corporations. Alongside the implementation of legislation and the establishment of the General Data Protection Regulation, the introduction of ethics into these processes is essential for fostering public trust.

2.5. The reproduction of biases in algorithms

Artificial intelligence techniques have incredible potential. Numerous opportunities are opened up, but a whole range of ethical questions are also raised.

These questions encompass a number of areas, including the use of personal data, the ecological impact and the emergence of AI systems capable of taking decisions for themselves.

Over the course of the working day, we send and receive any number of emails. To travel, we use our favourite route-finding app. We use digital services every day of the week. Most of these services are free of charge, but there is another aspect to them that should not be ignored.

As we use them, the resulting data is compiled and analysed, particularly for advertising purposes. The profile of each and every one of us is analysed based on our use of these applications, to suggest the most appropriate adverts for us, taking into account our observed habits and behaviour.

There's a saying that sums this up perfectly: 'If something is free, you are the product'. This can be observed in the vast majority of tech industry giants and often lies at the heart of their financial model. So, how can we ensure that our data will not be used to unwanted ends? Above all, by using our critical senses.

If you use a digital service, it is worth spending time identifying what data it gathers so as to understand how they can be used.

When a device, app or web page asks you to identify images to prove that you are not a robot, you are already becoming involved in image recognition algorithms.

When we agree to continue to surf a website by consenting to cookies, we are allowing algorithms to suggest to us advertising that has been filtered to meet our preferences. Some practices may be acceptable to us, while others may not, but the point is to be aware of what is happening.

At a Europe-wide level, the General Data Protection Regulation (GDPR) came into force in 2018. This Regulation has the aim of strengthening the framework for those organisations, be they public or private, that process personal data.

Certain countries have some usages that may place this in doubt. This is the case, for example, of China where, like in Jeremy Bentham's panopticon, the state apparatus monitors its citizens' Internet usage, as well as their movements and behaviour, based on image recognition.

Most cities are under video surveillance. Citizens' data and AI are used to give each member of the public a label based on their actions, to permit them to access, more or less easily, services such as loans or transport.

Huge companies gather huge amounts of data, which are used for advertising purposes, but not only that. With the appearance of artificial intelligence, certain tasks can be automated. Thus, the associated decisions are no longer made by human beings, but rather by algorithms.

How great is the risk of bias in this? Very great.

Algorithms can reproduce human biases. They are never axiologically neutral, as they work on the basis of training datasets. This data may be biased, providing an imperfect view of the world, as could happen to a facial recognition algorithm that is only trained using white people. If the data contain traces of discrimination, the algorithm will, in turn, display discriminatory behaviour.

Let's look at a specific example: recruitment. HR managers are nowadays making increasing use of AI solutions in this field. Logical systems automatically analyse documents such as CVs to save the profiles of the candidates best suited for specific roles or duties.

For example, in 2015, Amazon implemented a recruitment system to analyse applicants for its job postings. It had deliberately created an algorithm that was biased to identify women candidates.

Nevertheless, the system rejected, by itself, women candidates in favour of men. It did so because it had been trained on the basis of data from Amazon's employee organisation chart, 85% of which was made up of men. Once this fact was revealed, Amazon decided to abandon the system.

These problems are not systemic and there are a great number of extremely useful services created with AI. In the case of recruitment, the services offered are often very relevant to job hunters. For example, if someone is looking at job

postings, artificial intelligence services are very good at selecting those that best fit the particular seeker's profile.

However, we need to ensure that all these products are useful to the entire world and do not reproduce sexist or racist behaviour. So, we need to retain a critical mindset with regard to these AI solutions and demand transparency in such systems.

And there are still more problems we need to consider. We've all seen those surrealist videos in which Barack Obama insults Donald Trump and Mark Zuckerberg explains how he manipulates Facebook users.

These began to appear in 2018 and we know they are hoaxes, but they are astonishingly realistic. The technology making them possible is based on recent AI techniques. Such hoaxes are often described as *deep fakes*, but are not always so quickly discovered, nor is there always awareness of the fact.

The technology used is that of *deep learning*. Some might object by saying photos were manipulated before the existence of artificial intelligence. However, while this is true, such hoaxes are increasingly accessible to the public and the results are more and more convincing than ever before.

If you take a look at the website *This person does not exist* (https://thispersondoesnotexist.com/) you can see how important this issue is. It contains a series of generic portraits using this AI technique, and it is easy to believe that they are of real faces, even though they have been created artificially. Deep fake technology is also affecting, in turn, the manipulation of photos, audio and video.

In 2019, a Chinese app, Zao, caused uproar by allowing its users to replace the face of an actor in a music video or film with another of their choice.

There's no doubt that these new techniques create risks, particularly that of mass disinformation. What can users do? Are they at the mercy of all this? Can they any longer believe in anything that is presented as true? Should they be sceptical of everything?

To deal with all of this, it is essential to ask for the source of the information, if it is legitimate, and whether the media platform it is on is trustworthy. One should also compare and contrast information with other websites and use fact-checking pages like those suggested by some newspapers.

It has been said that data are the new oil of the 21st century and that AI drives this new black gold. What impact does this whole industry have from an ecological perspective? This is no idle question, particularly against the backdrop of a global environmental crisis and climate change.

As an economic sector, the digital world has a significant ecological impact. It was estimated to represent around 3.5% of worldwide CO_2 emissions in 2015. To put this into some perspective, the figures for air transport were 2% and, for telecommunications networks, about 0.5%.

To be able to work, the digital economy uses not only all our digital devices (smartphones, computers, tablets, etc.), but also a whole range of less visible infrastructure: fixed and mobile networks, business networks and data centres. Such centres host the servers that make the Internet and its services possible.

We often hear the term *cloud computing* to refer to these remote servers. The cloud represents all these devices that work remotely at all times to make out digital activities possible.

As soon as we start using the Internet or one of its services, we are asking a server somewhere in the world to process our request and provide us with a response. To do this, the server consumes energy.

How can we limit our ecological impact on a worldwide scale?

We can do something each and every day to reduce our digital ecological footprint. Simple things, like deleting old emails, cancelling subscriptions to unwanted newsletters or limiting the use of transmission platforms.

The development of artificial intelligence increases this environmental impact. It requires huge amounts of storage and computational capacity. The reason that algorithms work so well is that they have been trained for days using extremely powerful servers that consume energy.

It should also be pointed out that efforts are being made to ensure data centres become increasingly environmentally responsible. Their energy consumption can be optimised thanks to AI algorithms. The heat they give off is sometimes used to heat nearby buildings or swimming pools (which use a lot of energy).

Also, even though AI algorithms use a lot of energy, they are included in a wide range of applications that improve production chains, as well as our consumption of resources and energy.

2.6. The ethical assessment of algorithms

The word *algorithm* has become the subject of much public debate, but what it actually means has not been properly defined. Algorithms are popping up all over the place, in the new information and communication technologies, in cloud computing, in connected objects, in AI, in blockchains, and elsewhere.

They have become the subject of concerned questions around the manipulation and exploitation of digital data, around confidentiality and transparency and around the individual and collective good.

In the field of healthcare, the new sciences leveraging mathematical tools to process the huge amounts of big data are crucial. We hope for a better understanding of things like diseases and optimised decision making and improved forecasting.

Let's take a look at some examples of the use of algorithms.

The use of big data can help identify a health risk and therefore prevent it. For example, American actress Angelina Jolie found out, after sequencing her genome, that she was the carrier of a mutation in the BRCA1 gene that entailed a more than 90% chance of developing a cancer in the next ten years. She therefore decided on undergoing a procedure to prevent this risk.

In cancer treatments, expert systems can process in record time around ten billion data points of a DNA sequence from a patient's tumour, something of which human doctors, no matter how qualified, are simply incapable.

Google Flu Trends is a tool that looks at Google searches and measures how often the terms *flu*, *gastroenteritis* or *chickenpox* were searched for around the world.

The idea was that people tend to search for information on the flu or other illnesses when they think they have symptoms of it and are, in the vast majority of cases, actually sick. So, when Google announced an increase in searches about a flu epidemic, this would generally occur two weeks later.

Algorithms can also help us to prevent the spread of epidemics. For example, in Africa, geolocation data from mobile phones are extremely valuable in monitoring the population movements through their actual travels and thereby anticipate the development of diseases in a country. This helped, for example, in predicting the spread of the Ebola virus.

Big data affecting a large cohort of patients has the potential to speed up scientific research and experimental protocols for a wide range of pathologies and treatments. For example, the Facebook social media site has been used by researchers to create a map of Americans at risk of obesity. Research of this type without such tools would take up a huge amount of time and cost a fortune.

This new world of algorithms and big data is permanently studying the real world in the aim of creating predictability. However, this raises a serious question: how can we ensure that an algorithm is ethical?

This calls for debate on how best to ethically assess algorithms. To goal is endow all these tools with sense, transparency, security and trustworthiness, so as to be able to better design, use and control them.

Ethical action is, first and foremost, a response to a critical and complex situation. Generally, digital ethics come in the form of questions on individuals' behaviour regarding and use of new information and communication technologies and, subsequently, on the (increasingly autonomous) behaviour of the technological tools themselves.

Within this framework, ethics, as a way of governing behaviour based on respect for values, are essential for providing a framework for using algorithms. We should not forget the time-based and often irreversible nature of certain

decisions, which is why ethical considerations are so important.

Human responsibility lies at the heart of ethics. Companies, too, must know how to incorporate ethics in their digital actions.

Ethical considerations must form an integral part of the latter's mission and so be able to enable the building of ethical discussions. There is a need to go beyond a merely interdisciplinary approach to achieve a truly cross-disciplinary one, a merge of disciplines to achieve real digital ethics, in which social and moral questions are integrated within the new information and communication technologies.

It is essential to carefully craft specific ethical rules in the digital world without losing sight of the question 'can this digital world put our ethical behaviour at risk?'.

We need to assess the digital world as a continuous system without segmenting the different parts that make it up.

Studies of algorithmic processes are implemented on the basis of three interdependent categories:

a) To begin with, there are the ethics of data, which include the identification, construction, nature and characteristics of the data processed by an algorithm and the exchanges of these data.
b) Next, there are the ethics of algorithms, which include the workings, operations and processes associated with an algorithm over the course of the data's lifecycle.
c) Lastly, there are the ethics of practices, which includes explanations of the quality of the purposes and results of an algorithm.

There is thus a need to establish a framework for discussion and for good ethical practices around the creation, implementation and use of algorithmic systems, artificial intelligence and big data projects, so as to be able to better design, control and monitor them in companies.

Strict, sanction-based governance of the digital world is not, in our view, what is initially called for. Regulatory aspects should not be our first line of defence, but neither should they be forgotten. What is essential, though, is to examine the ethical dimension of algorithmic processing on an ex-ante basis.

Governance of the new information and communication technologies must make a decisive break with a rigid, purely technological and normative approach to embrace the cross-cutting, flexible, dynamic and evolving nature of the ethics of algorithms.

These ethics, based on the principle of 'ethics by design', are what need to be implemented in the coming years and must form the cornerstone of the rela-

tionship of trust that needs to be built with the public. This will help both users and the public authorities gain better awareness around the use of big data.

All this will help encourage members of public to ask themselves the extent to which they can hand over their private lives to digital services.

2.7. The responsibility of algorithms

Artificial intelligence and algorithms have suddenly appeared in our day-to-day existence and are spreading everywhere. Complex tasks are increasingly being delegated to increasingly sophisticated and autonomous applications, as machine learning and deep learning methods are being developed.

These exponentially increasing usages, associated with enormous amounts of data, available across almost all sectors, raise numerous ethical, legal and social questions.

We are witnessing AI-based applications being developed at a dizzying pace. They are being used in every sector: healthcare, sport, human resources, baking, insurance and defence. Such applications are frequently found integrated into robots.

The first question that pops up is that of responsibility. When an accident that could be caused by a self-driving car occurs, many questions are raised.

Who is responsible when a self-driving car runs over a pedestrian or crashes into a wall killing its passengers? The vehicle's owner? The program's designer? The program itself? Can a program be held civilly or criminally liable?

To try to answer this battery of questions, we need to take a deep dive into the world of consciousness. We accept that human beings have consciousness, that they can be held accountable for their acts, because they are capable of carrying out said acts freely, intentionally thought out and performed as a result of their will.

Can machines or algorithms acquire consciousness/awareness? Can they be held accountable? Must they answer for their acts and decisions? Are these really *theirs* or have then been programmed by a human being?

Before tackling an issue involving both philosophical anthropology and the fundamentals of ethics, let us consider another example. In the United States, automatic applicant selection applications based on the analysis of CVs have been described as discriminatory. The machine learning system was fed with the CVs of white males. Machines learn from what human beings feed them.

Artificial intelligence is a set of concepts and technologies that, when combined in the right way, gives rise to a system capable of simulating, if only partially, human intelligence in the fields of logical reasoning or learning. It has numerous applications in both assisting with diagnosis and ongoing care. It enhances pro-

cesses in many fields such as medicine, finance, human resources and customer relationship management.

To achieve these goals, AI employs logical or algorithmic problem-solving methods. The development of AI has been closely linked with the extraordinary growth in computers' calculation abilities (following Moore's famous law, which has yet to be refuted) and ever greater volumes of data, which we call *big data.*

The 1980s saw the birth of machine learning. During the 1990s, the development of expert systems began, mainly for finance-, health- or maintenance-related applications. These systems offered assistance with diagnostics thanks to the expertise of a man in the sector who allowed his experience in a particular field to be formalised.

Expert systems were made up of a knowledge base, rules and an inference engine. The goal, then, was to be able to model, formalise and conserve the expertise of the leading professionals in their particular fields. This was done so as not to lose this knowledge and by employing these applications to train future experts.

Unfortunately, expert systems did not have the impact they perhaps deserved, essentially for two reasons. At that time, the costs involved in the calculations and the volumes of data were prohibitive, but, above all, it was due to the fact that many experts refused to contribute their knowledge and allow it to be modelled to be transferred.

Nevertheless, the way forward had been shown. A few years later, in May 1997, the computer Deep Blue beat Garry Kasparov at chess.

From the early 2000s, the incredible growth in the volume of data available and the development of new computational powers and infrastructures meant that some computers were able to exploit big data to an unprecedented degree. And so it was that a new expression was coined: *deep learning*.

Deep learning means the exploitation of vast amounts of data (structured or otherwise), which allows a machine to create new rules based on the data made available to it. These days, a computer equipped with applications of this level is more capable of resolution, in some fields, than any human being, no matter how talented or qualified.

In May 2017, a computer beat the world's leading Go player. South Korea's Lee Se-dol, a Go world champion, was beaten by Google's AI AlphaGo, beginning in the first of match of five and ultimately losing 1-4. A few years later, Lee Se-dol decided to retire from playing competitive Go.

Another example of this superiority of the machine in some areas of human life is to be found in an article from *Usine Digitale* of 21 August 2020. In a competition organised by DARPA (the Defense Advanced Research Projects Agency) in the US, a number of companies pitted their AI self-piloting agents

against each other to establish which was best. The final duel was between the winner of this contest and a US Airforce pilot, who proved incapable of beating the AI agent.

To conclude these examples, which are becoming commonplace everywhere, we would refer to an article published in the *Journal du Net* (JDN) on 28 August 2020 that shows how new professions like that of the data scientist may be replaced by machines. The question asked by the articles is loud and clear: Will automated machines learning replace the data scientist?

The AI sector is currently seeking to create an AI capable of perceiving its surrounding environment, understanding the situation and, above all, take decisions.

After this digression, let's return to the case of the supposed responsibility of a self-driving vehicle. There is as yet scant jurisprudence on the matter. Responsibility in the case of an accident may be assigned to the user of the AI, or its owner, or its designer or a possible employee of theirs. Everything depends upon the contract. The legal debate is far from over and the issues are extremely complex.

Some speak of equipping AI with consciousness, while being cognisant of the fact that, again in this case, the awareness and degree of freedom associated with it will be programmed by human beings. As of today, this level of consciousness is difficult to program, because we do not code what we cannot control. We human beings ourselves are incapable of controlling our consciousness, our subconscious and our unconscious nature. And what does it mean, in the end, to say that 'I am aware that I am a man'? And how to we program this?

As things currently stand, the concept of artificial intelligence refers to a machine capable of certain forms of intelligence arising in the human brain, but this does not include being conscious of oneself or self-aware and nor is it an emotional life. Some researchers believe that, in the not-too-distant future, they will be able to create a conscious intelligence in a tangible or completely intangible medium. If such a day actually dawns, there will be no limit to the conception of all forms of artificial intelligence. The only limits will be those of humankind's ability to model ideas and abstract concepts and to develop the associated logical frameworks.

Insofar as consciousness can be modelled in the form of facts and rules, it is possible to conceive of it and implement it within a system. There is nothing, however, that could lead us to think that robots could become independently aware of their own existence.

To achieve a certain degree of awareness, an AI system's architecture must include in its core, necessarily but not exclusively, a basis of universal rules, the baseline rules accepted by humankind as a whole. Alongside these universal ethics, an AI must also possess, if at all possible, benchmarks reinforcing the mandates associated with a religion, culture or ideology.

A benchmark for each particular profession should also be included, such as the Hippocratic oath in the case of doctors. These benchmarks could vary and give rise to different outcomes depending upon the country, culture or ethnicity.

Three or four hundred years ago, some Native American tribes dealt with their elderly in a way that would be simply unacceptable today: when an old woman had lost all her teeth and could no longer work, she was given some bread and water and left to die on a mountainside. Even today, some countries do not allow women to drive. To us, both practices may seem discriminatory, but they do form part of their respective cultural frameworks.

There is no need to go so far from home, though. Take the example of a large London hospital: there, they have only one resuscitation bed for the cardiology service. If A&E calls them to bring a heart attack victim, the duty doctor asks for the victim's age, as they prefer to keep the place for a younger person rather than occupying it with someone of eighty-five or older.

If the goal is to create an artificial intelligence that works like the human brain and takes the same kind of decisions, it must be fed with the same rules, preferences and prejudices. Alongside a universal benchmark, it must be given other points of reference associated with religions, cultures and traditions grounded on the baselines provided by the rules of these universes.

An AI user could choose their usage benchmark in accordance with their desires and their own value system. The only thing else they would have to do would be to take out insurance with a company that advises on the consequences of their decisions, which may not always be in accordance with local law.

The great underlying question is whether we want to explicitly lay out all these possible rules to be made universally known. Designers could also grant an artificial intelligence free will, allowing it to make personal choices within its benchmark axiological frame of reference. Said free will could work, for example, on the basis of random chance or any other principle, but necessarily using algorithms programmed by human beings, albeit granting the machine a certain degree of freedom.

The clash between predestination and free will has, for a long time, fuelled debate in this field. Within the context of artificial intelligence, both can be programmed, with a greater or lesser degree of freedom assigned to the machine, while remaining perfectly aware of the fact that this decision-making freedom has been programmed by a human being.

This issue invites us the consider our own condition. It makes sense for us to asks ourselves, as human beings, if we are not, at the end of the day, an intelligence equipped with a consciousness and a sprinkling of free will, programmed by a higher power that some may call God, others evolution and others still random chance or pure necessity.

3. Artificial intelligence and the risk of gender bias

Can algorithms be called 'sexist'?

Tackling the risk of gender bias in AI is a complex issue, in two ways.

Firstly, as we have seen in earlier pages, there is the very notion of *artificial intelligence* itself and the ethical risks and challenges involved in its use in whichever field of human activity it is applied to. One of the key issues to bear in mind is that, despite the poor understandability and opacity afflicting these software resources, people tend to enthusiastically and completely uncritically embrace them (and, indeed, technology in general).

Added to the lack of intelligibility of the product/object of study, we need to add a second issue: that of conceptualising what gender bias actually is, as well as being able to appreciate it in practice and grasp its potential discriminatory risk.

Despite the fact that many of the texts we have reviewed mention gender studies as important for understanding the topic, we have found none that carries out any in-depth review of them as applied to the field of AI. Neither have we found texts that fully integrate the three fields of the ethics of algorithms, gender bias and recruitment and selection processes. Most of them focus on either gender bias or staff recruitment, but not on the intersection between the two. There is therefore an important gap that needs to be noted, thereby making our study ground-breaking, as it provides a critical and inclusive reading of these contributions, building bridges between the different fields.

Set out immediately below are some of the theoretical constructions of feminist theory, queer theory and gender studies, making potential connections with the ethics of algorithms and the world of work.

3.1. 'One is not born, but rather becomes a woman'

One of the first historical contributions to feminist theory was the creation of the constructivist perspective and, with it, the concept of the *social construct* when analysing the position and role of women in the social structure. Simone de Beauvoir, in her famous 1949 work *The Second Sex*, considered the historical and location-related implications of what it means to be a woman. 'One is not born, but rather becomes a woman', as her prescient statement put it (Beauvoir, 1953). These days, this expression is interpreted as reflecting a non-biodeterminist standpoint: in other words, being a woman, man, or any other kind of sex/gender option that may be offered by a given culture or social environment is, despite biology making its initial contribution, to a great extent learned. This learning is a result of an ongoing socialisation and enculturation, which begins at birth (within the family) and continues throughout one's life (within a wider social circle: school, work, society at large, etc.).

We can use this initial contribution by Beauvoir to extrapolate a premise for our analysis. When we think of AI algorithms, we do so by applying a constructivist perspective based on the premise that they necessarily incorporate biases, the result of both their construction and their socialisation. So, we need to ask ourselves when and how said biases are actually acquired, so as to be able to understand them.

There is more than enough literature on AI biases in general and gender biases in particular. A good part of it points us towards how we can detect them, properly characterises them and warns us that biases are to be expected and found throughout an AI's entire lifecycle (Leavy, 2018; West, 2020; Nadeem, Abedin & Marjanovic, 2020; Sun et al., 2019).

Most of these biases tend to be generically labelled *implicit biases* and described as unconscious or involuntary, as opposed to explicit and direct discrimination against a given group (De-Arteaga et al., 2019; Kleinberg & Raghavan, 2018). Some authors ascribe their appearance (at least in part) to the lack of diversity in the teams creating and producing AI-based products. According to this theory, an AI-based product may be unconsciously and involuntarily discriminatory against women if conceived of and designed by a team made up exclusively or mainly by men. Similarly, it may also be discriminatory against some ethnic and/or cultural groups if said team is exclusively or predominantly made up of 'Caucasian' people (Yarger, Payton & Neupane, 2020). Other texts make the same argument, while also expanding it to encompass the entire organisational environment and institutional culture. So it is that some works lay the blame on *geek culture*, defining this as a set of shared values and meanings found in tech company environments – and, more specifically, in computer programmers – where women rarely have a place or, at least, do not have one on an equal footing (Tassabehji et al., 2021).

3.2. The case of 'platform workers'

Along similar lines, the platform or gig work culture, a very important version of the 21st century's new (and often highly insecure) forms of self-employment, also takes up a good deal of space in the literature on biases and the ethics of algorithms. Without focusing specifically on gender discrimination, these texts highlight the great opacity of the associated algorithms and their enormous potential for discrimination, laying bare the lack of protection afforded workers when it comes to control and direct their professional careers or even just keep their jobs (Vyas, 2021; Jahanbakhsh et al., 2020; Kullmann, 2018; Rahman, 2021).

The argument that algorithms reflect the makeup of their design teams and the broader institutional culture, while widely supported, is also rejected by some authors, who state that being a woman does not automatically make one a gender expert (just as being an immigrant does not automatically make one an expert on ethnic minorities). This position should not be seen as denying the existence of possible biases, but rather the causes that contribute to them. It

does, though, impact upon proposals for mitigating them: while some authors recommend creating more diverse AI design teams (by implementing quotas such as the 'four-fifths rule' or the 'Rooney rule' in North America), others suggest hiring gender experts to raise awareness among organisations and work teams. In any case, these positions are not mutually exclusive and the debate is ongoing (Köchling & Wehner, 2020b).

Returning to the contributions made by Beauvoir, her constructivist focus provides us with a second important premise. It's not enough to say: 'being a woman is a historical, localised and situated construct'. If we look a little closer at her arguments, we can take away some political consequences. 'Being a woman' entails occupying a particular position and predetermined roles in the social structure and in the power and dominance relations of a given society. In the West, women have historically held subaltern positions and roles. Infants are socialised from birth in the difference between the sexes/genders. As part of this process, people with female biological attributes learn to become an invisibilised, dominated, objectified violated social group: in short, the 'second sex'. Only with an ongoing and high degree of coercion and violence (both physical and symbolic) by the dominant group (people with male biological attributes) has it been and will it continue to be possible for the subaltern group to carry out reproductive and less socially-acknowledged productive tasks, with a veneer of normality and even voluntariness. The feminist-constructive perspective teaches us to denature these assumptions and call them into question.

From this standpoint, in our analysis it can be seen that biases in algorithms not only incorporate the social gender biases of their creators and their surrounding social environment (as well as others, such as prejudices and stereotypes based on race, age, religious belief or different cultural characteristics), but, due to them – and this is extremely important – they reproduce power and dominance relations and different forms (some subtle) of exclusion (Crawford & Paglen, 2021).

3.3. Siri and Alexa: the case of personal voice assistants

One paradigmatic and sufficiently clear case of the reproduction of social stereotypes – and, with them, the dominance relationship they incorporate – is the design of personal voice assistants such as Siri and Alexa (to mention just two of the most common worldwide), which have female characteristics. Is this mere coincidence, the result of random chance? It would seem not, if we bear in mind that Siri and Alexa are both AI-based devices carrying out personal assistance duties that have traditionally held a subaltern – and thus pre-eminently female – position in the world of work (Adams & Loideáin, 2019). In this case, the authors go a step still further in their analysis, indicating how these devices with female names and voices help naturalise the idea the women exist to be 'used' by men.

3.4. A brief aside on intersexuality and non-binary gender

Having reached this point, some clarification is called for. In social theory, it is usually regarded as proper, following the constructivist line we have been explaining, to use the term *sex* to refer to biological features and *gender* for the social and cultural construction of stereotypical images and the roles and behaviours associated with each of the sexes. Both categories tend to be represented on the basis of a binary model, although, while for the *sex* category this binary nature is unquestionable, *gender* is often regarded as more flexible, with some contexts even admitting non-binary models. An examination of intellectual usages does not fall within scope of this brief review of gender theory[2]. Our review of the literature on AI and gender bias, we have not found any reference to discrimination against non-binary gender(s).

We therefore feel it is important to review the concept of gender by looking at recent contributions from queer theory and LGTBI+ studies. Historically, people with a non-conforming or non-binary gender have been the 'forgotten forgotten', even more invisibilised than women, relegated to the periphery or even completely and permanently excluded from society (Preciado, 2016). In this study, we do not wish to focus exclusively on practices that discriminate against women alone; we also want to look at anyone self-identifying with a non-conforming or non-binary sex or gender, as well as those who may be discriminated against by reason of sexuality. And, needless to say, men may also be discriminated against by reason of sex/gender in some job profiles. For example, duties traditionally considered as entailing care work (such as most healthcare or education-related professions) are regarded as female, as are housework, clerical duties, and those jobs with a requirement for an 'attractive appearance'.

3.5. The performativity of gender and symbolic violence

Continuing with our brief review of gender theory, Judith Butler, following in Beauvoir's footsteps, went further in the analysis of how this embodiment, this 'becoming a woman' occurs, highlighting the important role played in this process by language and, more specifically, its performative quality or power (Butler, 2007). In short, this consists in language's ability to create what it is naming. Social scientists often refer to this as the 'self-fulfilling prophecy' or 'Pygmalion effect', and it consists in people tending to learn and fit within the social roles and expectations of the terms (or social labels) attached to them – socioeconomic (rich, poor, middle class, homeless, etc.), medical (sick, crazy, bipolar, cancerous, etc.), educational (dropout, gifted, inattentive, hyperactive,

2 For an initial overview of this topic, see the article by Anne Fausto-Sterling, "The Five Sexes: Why Male and Female are not Enough" (Fausto-Sterling 1993), a benchmark for the non-binary and not exclusively biological vision of the category "sex". Also A. Bolin, "La transversalidad de género" (Bolin 2003) for a cross-cultural overview of the wealth of non-binary gender options in different social contexts and points in history.

etc.)[3], and so on. Thus, a woman becomes one when, from her very birth, she is called one (with her given name and the sex/gender assigned to her), to be followed over the course of her life with a succession of practices and representations, as well as rites of passage, of 'femaleness'. So, according to Butler, we should be aware of the language's creative and reproductive abilities, as well as its subversive potential.

Transposing the idea to the world of algorithms and their biases, we see that many studies focus on precisely this area, revealing the gender biases arising in translating or converting human language (natural language, text or audio), as well as images (such as photos or videos of professional interviews) into computer instructions (by means of labelling and coding processes) at different points of an AI's lifecycle. Thus, a great deal of the analyses in the literature reviewed focuses on the technical fields of *word embedding* and *natural language processing*, in the belief that the processes and ways of classifying, labelling and coding data (i.e. cramming the world's diversity into a certain number of pre-established categories) give rise to one of the most obvious forms of incorporating biases into algorithms (Leavy et al., 2020; Sun et al., 2019). Once incorporated into an AI, said biases are reproduced and even amplified during use. If not monitored and corrected, these could cause considerable real-life harm to certain people.

Different works point out different aspects of interest. Some of these are covered below.

If we examine the different points of an AI's lifecycle, we can see how biases can be included in one (or some or all) of the following processes:

Firstly, at the time of conception and implementation, incorporating the biases of the creators and their immediately surrounding environment (the work team, institutional culture and so on) in the very design of the project or its internal logic. These might include anything from the simple example of the characterising a virtual assistant as female to the case of an AI acting in support of a job postings website that, due to the way it has been constructed, ensure that some postings (such as those for jobs in tech companies or STEM careers) have less impact upon women: in other words, meaning women have fewer chances of actually seeing the relevant advert and applying for the job (Lambrecht & Tucker, 2019; Böhm et al., 2020).

Secondly, at the algorithm training/learning stage (with machine or deep learning) using a specific database for each given case (training datasets).

This training can be supervised or not, and the repercussions are different in each case. As might be imagined, unsupervised training is riskier in terms of the incorporation of uncontrolled biases (Crawford & Paglen, 2021; Nadeem, Abedin & Marjanovic, 2020).

3 We would highlight the socially pejorative tone of many of these descriptors, which we in no circumstances support.

Certain algorithm training datasets are pre-established: some are available for sale and others come in the form of open-source code. Algorithms can also be left to learn directly from the web, limiting them (or not) to certain sites (like Wikipedia).

Whichever the procedure, the data fed to the algorithms must first have been labelled by a person (a software company worker or someone else) to categorise what is relevant to the specific task assigned to the AI. This labelling will be affected by the subjectivity of the person carrying it out and act as a vector for the transmission of values and social stereotypes, of the worldview of this person and of the culture and society to which they belong. The bias stemming from the underrepresentation of groups in datasets (such as women, people of African descent, persons with disabilities and those belonging to specific age range) is known as representation bias (Gutiérrez, 2021; Leavy, 2018).

Furthermore, if we bear in mind that datasets are often historical, i.e. contain data from the past, we need to add another bias that some authors have dubbed historical bias. This means that an algorithm may be learning not only from the present, but also from the past. This was the case of Amazon's 'discriminatory AI', which 'learned' from a dataset of the company's own employees containing information covering the previous ten years.

3.6. Amazon's 'sexist algorithm'

In 2018, logistics giant Amazon hit the headlines due to the ethics of the algorithms it used in its human resources function. The company was accused of using a 'sexist algorithm' that discriminated against women and favoured male job applicants. Investigations revealed that the problem lay in the fact that the algorithm's training was unsupervised and employed a historical dataset of the company's workers covering the previous 10 years. This gave rise to a gender bias in its training, given that the majority of its workers over said period were men[4].

Therefore, with regard to those points at which an algorithm may acquire biases, we need to add the times of their specific uses, when these meet the right parameters for the performance of a specific production task and are fed with real-world data. At this point, we are no longer in the technical and expert environment of tech industry workers, but rather in organisational or production contexts with a wide range of players, both public and private (personnel departments, institutions, organisations, companies) that may purchase AI-based software for use in a particular task or process. One great example of this would be the initial screening of job applicant CVs to determine who moves on to the interview stage. In real-world settings, algorithms need to be adjusted, monitored and assessed in order to evaluate not only their performance in

4 Dustin, J. (2018). Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, 11-10-2018. Recovered from: https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G [last consulted on 10-6-2021]

terms of cost effectiveness but also their accuracy (lack of bias) using the real parameters of specific tasks and the characteristics of real-life data. We need to judge whether the results are not only correct (accurate) but also fair (Fabris et al., 2020; Fernández-Martínez & Fernández, 2020; Hangartner, Kopp & Siegenthaler, 2021; Chen et al., 2018).

Returning to our theorical approach, one concept that complements and helps us understand the notion of language's 'performativity' is that of 'symbolic violence'. This concept is particularly clarificatory when gauging the how deep-rooted the mechanisms are. As we have been suggesting, the dominance of one gender over another is not a natural phenomenon but rather a naturalised social one. The subaltern gender is not born in this state, but is subjugated, insidiously and violently. According to the thesis of Pierre Bourdieu in *La domination masculine* (1998), violence against women is not only physical but also and above all symbolic, discursive. It is by means of discourse, of the symbolic universe between dominator and dominated, oppressor and oppressed, which they share each and every day – the same worldview, values, ideals and lifestyles – that the subaltern classes subtly accept their fate. Thus, it is that worst form of domination is that imposed by women against themselves, through their own choices and self-censorship. So, let's link this with the world of AI: as shown by some studies, it is often women themselves who do not want or give up on jobs in very male or masculine environments, such as tech companies (Tassabehji et al., 2021) and STEM careers (Lambrecht & Tucker, 2019; Böhm et al., 2020). What's more, they point out how rarely women secure positions of leadership in such environments, and their need for guidance and support. As we have seen above, and to come full circle, work environments lacking diversity are more likely to give rise to discrimination, either by means of algorithms or directly.

3.7. The intersection of multiple forms of discrimination

We have seen how sex/gender becomes a social marker that places women on a lower rung on the social ladder, a practice that AI incorporates and reproduces, if not magnifies. In the social division of work, gender has been – and still remains – a leading discriminatory factor, but by no means the only one. Given this, any proper analysis of gender cannot ignore the latest contributions to the field, such as the theory of intersectionality. It was Kimberlé Crenshaw who first promoted this idea, based on a series of cases of workplace discrimination suffered by Afro-American women and who she herself, as a lawyer, represented before the courts and lost. The real problems experienced by her clients was that they suffered from different forms of discrimination at the same time, and that these overlapped, magnifying the extent of the harm caused. At the time, the American justice system did not permit the defence of a case on the grounds of both racial and gender discrimination at the same time: you had to choose one or the other. This was why the lawyer repeatedly failed to win her cases. The analogy used by Crenshaw, and which found favour in the social sciences, is still highly illuminating: 'if an accident happens in an intersection, it can be caused by cars traveling from any number of directions and, sometimes, from all of them' (Crenshaw, 1989).

So it is that we need to bear in mind that many forms of discrimination and hence biases in algorithms – age, gender, ethnicity, educational level, disability, religion, sexuality, etc. – can overlap or intersect, thereby giving rise to still greater injustices. According to this theory, which is supported by the studies in the literature reviewed (many of them experimental in nature, performing tests with simulations), a black woman will have a far greater chance of suffering from double discrimination due to the racism and sexism implicit in algorithms, when compared with a white man. A clear case exemplifying this theory in the world of algorithms is to be found in the inefficacy, despite the great efforts made in the field, of corrective mechanisms (i.e. antibias measures) created by the AI industry itself to counter its acknowledged problem (Bornstein, 2017; Raghavan et al., 2020; Vasconcelos, Cardonha & Gonçalves, 2018). The main reason antibias measures fail to work is that it is very difficult to 'wipe' gender markers from data. It is not enough to exclude a name from a CV, or remove pronouns from a text, or even to refrain from detailing the gender of given content (such as photos). Gender leaves a mark on women's lives, biographies and curricula that is easily detectable by an algorithm accustomed to processing vast amounts of data and finding correlation patterns in them. Some of the sources consulted speak, with regard to this issue, of a gender cluster (De-Arteaga et al., 2019). Let's examine one example of this provided by these authors within the context of personnel recruitment and selection. Despite the elimination of the most obvious gender markers from CVs, women's career histories often include temporary gaps or periods of lower productivity (part-time or lower-qualified jobs) in their professional experience section, easily attributable to maternity and childcare. An AI designed to rank job applicants scoring this section without any compensatory (debiasing) measure would be both objective but also unfair at the same time. The many discriminatory factors inexorably intersect and correlate.

3.8. The transformative dimension

Last but not least, we could not conclude this section without noting how the contributions of the authors we have discussed can provide a significant transformative dimension: what is socially constructed can be modified or deconstructed. This is therefore also the case of gender biases in algorithms, however deeply hidden and embedded they may be.

This – as we shall see in Part II – can be carried out at in least three stages in an algorithm's lifestyle, and perhaps more:

i. Taking steps to ensure that bias does not arise prior to its implementation, i.e. during the design and programming stages.
ii. Adopting measures to remedy bias if it has already arisen, i.e. with post-hoc and/or antibias corrective measures.
iii. Taking a more in-depth, openly political approach aimed at making society, from which algorithms arise, more globally just and egalitarian.

As one can easily imagine, this is a huge and by no means simple task. So let's examine it step by step.

4. The introduction of AI in the field of personnel recruitment and selection and the problem of gender bias

As we have already said, the automation – complete or partial – of staff recruitment and selection processes is steadily increasing worldwide, albeit to differing degrees in different countries and economic sectors. According to Hmoud and Laszlo (2019), there has been an emerging trend of using AI technologies within the business world over the last two decades. One recent report estimates that 98% of Fortune 500 companies uses some form of AI-based software to attract workers and in their recruitment processes.

To tackle the question of gender bias in AI and its discriminatory potential in staff recruitment and selection processes, as well as to put forward measures to prevent or at least mitigate it, we need to be able to conceptualise the issue without ignoring its complexity (nor becoming overwhelmed by it), while also conserving as far as possible the layers of meaning. The literature reviewed indicates a lack of definition around the central key concepts: 'there is no singular or unified way of interpreting the meaning of discrimination, or how it might feature in hiring practices, nor is there consensus on any computational criteria for how bias should be defined, made explicit, or mitigated' (Sánchez-Monedero, Dencik & Edwards, 2020). This complexity thus calls for a systemic examination.

4.1. The AI ecosystem

With this goal in mind, we shall hold that fitting within our analysis are a raft of stakeholders involved in different areas and at different levels, associated with the production, distribution and consumption of AI. It is thus obvious that these different stakeholders will have different interests and perspectives regarding the viability and potential pros and cons of the implementation of AI in the world of work, not to mention perceptions of their possible biases. The bulk of them form part of an interrelated whole in dynamic equilibrium that can be understood, metaphorically, as an ecosystem. In this study, we make use of and define the term 'AI ecosystem' as the frame of reference within which the different representations and practices around AI take on meaning. We shall see immediately below which stakeholders make up this system, presented in a random order which has nothing to do with their relative importance.

a) Firstly, there are those looking for jobs and who use an AI-based system to access job postings. Examples in Spain include the LinkedIn and Infojobs web platforms.

b) Secondly, there are the businesses or organisations looking for workers and using some AI-based system to carry out all or part of the recruitment and selection process, and taking take decisions on this basis.
c) Thirdly, there are the tech company workers who are involved at different levels in the design and implementation of AI-based software products (in this case software aimed at the so-called human resources sector of the market). This group could also include the distributors and vendors of these products.
d) Last but not least are the freelance workers whose day-to-day work is mediated (in whole or in part) by web platforms, known in literature on the topic as *platform workers* or *gig workers*. Although our study does not focus on this kind of work, it is a highly important and growing part of the job market, a pioneer in the automation of AI-based processes and which a large number of the studies reviewed examine. Indeed, terms such as 'platform work', the 'gig economy' and even the 'gig culture', have become increasingly popular as descriptors for some of the new forms of precarious work in the 21st century.

4.2. Some representations and tasks assigned to AI

As stated above, within the growing trend of using AI technologies, the perspective taken by the different stakeholders with regard to the AI resources available and their perception of the pros and cons of automating given processes and functions vary significantly depending upon the interests of the party in question.

So, for example, the field of workforce management is seeing a strong boost to the inclusion of AI-based solutions, particularly in recruitment and applicant selection processes (Laurim et al., 2021). Here, AI is strongly represented, and the tendency is only to take into account and evaluate its presumed benefits, neglecting more critical viewpoints, to such an extent that a great deal of literature supports the notion at AI 'is here to stay' (Hmoud & Laszlo, 2019; Upadhyay & Khandelwal, 2018).

Some of the proclaimed advantages are: AI is allegedly more objective (even conceding that it may have biases) than human beings. In this regard, AI-based solutions tend to be marketed and sold to organisations as a resource to neutralise the biases of those people traditionally carrying out recruitment and selection duties, avoiding the recruiter's subjectivity and thereby enhancing companies' talent acquisition. It is worth noting that, in the IT and tech sector, this is something remarkable, given that one of its inherent problems is precisely that of finding optimally qualified workers (Laurim et al., 2021). Indeed, the aforementioned authors speak of a 'war for talent' (a highly evocative image) in which AI is depicted as a good ally for recruiting the best candidates. From this same standpoint, another advantage would be that such software can prevent discrimination against legally-protected groups (women, ethnic minorities, people with some kind of disability, etc.), while also fostering diverse and inclusive workplaces. As we shall see later on, this latter point has become particularly controversial, if not actually questionable, given that most software

currently on sale includes no antibias measures (Sánchez-Monedero, Dencik & Edwards, 2020).

AI also tends to be regarded, in this uncritical and positive representation, as considerably more efficient and effective than people: quicker, more accurate (committing fewer errors), uncomplaining and tireless. A number of arguments are thus frequently couched in terms of its potential for freeing HR staff working on recruitment and applicant selection from repetitive tasks that they themselves regard as more tedious, such as reviewing and carrying out preliminary screenings of hundreds if not thousands of job applicant CVs.

Last but not least, AI is positively portrayed as cost-effective, a euphemism for saying that, in practice, by investing in AI, an undertaking can save on the salary costs of one or more workers.

Turning to the tasks and processes that tend to be assigned to AI-based systems, the context we are examining includes the following examples: recruitment of candidates by means of search engines, and the filtering, verification and evaluation of their credentials and shortlisting those considered best suited or most promising for the interview stage (via so-called *ranking algorithms* or *recommender engines*). From a recruiter's standpoint, such systems can be autonomous or hybrid depending upon the degree of man-machine interaction and human oversight over the processes and their outcomes. The data and inputs can be either structured (e.g. by means of forms supplied by the recruiter and completed by applicants) or unstructured (material submitted by applicants that does not fit a strict template, such as a CV or a cover/motivational letter). The sheer range of formats, language employed and incompleteness of the data furnished can be the source of bias.

Aside from this first screening stage, another practice that is becoming commonplace is for such systems to create and analyse appraisals of applicants' aptitudes by means of different tests or video games. These are based on knowledge of neurosciences and industrial and organisational psychology, and are used to analyse applicants' reactions and behaviour. To this end, metrics for certain cognitive, emotional and social traits are created, on whose basis applicants are profiled and given scores, to in the end be compared with the top performers, whose profiles have been used to train the algorithm (Sánchez-Monedero, Dencik & Edwards 2020).

It is also common, these days, for interviews to be carried out via the Internet, either synchronously or asynchronously, meaning they can be recorded and subsequently analysed by AI-based systems, using a logic similar to that which we have seen for video games. With regard to this latter point, set out immediately below is an interesting case study calling for some discussion of the issue.

4.3. The disturbing case of the analysis of video interviews

What are the questions and what, exactly, is analysed in these interviews recorded and encoded by AI? This is no idle question. Some authors hold that

such analysis processes are a 'black box', an opaque part of the recruitment and selection process that raises a number of questions and doubts, both ethical and legal (Köchling et al., 2020a; Kim & Heo, 2021; Gutiérrez, 2021).

The software designed for this kind of analysis is quite worrying, as some of the systems are aimed at capturing, coding and inferring (based on their biometric data, facial expressions and non-verbal communication) a candidate's feelings, state of mind, personality traits and even sexuality (Crawford & Paglen, 2021).

Based on these studies, we would like to highlight a number of issues that are of general interest, given that these kinds of practices are on the increase:

i. The theoretical minefields implicit in such an inference process raise a whole range of real risks for people. To explain: inferring from a facial expression a state of mind or emotion (or, worse still, an ongoing personality trait) entails the presumption that a necessarily limited number of facial expressions (AIs tend to encode 6 or 7) betray a limited and exact number of feelings. This, in turn, presupposes that absolutely everyone around the world, irrespective of their age or culture, of all genders and at all times, expresses their emotions in the same way (with a limited number of possibilities and thus with no room for individual nor circumstantial variations nor hybrid or simply different forms). What's more, it also implies that one facial expression, captured in one instant by a camera can, after an entire chain of inferences, end up treated as the objective manifestation of a stable psychological trait. This is, therefore, a conceptual model that, in our view, is excessively biologicist, reductionist, determinist and closed.
ii. Nevertheless, for us, the most disturbing aspect of all is the fact that an undertaking feels itself entitled to use this kind of analysis for a job interview and has the freedom (or unregulated room for manoeuvre) to employ it. This point is extremely important, as it entails an intersecting infringement of human rights: privacy, equality, personal data and employment rights. Such an intersection of so many fields and potentially discriminatory factors makes it difficult to grasp and, even more so, lodge any possible claim.

4.4. What does 'solving the gender bias problem in personnel recruitment and selection processes' mean? The cases of HireVue, Pymetrics and Applied

Sánchez-Monedero, Dencik and Edwards (2020) carried out an in-depth analysis of three AI-based recruitment and selection software solutions, produced in the United States (HireVue and Pymetrics) and the United Kingdom (Applied) and highly popular throughout the world. These products were chosen by the authors as being the only ones in their field providing some public information (on websites and in registered patents) on how they work (how they are designed, validated and audited) and on the antibias measures they incorporate. Such information is by no means common and most commercial products do not provide it.

When analysing this software, the authors provided a highly interesting review of the conceptualisation of debiasing measures in algorithms, of which we would like to highlight the following aspects:

i. As we have previously noted, there is no unified definition within the AI ecosystem of the meaning of bias and much less of how it can be operationalised in IT terms. It is clear that maths has difficulties in capturing the meaning of philosophical/sociological concepts such as fairness and discrimination. Even when considering one statistical definition for bias such as 'there be no imbalance between groups (by reason of gender, ethnicity, age, etc.)', it could be called into question and subject to many qualifications depending upon the specific contexts in which it needs to be applied. Accordingly, the conception and operationalisation of what is debiasing is equally or even more difficult. Nevertheless, looking at the kinds of debiasing measures they apply, the commercial software packages analysed lead one to deduce that bias is regarded as a synonym of inequality or discrimination. Such measures take one of the three following approaches:
   a) Anti-classification: trying to correct bias by means of the omission of group variables in the decision-making model.
   b) Classification parity: trying to correct bias by means of the application of equal (or an established proportion of) passing rates between groups.
   c) Calibration: establishing that the outcomes of a screening process are independent of the group variables.

. Notice that all the categories depend on clear definitions of groups. Examples of group definitions are gender (binary or other), ethnicity (using definitions informed specifically by the official categories of a given country's demographic representation system), age interval, etc. It is therefore clear that the definition and delimitation of these groups provides significant opportunity for the introduction of biases, even when the exclusive goal of creating these groups is the application of these supposedly mitigating measures.

ii. Much as with the meaning attributed to the term bias, antibias measures necessarily reflect the social and legal conceptions of the context in which they are created (here, the US and the UK). This means that the commercial exportation of these software systems to other contexts, with other axiological and legal systems for the hiring of workers, raises a large number of questions, as it may give rise to unforeseen mismatches in the outcomes. For example, the bias auditing included in HireVue and Pymetrics uses the 'four-fifths rule' when assessing the fairness of candidate selection. This proportion, which is a legal tool in the US that guarantees that no legally protected group (by reason of gender, age, belief, etc.) can be discriminated against, provides a clear example of how such software features the values and legislation in force in the place in which it has been created. In Europe and other contexts, this rule continues to be applied even though it does not correspond with the applicable legal framework.

Now, from our viewpoint, the authors' most interesting suggestion is that the ethical challenges presented by algorithms, including gender (and other) discrimination, can be tackled most efficiently from the perspective of data protection rather from that of equality or non-discrimination.

We therefore follow these authors in their analysis of Europe's General Data Protection Regulation (GDPR EU 2016/679), and particularly its controversial Article 22, which has been the subject of a fair deal of literature, with regard to the consequences that may arise from it when it comes to constructing the ethics of algorithms and advocating the right to transparency in them. Given the sheer extent of all the issues associated with this subject, we shall limit ourselves to providing brief notes on each and indicating where the hot topics meriting our attention lie. We shall return to these points, laid out here in the form of challenges and questions, in Part 2, when providing the grounds for our recommendations.

1) The right to human intervention in the process.

This right means, strictly, that no process that might have important repercussions/consequences for people, such as the case under review of hiring algorithms, may be left entirely in the hands of a software system without human oversight.

This point quickly gives rise to some ethical questions/challenges, such as: what can be regarded as the minimum and sufficient level of intervention by a human supervisor? For example, in a selection process, is it enough to validate the suggestions made by the AI?

Additionally, as noted in our introduction to the ethics of algorithms, we are faced with the issue of responsibility. Who is ultimately (and to what extent) responsible for possible biases: the software's designers, the supervisor, or the AI itself, regarded as a separate legally-answerable entity?

This also means that if an undertaking seeks to implement a process without any human supervision (in other words, one in which – for example – a hiring decision is completely automated) the data subject must be duly informed and explicitly provide their consent.

2) The right to an explanation.

In the case of hiring algorithms, this right means that the data subject is entitled to receive clear and relevant information on whether parts of a process are automated. This explanation must be comprehensible to someone who is not an IT expert and cannot be, as the authors ironically note, 'a simple regurgitation of source code'.

Thus, a plausible explanation should include meaningful information about the selection process and the logic employed, such as general information on the factors taken into account for the decision-making process and on their

weight in the final assessment. Indication must also be provided on whether the process (including applicant profiling) is totally or partially automated.

The underlying goal is clear: that the processes for which algorithms are responsible not be opaque, that they be made transparent and therefore open to questioning and challenge should the case call for it.

5. Conclusions

Discriminatory practices with regard to gender in staff recruitment and hiring processes affect, among other areas, people's ability to find and hold onto a job, an obviously key aspect for both their subsistence and participating in society. Nowadays, the ways in which decisions are made on who is eligible for which job and why are changing rapidly with the appearance and adoption of automated selection and hiring systems, i.e. AI-based software. As we have noted in preceding sections, this process is becoming opaque and runs the risk of infringing fundamental human rights that have become deeply embedded in democratic societies. This trend has become rooted and is being fed, in part, by a representation and social perception of these systems as more efficient and economical than human workers.

For their part, emerging concerns include a lack of transparency and potential limitations on accessing jobs for specific profiles: in other words, forms of discrimination that, rather than direct, are often indirect, latent and implicit.

This point is crucial, as it can entail the intersecting infringement of people's rights: privacy, equality, personal data and employment rights. This intersection of so many fields and discriminatory facts makes it really difficult to get a proper hold of.

We therefore wish to stress the urgent need to oversee these processes and them make transparent, especially those that, as we have seen, entail the collection and analysis of applicants' biometric and proxemic data, so as to properly protect the public against infringements of their fundamental human rights. These include, for example, the right not to be discriminated against by reason of gender – not in its most reductionist form but rather, as we have tried to show throughout this report, in all its complexity. They also include, however, many other rights arising from the in-depth examination of hiring algorithms from an ethical standpoint, such as the right not to be assessed on the basis of aspects not directly related to the job applied for (as seen, for example, in the unwarranted performance of psychological profiling on candidates). Or applicants' right to a guarantee of privacy and the security of their personal data after the recruitment process has been completed (who will view any interview recordings, what they will be used for, how they will be analysed, what will be done with them later, etc.). And even the right to refuse to be assessed by AI-

based software and opt for human assessment, without this becoming, in turn, a further source of discrimination.

All these rights, previously consolidated in democratic countries, are currently being strained, if not opening violated, with the replacement of some human processes by AI-based systems. Despite the fact that the marketing for much of this software trumpets how it can detect and mitigate discrimination against protected groups and foster diversity and inclusivity in the workplace, such claims are rarely subject to close scrutiny or assessed on the basis of ethical principles, national legislation or international guidelines on the different intersecting issues.

The second part of this report, attached, sets forth the key aspects to be taken from all of this and that in our view merit closer oversight and intervention. It contains a number of general principles and specific measures that can help tackle, in the real world, gender discrimination in hiring algorithms.

PART II. From opacity to transparency

6. Proposed principles and measures

6.1. Proposed general ethical principles

In addition to the expected (and, indeed, compulsory) compliance with the law, we have some ethical requirements that legislation has yet to stipulate or with regard to which is ambiguous. These requirements are framed in terms of some general principles that we shall apply in our study to the recruitment and hiring of workers: those of disclosure, transparency and responsibility.

The reasoned and contextualised application of these principles to every given case should help guarantee that undertakings' staff recruitment and selection processes are not merely lawful but also ethical: that is, that they take into account and care for the interests of all those involved, without committing any discrimination.

The main idea underlying the different measures proposed is that processes wholly or partially assisted by AI must be, as a result of the application of these general principles, understandable, traceable, and auditable (with a view to any potential claims for ethical or legal damages), as well as studiable (with a view to improving or, as the case may be, rejecting them).

We shall then break these general principles down into some of their more specific aspects, aware of the fact that, in many situations, the former will have to be resorted to once again, given the impossibility of predicting every case that may arise in practice.

6.2. Proposed general measures

— Measure 1

In general, recruitment and selection processes must comply with applicable law (country-level and EU directives) on the issue: employment, equality, transparency and data protection and digital rights legislation. Although such a statement may appear self-evident, its practical implementation is no easy task, hence the need for to take the remaining proposed measures into account.

— Measure 2

We recommend creating work teams that are as diverse and heterogeneous as possible in terms of characteristics, including gender parity. We would also encourage the implementation of quotas and ratios, which can be varied in accordance with the given context and technical requirements.

This recommendation tends to be countered by arguing that a woman is not necessarily, due to the fact of simply being one, a gender expert. Nevertheless, we believe that diversity/heterogeneity in a work team provides organisations with a richness (in terms of viewpoints and personal and professional expe-

rience) that is positive in itself. As we saw in Part I, the academic literature reviewed indicates that one of the potential causes of gender biases in algorithms, and hence one of the possible ways of tackling them, is the low presence of women and their lack of job continuity in the AI sector.

— Measure 3

Stemming from the above measure (and the arguments against it), using the services of gender experts and providing training for HR management are measures of proven efficacy in raising awareness of and providing tools and resources for identifying gender biases in both algorithms and human beings.

Continuing with the general principles, but focusing a little more in the intrinsic characteristics of AI-based technology:

— Measure 4

We cannot begin to use AI-based systems and leave them without human oversight. Academic literature provides evidence justifying the monitoring of AI projects throughout their lifecycle, which entails doing so at many different points in time. This obviously causes difficulties, as no one person can be made responsible for all of this. With regard to this oversight, we need to foster a cross-cutting approach involving all stakeholders: in other words, we should not leave this oversight solely in the hands of AI system suppliers, but rather those of supervisory teams and processes featuring the presence of management from the companies using said systems in their recruitment process, and other key players such as workers' representatives and trades union, as well as those responsible for ethical compliance (if existing).

— Measure 5

There is a need to monitor algorithms' construction and training process to prevent or mitigate gender bias: in their design, construction and in the training datasets employed. We once again have to emphasise how algorithm construction and training processes need to benefit, to some degree, from the participation of the players who will subsequently use said algorithms, so as to be able to adapt them into line with the specific context of the relevant company and sector, local legislation and even with appliable specific equality or discrimination policies.

— Measure 6

Algorithms need to be monitored during their on-the-ground implementation for a particular task, so as to avoid/mitigate gender biases due to a particular demand for a given position (bias in the announcement, explicit/implicit/proxy bias in the selected requirements) and/or bias in the real-world datasets from which applicant information is extracted.

— Measure 7

Algorithms need to be monitored after their implementation: a review of the outcomes, both recommendations (matches and predictions) as well as the rejects with regard to the job position's explicit, published requirements, and both

false positives and, particularly false negatives analysed. Here, we advocate the opening up of the information to academic study and technical/legal auditing.

— Measure 8

There is a need to work on formally approving industry standards and independent audit agencies so as to normalise their use. Auditability is now becoming a significant social demand. Undertakings are increasingly expected to implement the good practice of auditing their AI systems and making them transparent.

— Measure 9

Alongside the normalisation of internal audits, information and complaint mechanisms around fully or partially algorithm-based complaint processes need to be implemented. These should be of two types: (*a*) clear and efficient information and complaint options for the company using AI systems, and (*b*) public complaint mechanisms.

— Measure 10

We cannot use AI-based systems if we are unaware of whether they include antibias mechanisms and, if so, what the specifications of the latter are.

In the case of doubt, we should either request the technical specifications and their explanation in accessible language (understandable to those outside the tech sector) or decide not to use the product in question and instead opt for an alternative.

After the sheer amount of research on the implicit bias of algorithms and the proven and validated work of antibias or debiasing measures, we simply cannot allow these measures not to be implemented on compulsory and proactive basis. We cannot remain unaware of whether an AI carrying out an automated process is discriminating (or not): we need to be sure of the fact.

6.3. Proposed specific measures

Operationalisation of the principles set forth in the preceding section could entail at least the following obligations:

— Measure 11

Properly advertise the job, with the most accurate description possible of:

— The details of the position.
— Applicant characteristics that will actually be assessed.
— The details of the selection process.

If a selection process is wholly or partially carried out using AI-based software and entails automated or semi-automated decision-making (i.e. without or with only partial human supervision), this should be disclosed.

— Measure 12

If a selection process is wholly or partially carried out using AI-based software, such software must be identified, providing – at the very least – the following information:

— The product's commercial name.
— The company producing it.
— The software's antibias measures (in both technical and accessible language).
— The training dataset with which the AI has been trained.

— Measure 13

If a selection process is wholly or partially carried out using AI-based software, the applicant characteristics that have actually been coded and analysed must be identified, as must their relationship with the requirements of the job advertised.

— Measure 14

If a selection process is wholly or partially carried out using AI-based software, the parts of the selection process for which AI is responsible must be identified, as must its degree of responsibility (complete or partial) – unsupervised or supervised selection of candidate's CVs, analysis of video interviews, etc. – and the kind of supervision in question.

— Measure 15

If a selection process is wholly or partially carried out using AI-based software, once the process has been concluded, the outcomes (both positive and negative, and duly anonymised) must be published, so that they may be studied and revalidated or, if not, challenged.

A staff selection process must be transparent, particularly in the public sector. If AI systems are used, it should also be open and accessible to study and research, with a view to both improving the software itself and to ruling it out if it is found that it does not provide ethically or legally acceptable outcomes. It must always be possible to complain about or challenge a recruitment and selection process if applicants consider there has been a lack of transparency and improper protection of personal data and/or that discrimination has occurred.

7. Bibliography

Adams, R.; Loideáin, N. N. (2019). «Addressing indirect discrimination and gender stereotypes in AI virtual personal assistants: the role of international human rights law». *Cambridge International Law Journal*, *8*(2), pp. 241-257.

Beauvoir, S. (2015). *The Second Sex* (1949). London: Penguin Random House.

Bolin, A. (2003). «La transversalidad de género. Contexto cultural y prácticas de género». *Antropología de la sexualidad y diversidad cultural*, pp. 231-260.

Bourdieu, P. (2001). *La dominación masculina*. Barcelona: Anagrama.

Bornstein, S. (2017). «Reckless Discrimination». *Calif. L. Rev.*, 105, pp. 1055-1110.

Buolamwini, J.; Gebru, T. (2018). «Gender Shades: Intersectional Accuracy Disparities on Commercial Gender Classification». In: S. A. Friedler & C. Wilson (Eds.), *Conference on Fairness, Accountability, and Transparency (Proceedings of Machine Learning Research)*, *81*, pp. 77-91). Nova York: New York University Press. Retrieved from: http://proceedings.mlr.press/v81/buolamwini18a.html

Butler, J. (2007). *El género en disputa: el feminismo y la subversión de la identidad*. Paidós.

Chen, L.; Ma, R.; Hannák, A.; Wilson, C. (2018). «Investigating the Impact of Gender on Rank in Resume Search Engines». In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1-14.

Crawford, K.; Paglen, T. (2021). «Excavating AI: The politics of images in machine learning training sets». *AI & SOCIETY*, pp. 1-12.

Crenshaw, K. (1989). «Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics». *University of Chicago Legal Forum*. *1989*(1), article 8.

De-Arteaga, M.; Romanov, A.; Wallach, H.; Chayes, J.; Borgs, C.; Chouldechova, A.; Kalai, A. T. (2019). «Bias in Bios: A Case Study of Semantic Representation Bias in a High-Stakes Setting». In: *Proceedings of the Conference on Fairness, Accountability and Transparency*, pp. 120-128.

Fabris, A.; Purpura, A.; Silvello, G.; Susto, G. A. (2020). «Gender Stereotype Reinforcement: Measuring the Gender Bias Conveyed by Ranking Algorithms». *Information, Processing & Management*, *57*(6), article 102377.

Fausto-Sterling, A. (1993). «The Five Sexes: Why Male and Female Are Not Enough». *Sciences*, 33, pp. 1-20.

Fernández-Martínez, C.; Fernández, A. (2020). «AI and Recruiting Software: Ethical and Legal Implications». *Paladyn, Journal of Behavioral Robotics*, *11*(1), pp. 199-216.

Floridi, L. (2013). «The Ethics of Information». Oxford (UK): Oxford University Press.

Gutiérrez, M. (2021). «Algorithmic Gender Bias and Audiovisual Data: A Research Agenda». *International Journal of Communication*, 15, pp. 439-461.

Hangartner, D.; Kopp, D.; Siegenthaler, M. (2021). «Monitoring Hiring Discrimination Through Online Recruitment Platforms». *Nature*, *589*(7843), pp. 572-576

Hmoud, B.; Laszlo, V. (2019). «Will Artificial Intelligence Take Over Human Resources Recruitment and Selection?». *Network Intelligence Studies*, *7*(13), pp. 21-30.

Jahanbakhsh, F.; Cranshaw, J.; Counts, S.; Lasecki, W. S.; Inkpen, K. (2020). «An Experimental Study of Bias in Platform Worker Ratings: The Role of Performance Quality and Gender». In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1-13.

Kim, J. Y.; Heo, W. (2021). «Artificial Intelligence Video Interviewing for Employment: Perspectives from Applicants, Companies, Developer and Academicians». *Information Technology & People*. DOI: 10.1108/ITP-04-2019-0173.

Kleinberg, J.; Raghavan, M. (2018). «Selection Problems in the Presence of Implicit Bias». arXiv. Preprint arXiv:1801.03533

Köchling, A.; Riazy, S.; Wehner, M.; Simbeck, K. (2020a). «Highly Accurate, but Still Discriminatory: A Fairness Evaluation of Algorithmic Video Analysis». In: *Academy of Management Proceedings*, *2020*(1), p. 13339). Briarcliff Manor (New York 10510): Academy of Management.

Köchling, A.; Wehner, M. C. (2020b). «Discriminated by An Algorithm: A Systematic Review of Discrimination and Fairness by Algorithmic Decision-Making in the Context of HR Recruitment and HR Development». *Business Research*, pp. 1-54.

Kullmann, M. (2018). «Platform work, Algorithmic Decision-Making, and EU Gender Equality Law». *International Journal of Comparative Labour Law and Industrial Relations*, *34*(1).

Lambrecht, A.; Tucker, C. (2019). «Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads». *Management science*, *65*(7), pp. 2966-2981.

Laurim, V.; Arpaci, S.; Prommegger, B.; Krcmar, H. (2021). «Computer, Whom Should I Hire? Acceptance Criteria for Artificial Intelligence in the Recruitment Process». A *Proceedings of the 54th Hawaii International Conference on System Sciences*, p. 5495.

Leavy, S.; Meaney, G.; Wade, K.; Greene, D. (2020). «Mitigating Gender Bias in Machine Learning Data Sets». In: *International Workshop on Algorithmic Bias in Search and Recommendation*, pp. 12-26). Springer, Cham.

Leavy, S. (2018). «Gender Bias in Artificial Intelligence: The Need for Diversity and Gender Theory in Machine Learning». In: *Proceedings of the 1st International Workshop on Gender Equality in Software Engineering*, pp. 14-16.

Nadeem, A.; Abedin, B.; Marjanovic, O. (2020). «Gender Bias in AI: A Review of Contributing Factors and Mitigating Strategies». *ACIS 2020 Proceedings*.

Preciado, P. B. (2016). *Manifiesto contrasexual*. Barcelona: Anagrama.

Raghavan, M.; Barocas, S.; Kleinberg, J.; Levy, K. (2020). «Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices». In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 469-481..

Rahman, H. A. (2021). «The Invisible Cage: Workers' Reactivity to Opaque Algorithmic Evaluations». *Administrative Science Quarterly*, article 00018392211010118.

Sánchez-Monedero, J.; Dencik, L.; Edwards, L. (2020). «What Does It Mean to Solve The Problem of Discrimination in Hiring? Social, Technical and Legal Perspectives From the UK on Automated Hiring Systems». In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pp. 458-468.

Sun, T.; Gaut, A.; Tang, S.; Huang, Y.; Eisherief, M.; Zhao, J.; Wang, W. Y. (2019). «Mitigating Gender Bias in Natural Language Processing: Literature Review». arXiv e-prints, arXiv-1906.

Tassabehji, R.; Harding, N.; Lee, H.; Dominguez-Pery, C. (2021). «From Female Computers to Male Computers: Or Why There Are So Few Women Writing Algorithms and Developing Software». *Human Relations*, *74*(8), pp. 1296-1326.

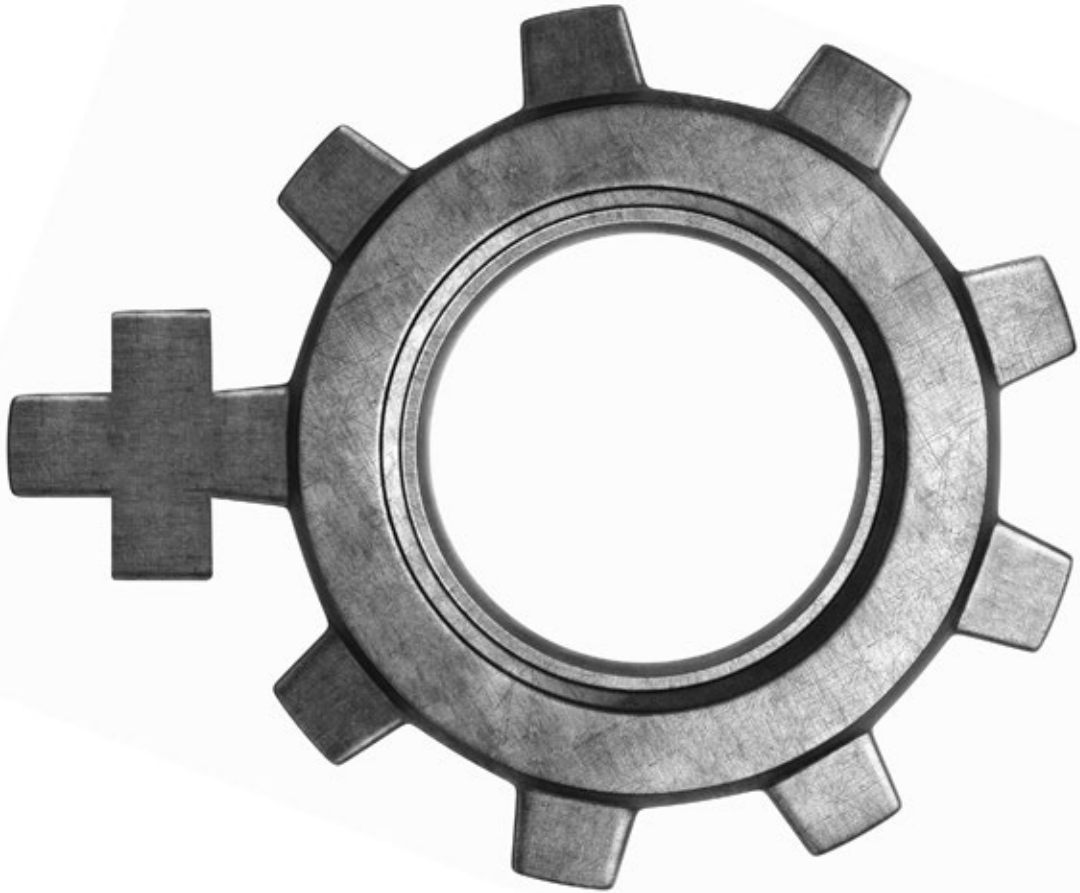
Upadhyay, A. K.; Khandelwal, K. (2018). «Applying Artificial Intelligence: Implications for Recruitment». *Strategic HR Review*, *17*(5), pp. 255-258.

Vasconcelos, M.; Cardonha, C.; Gonçalves, B. (2018). «Modeling Epistemological Principles for Bias Mitigation in AI Systems: An Illustration in Hiring Decisions». In: *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 323-329.

Vyas, N. (2021). «Gender Inequality - Now Available on Digital Platform: An Interplay Between Gender Equality and the Gig Economy in the European Union». *European Labour Law Journal*, *12*(1), pp. 37-51.

West, Sarah Myers (2020). «Redistribution and Rekognition: A Feminist Critique of Algorithm Fairness». *Catalyst: Feminism, Theory, Technoscience*, *6*(2), pp. 1-24. Retrieved from: https://doi.org/10.28968/cftt.v6i2.33043

Yarger, L.; Payton, F. C.; Neupane, B. (2020). «Algorithmic Equity in the Hiring of Underrepresented IT Job candidates». *Online Information Review*. *44*(2), pp. 383-395. DOI: 10.1108/OIR-10-2018-0334.



JORDI DURÓ

The Role of Leadership for Egalitarian and Ethical Organisations

Sara Berbel, Gemma Calvet, Àngel Castiñeira, Oriol Estela, Vittorio Galletto, Jaume Garcia, M. Eugènia Gay, Anna Gener, Pilar Molina, Carme Poveda and Juliana Vilert

1. Project philosophy and context

At the AMB Transparency Agency, promotion of gender equality is a priority line of action in innovation for good metropolitan governance. The Demeter Programme was created and designed in this context and is now subscribed to by the AMB Metropolitan Council, many metropolitan municipalities and associated organisations. Through this programme, presented in 2019 to the United Nations Ethics Office and Department of Social and Economic Affairs, the Agency has pioneered the use of good governance and transparency as a strategic instrument for advancing gender equality in organisations. In addition, the Agency adheres to the United Nations 2030 Agenda, which includes gender equality as one of the Sustainable Development Goals (Goal 5).

Good governance and transparency have been shown to promote gender equality, resulting in better financial management in organisations, a more collaborative and creative working environment and greater use of the talents of everyone involved. Successful organisations of the future will be those that attract and retain talent. In this competition to attract the best talent, companies will have to offer a new, more inclusive, participative, broad, sustainable and egalitarian style of leadership and a working environment conducive to the professional development of everyone who is a part of it.

This year we celebrate the 25th anniversary of the Beijing Declaration and Platform for Action, as well as 10 years since the founding of UN Women, and although there have been major achievements (such as the enactment of the Organic Law for the Effective Equality of Women and Men in 2007 and the Law of Effective Equality of Women and Men in Catalonia in 2015) gender equality and parity are still urgently needed goals.

As a result of certain sexist stereotypes and gender-based role assignments, women continue to bear the brunt of raising children, caring for family members when they are ill and managing the home, tasks that ought to be treated as co-responsibilities, whose functions are equally distributed among all members of the family.

If not, it becomes a barrier – and therefore discrimination – in work and personal terms, as it is women who usually chose to work shorter hours or take a

leave of absence to care for children and family members. Thus women have to make a much greater effort to reach certain positions of responsibility, and this is one of the reasons for, and consequences of, women's under-representation in the decision-making bodies of both public and private companies and organisations. It is worth reiterating that women make up approximately 47% of the workforce in public and private organisations and 55% of graduate employment, but hold only 22% of managerial positions.

Technology has served to create more dynamic and flexible working environments in the context of a global economy. Nevertheless, women generally continue to endure more job insecurity, leading to an intolerable wage gap, the so-called glass ceiling and higher unemployment rates compared to men. The impact of the COVID-19 crisis will also mark a set-back for gender equality.

Is it possible to envision a more gender-equal future in organisations (public and private)? What resistance and barriers need to be overcome? What needs to happen to make it possible? How should the principles of ethics and good governance be applied in terms of gender equality in companies? What good practices exist? What indicators can we use to measure progress?

The project 'The Role of Leadership for Egalitarian and Ethical Organisations' was designed to answer to these questions and its results are presented below.

2. Working methodology

The starting point was the seven objectives of the Demeter Programme and the project focussed on identifying resistance and implementing corrective measures through innovation.

The work was structured in four phases:

— Identifying participants, with the director of the Transparency Agency, based on the criteria that they should be men and women in managerial positions in the public, private and academic sectors, with careers characterised by a commitment to equality.

— Drafting and sending the questionnaire and a letter of invitation by e-mail to the people participating in the project. Ten responses were received. The aim was to learn the opinions of people in leadership positions to contribute their vision and thoughts on issues for improving governance to achieve gender equality in organisations.

— Deliberative meeting with the 10 participants to share the conclusions drawn from the answers to the questionnaires (held online on 1 October 2020).

— Preparation of a final document to compile the answers to the questionnaires, some of them verbatim, and the conclusions from the meeting.

3. Participants

The Working Group was made up of 10 people with long professional careers in positions of responsibility in different professional and functional areas in both the public and private spheres, all of them committed to gender equality.[1] The group complied with the principle of gender equity (60% women and 40% men).

— Director

Gemma Calvet i Barot, Director of the AMB Transparency Agency.

— Coordinator

Carme Poveda, Director of Economic Analysis at the Barcelona Chamber of Commerce and Associate Professor at the Universitat Autònoma de Barcelona (UAB).

— Members

Sara Berbel Sánchez, Municipal Manager of Barcelona City Council and President of the BCN Vocational Training Foundation.

Àngel Castiñeira, Director of the Chair in Leadership and Professor in the Department of Society, Politics and Sustainability at ESADE (Ramon Llull University).

Oriol Estela Barnet, General Coordinator of the Barcelona Metropolitan Strategic Plan (PEMB).

Vittorio Galletto, Head of the Regional and Urban Economy Department of the Institute of Regional and Metropolitan Studies of Barcelona (IERMB).

Jaume Garcia, Director of the Transports Metropolitans de Barcelona (TMB) Legal and Good Governance Department.

M. Eugènia Gay Rosell, Dean of the Barcelona Bar Association (ICAB) and President of the Catalan Bar Council (CICAC).

Anna Gener, President and CEO of Savills Aguirre Newman Barcelona.

Pilar Molina Mesa, Director of Services of the Metropolitan Taxi Institute (IMET).

Juliana Vilert,[2] Director of Organisation and People at Ferrocarrils de la Generalitat de Catalunya (FGC).

1. See the biographical profiles in the Contributors section.

2. The opinions in the questionnaire were discussed and agreed on with one of the coordinators of the FGC Equality Plan, Beatriz Pérez Llorca, head of the Legal and Labour Standards and Procedures Department.

4. Diagnosis of resistance and barriers

Diagnosis and strategy

Although equality is widely advocated and a legal requirement, empirical evidence overwhelmingly shows that: 1) leadership positions in companies and organisations (public and private) are almost entirely held by men; 2) this has consequences for women's material living conditions (an illustrative example can be found in the data on the wage gap in companies and organisations); and therefore, 3) in practice, equality in companies and organisations is not effective if certain measures are not applied to reverse these gender inequalities (positive action in selection and promotion, measures to ensure gender parity and work-life balance and working patterns that help sustain life and protocols for sexual and gender-based harassment, among others).

Firstly, it is important to make clear that such resistance or barriers are not chance situations or the result of behaviours attributable to each person's individual character; on the contrary, they are behaviours, organisational systems, resistance and barriers with social and structural roots (gender stereotypes, traditional discriminatory practices, patriarchal structures and so on).

It is therefore useful to provide data and empirical evidence that demonstrate the structural nature of gender bias. For example, gender inequality and bias in leadership positions in companies and organisations can usefully be illustrated with statistical data on parity in decision-making bodies and power or wage gap figures, among others.

Secondly, once the problem has been framed – and made visible – it is necessary to move on to action to reverse existing gender inequalities. To do this, it is essential to ensure that all companies and organisations have:

— Equality plans in place.
— Committees or reference persons and experts in the field of equality.
— Training in this area for the entire organisational structure and all staff to ensure gender mainstreaming in policy and working practices.
— Specific prevention, detection and protection measures for situations of sexual and gender-based harassment.

How is good leadership defined?

Good leadership has two levels: organisational and individual. The former occurs when the organisation's leaders or managers empower and respect attributed leadership roles, and this is made effective through the design of procedures, circuits and service charters. Lack of a clear, defined framework and recognition from leadership within the organisation affects the correct performance of functions and mission.

In individual terms, as reflected in the conceptions of good governance, good leaders are characterised by their strength in teamwork, the merit of their professional qualifications and their ability to manage issues arising in the team with respect and empathy. In short, good leadership stems from acting ethically and responsibly, setting a good example to partners and promoting positive management and compliance.

It should be stressed that good leadership is not directly related to gender (although women's leadership tends to be more inclusive and collaborative and more geared to transformation), but rather to a virtuous process that simultaneously activates a good project (imbued with values), a good team of followers (who assume the ethics of this role) and a group of leaders who provide vision, motivational relationships and the capacity to transform reality. Talent is diverse by definition, because so is society. Therefore, good male and female co-leadership has to be achieved, because the combination of both will ensure the best decision-making.

Good leadership is leadership that meets three conditions:

— It is efficient and sustainable on a human scale; i.e. it promotes organisational development and people's empowerment, while ensuring that set goals are achieved.

— It recognises the existence of formal and informal power, is aware of its dynamism and complexity and has the tools to deal with and process any difficulties that arise.

— It promotes operational cultures and systems that allow people to improve their ability to contribute and which bring out the best in them.

What resistance and stereotypes exist in the internal promotion of women?

— Preserving the known leadership style (masculine) and the control of the person in power. Resistance to women's promotion stems from a generalised, cultural way of doing things characterised by the decisions of those who dominate the workplace: men. The approach is to promote the person who meets the promotion decision-maker's standards of comfort, thereby avoiding creating uncertainty and insecurity.

— The myth of full commitment in which women always prioritise caring for the home and are not fully committed to work. By contrast, men are considered completely available for full-time employment. It is assumed that women will take most (or all) responsibility for the family and this slows down progress in assigning responsibilities, limits chances of gaining experience and thus negatively affects opportunities for promotion.

— The mirage of equality. Belief that women have already achieved equality because the law says so. This belief assumes that equal opportunities come naturally and therefore it is only a matter of time until women reach managerial positions.

— Women's taking on of leadership and work responsibilities triggers conflicts of priorities and a sense of guilt, because there is a false belief that the gendered role of family responsibility lies exclusively with women as caregivers.

— Lack of motivation and encouragement for women. In some cases it may be women themselves who assume their professional development will be more difficult, and voluntarily accept working conditions (such as shorter working hours, limited availability to travel) that compromise their professional career, possibly turning a potential into a real difficulty. The lack of effective incentives and measures to promote work/family balance prevent access to positions of responsibility based solely on merit.

Overview of resistance in the public sector

The dynamics of public services are based on a rigid, hierarchical, and highly protected model, where the role of trade unions conditions many actions (trade union organisations are mostly led by men) and the status of civil servants and a context in which management is temporary and provisional in nature affect the capacity for managerial leadership.

Appointments to managerial positions in the public sector are still affected by inertias in posts that extend over time, where the posts are occupied by men who, directly or indirectly, come from the political sphere. There is little turnover and the positive action (discrimination) criteria stipulated in regulations, which should balance the presence of women in managerial positions, are not applied to new appointments. This results in a minimal, very weak presence of women. Talk of equality is not matched by specific measures in this area. With regard to technical posts (where women are the majority), these are maintained, as if set in stone, at the same professional category and pressures are solved by raising professional grades.

The management style in government agencies with little turnover in managers is highly marked by sexist and patriarchal stereotypes, with insensitive attitudes (in language, style and methodology) toward promoting women in organisations and in how managerial positions held by women are considered. The few women who do hold managerial positions adopt adaptive attitudes in order to survive and those who do not are clearly ignored or disapproved of. There is still a fear of playing the role of an 'efficient, proactive manager', which entails the risk of clashing with the hegemonic dynamics of highly centralised, top-down control. On reaching managerial positions, women's leadership capacity can easily shift toward a more submissive and adaptive role. When they display this capacity, criticism and questioning of 'style' is latent in an organisation marked by patriarchy. Stereotypes are set in motion.

What barriers should be addressed?

We must aim to shape truly meritocratic environments, so that both women and men have equal opportunities to get where they want to be professionally.

Unfortunately, it is clear that economic and entrepreneurial power today still predominantly lies in the hands of men, due to the following barriers:

1. The first barrier is the difficulty/impossibility of implementing governance-related changes in equality, in dynamic and horizontal parameters, with greater transparency and capacity for innovation.
2. The second is the difficulty in raising the visibility of these situations of 'pressure' on women managers who want to promote change. Control or assurance bodies are also mostly occupied by men, who are complicit in the prevailing dynamics. The experience is one of isolation and loneliness, with little executive capacity to turn the situation into an opportunity for overall organisational improvement.
3. The third barrier is the realisation that the diagnosis has not enabled us to make as much progress as we would like. In other words, data and indicators on gender equality has not, in most cases, been accompanied by corrective measures. Strategic intent in these actions must come from a determined political intent (top-down).
4. The fourth barrier is structural and has to do with the slowness of internal promotion and lack of renewal ('generation block'), worsened in many organisations by either their small size or lack of desire for renewal, especially with regard to positions of trust.
5. The fifth barrier is the difficulty in balancing family life, especially in the age bracket where internal promotions may be more likely ('now or never'), leading to a failure to get everything done and thereby complicating internal promotion. Furthermore, social and work hours are not conducive to work-life balance.
6. Related to the above, there is a lack of social resources for balancing personal and professional life, such as community care services that facilitate personal autonomy.

The approach to overcoming these barriers requires:

— More responsibility on the part of the organisation and the company, facilitating flexibility measures for women who are caring for children.

— More family responsibility in sharing family care.

— More social responsibility: more nurseries and more flexible childcare arrangements are needed while mothers work.

In the specific field of companies and organisations, identifying resistance and barriers to women's internal promotion requires organisations to look inwards. Answers must be sought in the organisational culture which in one way or another means the situation of the people who work in it. Organisational development and personal development are one and the same in how they in-

terrelate: if the organisation does not grow, neither do people, and vice versa. Relevant questions for the initial diagnostic process might be:

— Beyond discourse, what are our practices?
— We have democratic structures; but does everyone feel free to express themselves?
— How does sexism manifest itself?
— Can we look for effective and cost-efficient forms of organisation that care for people's health and well-being?
— What stress levels are we willing to accept?
— What are our decision-making systems like?
— What about our formal power relations? And the informal ones?
— How do we resolve conflicts?

How can we eliminate stereotypes and overcome barriers?

Stereotypes do not change on their own (or they change very slowly). This makes it necessary to raise awareness, train, insist and often establish policies. This means applying gender mainstreaming to everything we do. In the work environment, equality has to be promoted based on policies with specific actions and sometimes with positive (or coercive) action to correct situations of inequality.

The following proposals are aimed at eliminating stereotypes:

a. Designing career development policies through systematic criteria for promotion and managerial renewal. Often, when the need arises, the most suitable person is chosen and appointed from among those closest to management. There is a need for equal participation in administrative decision-making bodies, greater transparency in decision-making and the signing of a code to promote and respect gender equality.

b. Given the lack of female candidates in the initial stages of career advancement processes, efforts should be made to lengthen the list of female candidates by consciously and proactively taking into account or even soliciting the names and profiles of potential female candidates. Breaking stereotypes or prejudices implies, in the initial and intermediate professional stages, empowering professional women with managerial responsibilities and accepting (as with men) that mistakes may be made. At these levels, this helps promote micro- and meso-leadership in organisations and assessment of the degree of achievement.

c. Equal opportunities should start in the staff selection process, applying the rule, procedure or protocol with respect for diversity and without discrimination on grounds of ethnicity, gender, religion, sexual orientation, functional diversity or other similar differences, and maintaining the same criteria in remuneration policies. As only those selected in advance can be promo-

ted, the selection process is key. In culturally very homogeneous workforces or ones where a single sex is in the majority, it may make sense to favour, for a period of time, selection criteria that positively discriminate in favour of under-represented groups. For these mechanisms to work, the recruiting (followed by the promoting) team must be both plural and diverse and incorporate diversity as an explicit organisation value.

d. Best practices in raising visibility for diversity require leadership by example from top management. Organisational culture is primarily determined by management. It is therefore up to management to decide to promote female managers and regularly report (internally and externally) on progress. It should be possible to demonstrate continuous improvement in these areas over time.

e. Raising the visibility of women in the organisation, in its day-to-day activities and achievements, especially in representative and corporate events, functions and communications.

f. Supporting women in their career plan. Promoting with training, covering both knowledge and also leadership building.

g. Women over 45 can act as role models and mentors for younger women to encourage them to be adaptive in their careers, not to be afraid to take responsibility and not to take a negative view of exercising power.

h. Internal information and communication mechanisms are also an excellent platform to exemplify, disseminate and provide visibility for this commitment to diversity.

i. Valuing and encouraging employees to adopt full personal autonomy, comprehensive life projects and take co-responsibility in the social and family sphere.

j. Raising awareness and training employees in gender roles and stereotypes.

k. Encouraging men to participate in reflection groups and co-responsibility workshops.

5. Good practices for promoting equality

What measures should be taken to promote gender equality in organisations?

Internationally, the field of equal opportunities has been implemented as a public policy strategy since the 1970s. Likewise, in Catalonia, current legislation endorses the application of positive action and considers it necessary for achieving equality. However, this is not the only tool for achieving effective equality at our disposal, and nor should it be, because it is not enough to impact on the social structures that generate gender inequality and discrimination.

Promoting equality requires a multi-pronged strategy. Equal opportunity policies and positive action must be combined with gender mainstreaming and the principle of gender equality must be incorporated as a structural element that comprehensively modifies everything we do in all areas, to dismantle the structures that generate inequalities. Therefore:

— Firstly, gender mainstreaming must be applied throughout the organisation.

— Secondly, specific actions to promote gender equality should be promoted.

Among the gender mainstreaming actions, the gender perspective should be included in different areas, especially:

— The institutional culture (some indicators in this area are: are there people or bodies responsible for equality within the structure? Is information broken down by gender?).

— Distribution of staff (possible indicators: distribution of women and men by job position, job category, management and department, professional grade, qualification group, employment status, etc.).

— People management processes (hiring, access to training, promotions, etc.).

— Pay policy and structure (analysis of the gender wage gap).

— Personal, family and working life balance (available measures, people who use them, etc.).

— Communication, language and corporate image (available measures to promote inclusive language).

— Occupational health and safety (risk analysis from a gender perspective, occupational accidents, measures to prevent and tackle sexual and gender-based harassment, etc.).

All these areas are often incorporated in the organisation through equality plans (now required by law for companies with over 50 employees) and protocols against sexual and gender-based harassment.

As for the second group of actions, those specifically aimed at promoting gender equality, positive actions are particularly important. As defined by the Council of Europe's Gender Equality Commission, these actions are 'strategies aimed at establishing equal opportunities through measures to counteract or correct discrimination resulting from social practices or systems'.

It should be noted that the term *positive action* should be used rather than *positive discrimination*. The expression *positive discrimination* has become obsolete and outdated for several reasons. Firstly, there is the paradox in the meaning of the words, since *discrimination* refers to treating people differently (and removing their rights and opportunities) because of their ideology, gender, ethnicity and so on. Secondly, it is also clear that policies for advancing gender equality do not discriminate against anyone and we cannot fall into this mistaken position. By contrast, we propose strategies and tools to redress existing discrimination (which is what positive action does). Therefore, the preferred term is *positive action*.

The importance of promoting positive action in favour of women, especially in promotion and selection processes, but also in seeking ways to make women's leadership visible through internal or public events and through quotas for women's representation, was uniformly supported by all members of the working group. Their necessity will depend on the analysis and purposes of each company. If the analysis shows that the percentage of women is non-existent or very low, parity must be achieved through rigorous measures to comply with the law. In these cases, the quota policy has proven to be effective. It has also been shown that when laws just 'recommend' but do not enforce compliance with quotas, little or no progress is made in closing the gender gap. In short, the voluntary approach does not work.

Examples of good practices

The following is a list of proposed measures to increase visibility for and promote female talent:

— The importance of increasing the reach of the Equality Act and providing it with much more active implementation tools. Unfortunately, more than 10 years after state-wide law was passed, the trends are not encouraging.

— Developing proactive equality actions that go far beyond those outlined in equality plans.

— Taking positive action in managerial recruitment processes. Imposing quotas.

— Implementing protective measures for female managers in situations of harassment and role restriction.

— Encouraging the publication of information on professional profiles, types of contract and salaries (hourly and total) by sex, assuming that the first step in reducing imbalances is knowing the full extent of their existence and magnitude.

— Establishing flexible working hours and other measures to facilitate work-life balance and encouraging co-responsibility in their use.

— Facilitating access to services that facilitate caring and work-life balance for company employees (nurseries, canteens, day centres, subsidies, grants).

— Promoting the visibility of female managers, providing them with resources and support to boost their participation in the media, events, professional networks, etc.

— Raising awareness of unconscious biases that may be impacting on professional activity, with training plans on this issue for as many employees as possible, especially those in management positions.

— Raising visibility of the examples of women who have been personally and professionally successful in different sectors of the economy or politics, to serve as a guide and inspiration for other women.

— Raising awareness of the importance and nuances of truly inclusive language.

— Establishing specific, quantitative KPIs for greater visibility of female talent in the organisation.

— Structuring internal and external positioning programmes for women with managerial potential.

— Offering women leadership of high visibility projects and recognising their role.

— Implementing the 2012 Directive on improving gender balance on company boards of directors, which sets a minimum target of 40% of non-executive board members of the under-represented sex. Making the measure compulsory; it is still voluntary in this country.

— Influencing training and education in order to overcome social stereotypes.

— Encouraging greater participation by women in scientific publications and especially in registering patents, e.g. by introducing a clause to improve the evaluation of a project if the principal investigator is a woman.

— Monitoring the situation of women in society and the economy using key indicators and, above all, publicising results to raise awareness of the existence of discrimination and its cost to society.

— In the public sector, signing an agreement between the government body and trade unions to define the categories most affected by inequality and measures to reduce it. Monitoring rises in staff grades in general and for women in particular. The pay gap between men and women is sometimes hidden by professional grades. They have the same status, but are paid differently because the grade raises the salary.

Gender mainstreaming in all actions and policies may be of particular interest at the current time, in relation to designing post-COVID-19 measures and facing challenges such as climate change, the ageing population, digitalisation (women are less represented in fields supposedly with better future prospects, such as STEM, and in innovation, although women perform better in education according to OECD-PISA results).

6. Principles of ethical leadership in gender equality and governance

How should application of the ethical principles listed below be approached? Can an order of priority be established when implementing them?

The expert group considers that all the ethical principles identified by the AMB Transparency Agency are relevant and need to be implemented. Establishing compliance indicators is sometimes considered complicated, but a number of ways for implementing them and, consequently, measuring the degree of achievement are given below. Stress is also placed on the importance of achiev-

ing these principles through habits, example and daily practice, over and above any positive actions that may be taken to advance their achievement.

It is proposed that all these principles should form part of a compliance programme that implements specific procedures and protocols for action, through the figure of the *compliance officer* with the help of the other areas in the organisation. To be complete, they must include measurement and publication of equality indices in their structures. In the public sector, this would fall under integrity policies.

It is also suggested that there could be a certain order of priority among the principles. The first would be legality, as this concept implies knowledge and responsibility. Accepting obligations and accounting for having done so: accountability. Acting this way, through legal compliance, with knowledge of and responsibility for actions and decisions, also means acting with integrity. Likewise, if organisations are responsible, comply with current legislation and provide explanations for their management, they will be acting from a perspective of gender equality, because this is required by current law and because management will be more efficient, balanced and satisfactory. In short, compliance with the legality principle would include much of what the other principles mentioned above involve.

The second priority would be the group of principles regarding the right to information, objectivity and public ethics, concepts that are more closely linked to the role of the public sector. This is not to say the business world is characterised by contrary precepts, but rather that these three concepts are clearly obligations for those in public employment. Even more so for those who manage and make decisions within the public system that has to serve the public. However, these principles will become increasingly important in private organisations in compliance with the 2030 Agenda.

Principles (and definition of the AMB Transparency Agency) with proposed actions/indicators

— **Legality:** commitment to interest, knowledge and promotion of due diligence in regulatory compliance.

- Identifying the most common cases of legal non-compliance in the organisation's work areas.
- Developing protocols and procedures (and subsequent training in good practices) to ensure regulatory compliance at all times.
- Setting bonuses for managers who fully meet yearly compliance targets.
- Catalogues of positive actions arising from compliance with the Catalan Equality Act.

— **Responsibility:** taking on the obligations and challenges of implementing transparency and good governance and promoting professionalism and service quality.

- Providing the necessary tools to carry out work professionally and ensuring service quality. Monitoring implementation and accountability.

- Drawing up citizen satisfaction questionnaires.
- Establishing a citizens' complaints and advice service.
- Establishing management-by-objectives models and providing regular personalised assessment of their achievement.
- Identifying objective elements that help identify barriers to equality.

— **Integrity:** individual and service actions must be based on the values of independence, honesty, loyalty, respect and firmness.

- Qualitative monitoring of managerial skill development, with protection and empowerment dynamics for the few female managers.

— **Accountability:** commitment to providing information on public management and submitting it to public scrutiny.

- Using annual reporting models that include indicators and metrics providing economic-financial, sector-based, social impact and environmental impact information.
- Online tools help to achieve these necessary, mandatory objectives.

— **Right to information:** recognising the legitimate right of citizens to information on the operation and activity of government agencies.

- Putting the general interest and citizens at the centre of operations.
- Approving and publishing administrative procedures.
- Facilitating citizens' access to the government, without such access requiring a master's degree in computer science, using clear, plain and comprehensible language.
- Developing a model of open administration.

— **Objectivity:** acting impartially and making decisions without personal, subjective or opportunistic considerations.

- Making decisions collectively, in joint bodies.

— **Public ethics:** service in the public interest and for the common good, applying values and principles that must guide government actions.

- Creating a unit or committee in charge of compliance and monitoring the organisation's code of ethics.
- Establishing and developing all elements pertaining to the ethical infrastructure, such as: training and socialisation; conduct guidelines, monitoring and sanctioning systems; leadership and mentoring; human resource policies and value-based promotion and recognition criteria; deliberative spaces on ethical issues, etc.
- Including in the annual report a report based on the degree of compliance with the organisation's values.
- Drawing up questionnaires on code compliance in relation to equality as a tool for good governance.

— **Gender equality:** a fair distribution of rights, benefits, obligations, oppor-

tunities and resources between women and men, oriented towards equality, based on recognition of and respect for differences.

Out of all the ethical principles listed here which are fundamental pillars for ethical leadership and good governance, the one generally most emphasised by the group was the principle of gender equality and gender mainstreaming in company and organisation operation in all policy areas, at all levels and at all stages of policy-making.

The most important measure, agreed on by everyone, as established by law, is the approval of equality plans with specific objectives and an annual review of their achievement. These plans should be implemented in all companies and organisations, but special attention should be paid to the tech industry, which is currently characterised by an under-representation of female talent and which has great future potential.

We must understand that gender equality is not just a question of ethics or justice, but also a means of providing value. Among the conclusions of the 3rd Women Business & Justice European Forum, held at the ICAB on 27-28 June 2019, it was revealed that 60% of companies recognise that the incorporation of women generates profits and that projects led by women are more profitable.

Questions for the design of measures and compliance indicators

— Is there equal representation in management bodies?

— Joint representation: is there a situation that guarantees a presence of women and men whereby neither sex exceeds 60% or is below 40% of the total number of persons involved, and which should aim to reach 50% of each sex?

— Are there work-life balance measures that favour a co-responsible distribution between women and men of market work and domestic and care work, within the framework of a sustainable society model?

— Are there protocols for sexual and gender-based harassment?

— Is there an equality plan in the company or organisation? Is it in force? Are there experts and personnel of reference on equality issues? Is there a specific committee responsible for promoting and ensuring gender mainstreaming?

— Is equality training offered to staff and management?

In conclusion, the seven current objectives of the Demeter Programme are the linchpin of corporate and managerial leadership policies for gender equality in organisations.


JORDI DURÓ

The Challenge of Equality in Organisations. Final Report on the Demeter Cycle

50a50 Association

Introduction

The 50a50 Association organised and conducted during 2020 the cycle on 'The Challenge of Equality in Organisations', as part of the Barcelona Metropolitan Area (AMB) Transparency Agency's Demeter Programme, which aims to promote transparency in gender equality and innovation in good governance.

The Transparency Agency has assumed responsibility for promoting gender equality in the AMB as a priority line of action for innovation in good governance and has promoted the Demeter Programme as an instrument for gender mainstreaming in transparency policy and innovation projects in good governance to contribute to achieving a balance and co-responsibility between women and men.[1]

In this context, the training met two of the seven objectives of the Demeter Programme:

— Promoting the diagnosis of the factors impeding women's equal access to managerial positions.

— Developing instruments and encouraging training for the continuous promotion of equal opportunities.

The cycle consisted of four workshops aimed at identifying the elements needed to reinforce the Demeter Programme objectives and offer reflections and instruments to help advance towards real and effective equality in organisations, through equality plans.

The workshops took place online on 29 October and 3 and 12 November and were attended by 35 to 45 people, depending on the session, mostly from Catalan local and general government agencies.

The cycle was designed by the 50a50 Association, always under the supervision of the Transparency Agency.

One of the unique features of the cycle is that, during the sessions and in conjunction with the workshop participants, the 50a50 Association carried out a

1 http://transparencia.amb.cat/web/bon-govern/programa-demeter.

SWOT analysis, looking at internal characteristics (strengths and weaknesses) and the external situation (threats and opportunities) in the face of the challenges for equality in organisations. Thus, points for improvement were identified in order to continue advancing towards real, effective gender equality in organisations.

From this diagnosis, a series of actions were proposed to the Transparency Agency, with the aim of further developing the Demeter Programme.

SWOT analysis of 'The Challenge of Equality in Organisations'

During the course of the cycle and in conjunction with workshop participants, the 50a50 Association drew up the SWOT analysis for 'The Challenge of Equality in Organisations', which is reproduced below:

INTERNAL STRENGTHS	INTERNAL WEAKNESSES
• Equal access to opportunities • High percentage of women, also in management • Progress toward parity in leadership • Ability of women in management positions to bring about change • Capillary nature of the feminist perspective throughout the organisation and staff's desire for change	• Management denial or disinterest • Management resistance to change • Lack of gender training in management • 'Invisible' wage gap • Shortcomings in monitoring the equality plan • Lack of staff to implement plans • Difficulty in collecting data by sex
EXTERNAL OPPORTUNITIES	**EXTERNAL THREATS**
• Equality legislation • Transparency legislation (open data) • Political support for and incorporation of gender mainstreaming in public programmes • Impact on society of policies applied to government agencies • Feminism rooted in society • Growing culture of diversity • Boost to reforming work hours	• Prevalence of masculine models • Lack of co-responsibility • Link between poverty and gender • Presenteeism in the work culture • Quotas wrongly understood as contradicting the meritocratic principle • COVID-19 as an excuse

Internal weaknesses

The internal weaknesses that were detected are described below:

— **Management denial or disinterest.** The opinions gathered during the cycle suggest that sometimes people in positions of responsibility in organisations do not show interest in the subject or directly question whether equality plans are necessary, as they deny there are inequalities in organisations. In other cases, despite promoting the equality plan (probably for legal reasons), no real policies are implemented to enforce it; thus, there is a mismatch between intention and action, between willingness and efficiency.

— **Management resistance to change.** People in positions of responsibility, especially men, are resistant to accepting a different reality.

— **Lack of gender training in management.** People in positions of responsibility in organisations are not trained in gender mainstreaming. In the case of men, the participants thought their perception of the problem was almost non-existent. In the case of women, they are often aware of the problem, but need training to understand its causes and obtain the knowledge and tools to tackle it successfully. In this context, a need to train people in gender equality was detected, especially for team managers.

— **'Invisible' wage gap.** This weakness refers to the belief that there is no pay gap in government agencies. Once again, there is a need to train staff and educate them on the issue, which, although to a lesser extent, also affects public organisations.

— **Shortcomings in monitoring the equality plan.** This refers to the fact that equality plans are often drawn up as a legal obligation, after which the necessary monitoring is not provided.

— **Lack of staff to implement the plans.** When the will exists within the organisation to carry out and develop equality plans correctly, the human resources to implement them are often lacking.

— **Difficulty in collecting data by sex.** The need to implement mechanisms to provide data broken down by sex was explicitly stated.

Internal strengths

Participants in the Demeter Cycle identified the internal strengths of their organisations in moving toward gender equality:

— **Equal access to opportunities.** Recruitment processes in government agencies guarantee equal opportunities for women and men in access to employment.

— **High percentage of women, also in management.** There is a high percentage of women working in government agencies, including in senior positions.

— **Progress toward parity in leadership.** As a result of the two previous points, progress has been made towards parity in senior positions in government agencies.

— **Ability of women in management positions to bring about change.** Women in management positions in organisations, who are aware of the problem and also trained in gender issues, have, by their nature, the capacity to bring about change and implement measures that help to move toward equality in areas such as time management, replacing presenteeism with more rational work hours and work by objectives, and prioritising traits such as trust and responsibility.

— **Capillary nature of the feminist perspective of the whole organisation and staff's desire for change.** In some organisations, gender training and a feminist perspective have been prioritised, facilitating gender mainstreaming throughout the organisation and getting staff involved.

External threats

Among the external factors affecting gender equality, the following are environmental threats:

— **Prevalence of masculine models.** Society is still marked by the prevalence of masculine models, especially in how roles and responsibilities are allocated from childhood.

— **Lack of co-responsibility.** The burden of domestic and care work and the lack of co-responsibility still weighs heavily on women. Working a professional and a domestic shift, difficulty in balancing work and life and the fact that women are more likely to use reconciliation measures are detrimental to their access to promotion, which undoubtedly influences the pay gap.

— **Link between poverty and gender.** Women, and single parents in particular, are most affected by poverty or the risk of poverty.

— **Presenteeism in the work culture.** Work culture still associates full-time work or physical presence with greater efficiency and productivity.

— **Quotas wrongly understood as contradicting the meritocratic principle.** Temporary measures that help boost parity, such as quotas, are still wrongly seen as contradicting the meritocratic principle.

— **COVID-19 as an excuse.** The coronavirus pandemic has meant that many planned and very important actions related to equality have been put on the back burner, thus serving as the 'perfect excuse'.

External opportunities

External opportunities are discussed below:

— **Equality legislation.** Law 17/2015, of 21 July, on effective equality between women and men, approved by the Parliament of Catalonia, establishes a clear and specific legal framework to advance towards gender equality. The latest royal decrees are also a major step forward.

— **Transparency legislation (open data).** Transparency and open government legislation, known as *open data*, help raise the visibility of inequalities and thus coordinate policy with measurable results.

— **Policy support and gender mainstreaming in public programmes.** Political parties have incorporated equality into their political agenda and gender mainstreaming into public plans and programmes.

— **Impact on society of policies applied to government agencies.** Policies and measures defined by government agencies have an effect on the rest of society, so greater awareness in the management of public organisations represents an opportunity.

— **Feminism rooted in society.** Over recent decades, women have become empowered and feminism has matured to become a well-rooted movement in society.

— **Growing culture of diversity.** There is a growing realisation that diverse teams are more efficient and cost-effective.

— **Boost to reforming work hours.** Although delayed, the reform of work hours to rationalise and optimise time management from a clear gender perspective is now being receiving a strong boost.

Conclusions

Many of the internal weaknesses that hinder progress in gender equality in organisations are a challenge to organisation management. There is a perceived lack of interest or involvement, a lack of resources allocated to developing equality plans or a lack of training in gender mainstreaming on the part of senior officials. In this regard, **awareness-raising and training for top management** in organisations need to be strengthened. As Sara Berbel stressed in her presentation at the second workshop in the Demeter Cycle, it is essential for senior management to be trained in gender mainstreaming, through high-profile, compulsory training. It is also very important to integrate gender mainstreaming and equal opportunities as strategic objectives for the organisation.

In terms of external threats, people stated that the causes partly stem from lack of awareness in society, the predominance of masculine models, lack of co-responsibility and a work culture associated with presenteeism. These are

elements that could also be improved through **awareness-raising campaigns**.

Internal strengths and external opportunities need to be strengthened to move towards gender equality in organisations.

Proposals to develop the Demeter Programme

Based on the conclusions from the SWOT analysis, the 50a50 Association proposed that the Transparency Agency implement a number of initiatives for the Demeter Programme to continue progressing toward its objectives:

— **Awareness-raising campaign aimed at Catalan society**, seeking the cooperation of people already actively working and promoting shared leadership to achieve a fairer and more balanced society.

— **Training programme aimed at senior management and officials in public organisations.** The involvement of organisation decision-makers in the training is essential.

— **Training in diagnosis and design of equality plans**, accompanying public organisations in this project.

— **Mentoring programmes** for organisations having difficulty implementing their plans.

JORDI DURÓ

Recommendation 3/2020, on promoting effective equality between men and women as a metropolitan good governance policy[1]

AMB Transparency Agency

Legal framework

The drafting of Recommendation 3/2020 is based on the following regulations:

- Spanish Constitution
- Statute of Autonomy of Catalonia
- Law 31/1995, of 8 November, on occupational risk prevention
- Organic Law 3/2007, of 22 March, on effective equality between women and men
- Law 7/2007, of 12 April, on the basic statute of the public employee
- Law 5/2008, of 24 April, on the right of women to eradicate sexist violence
- Law 31/2010, of 3 August, on Barcelona Metropolitan Area
- Law 19/2014, of 29 December, on transparency, access to public information and good governance
- Law 17/2015, of 21 July, on effective equality between women and men.
- Royal Decree 901/2020, of 13 October, regulating equality plans and their registration and amending Royal Decree 713/2010, of 28 May, on the registration and deposit of collective bargaining agreements and arrangements
- Royal Decree 902/2020 of 13 October on equal pay for women and men

Objectives of the Recommendation

The objectives of this Recommendation are:

- To promote effective equality between men and women in metropolitan organisations, by means of positive action specific to quality policy instruments, through the Demeter Programme.

1 Recommendation approved on 8 March 2021 by the AMB Transparency Agency, in the context of the Demeter Programme.

— To strengthen the tools for diagnosis, active publicity and strategic definition of equality plans.

— To implement action relating to gender mainstreaming in organisations through equality plans and protocols on workplace, sexual and gender-based harassment.

— To implement positive action specifically aimed at promoting gender equality and women's leadership.

— To develop protocols, procedures and training to act in accordance with the organisation's gender equality and good governance regulations and instruments.

— To offer citizens quality services, transparency and accountability within the framework of good governance, gender equality and transparency.

— To encourage regular assessment in metropolitan gender equality policies.

Background

Contextualisation in the Demeter Programme

This recommendation aims to promote effective equality between men and women as a good governance policy, within the framework of the Demeter programme, promoted by the Transparency Agency and subscribed to by the Metropolitan Council, 19 other metropolitan municipalities, five related bodes and a service provider company.

The Demeter Programme consists of adhering to a gender equality process that includes seven commitments:

1. Incorporating the gender dimension in all data, promoting transparency in diagnosing equality and reinforcing the principles of legality, equity and the promotion of equality.
2. Promoting the diagnosis of the factors that block equality in women's access to management positions, applying the methodology and instruments of the Government of Catalonia that implement Law 17/2015.
3. Publishing and disseminating the results of the gender equality diagnosis.
4. Developing instruments and encouraging training for the continuous promotion of equal opportunities.
5. Establishing effective protective measures against sexual and gender-based harassment.
6. Promoting special performance conditions in public procurement on the promotion of effective equality and a preference toward awarding contracts, in certain circumstances, to tenderers with such measures in place.
7. Assessing the impact on gender equality at the metropolitan level.

AMB joined the Demeter Programme by means of Presidential Decree of Barcelona Metropolitan Area (AMB) of 21 November 2017. The AMB Metropolitan Council was subsequently informed, in the session on 28

November 2017, of the aforementioned Decree approving the implementation in AMB and its related bodies of the proposals contained in the Demeter Programme for transparency in gender equality as an instrument for innovation in good governance.

Participation in identifying lines of action in metropolitan gender equality

Following a quantitative diagnosis, the Transparency Agency created a space for participation and deliberation made up of female workers from AMB and related bodies, who met throughout 2018 and 2019 to establish needs and a qualitative diagnosis.

In 2020, the Agency launched the Gender Equality Working Group in the metropolitan governance area, made up of several public-private decision-makers, which focused on identifying resistance and developing corrective measures through innovation with regard to gender equality in organisations.

From the working group results report, as well as work carried out in the Demeter Programme, conclusions and measures on leadership and equality were drawn up for implementation at AMB:

1. Promotion and retention of talent through the positive action and measures contained in the various instruments relating to organisational equality policies.
2. Planning equality policies. Ensuring that AMB has the instruments and actions listed below:
   - Equality plans in place.
   - Committees or reference persons and experts in the field of equality.
   - Training in this area for the entire organisational structure and all staff to ensure gender mainstreaming in policy and work practices.
   - Specific prevention, detection and protection measures for situations of sexual and gender-based harassment (protocols).
3. Implementation of actions for gender mainstreaming in organisations through equality plans and protocols on workplace, sexual and gender-based harassment, specifically in the following areas:
   - Institutional culture.
   - People management and recruitment processes.
   - Remuneration policy and structure.
   - Personal, family and working life balance.
   - Communication, language and corporate image.
   - Occupational health and safety.
4. Implementation of positive actions specifically aimed at promoting gender equality, in particular:
   - Establishing quotas in appointment, selection and promotion processes for managerial positions and positions of responsibility in order to overcome current discrimination.

- Implementing protective measures for women holding positions in organisations when they experience situations of any kind of harassment, role limitation, etc.
- Encouraging the publication of information on professional profiles, types of contract and salaries by gender.
- Establishing flexible working hours and other measures to facilitate work-life balance and encouraging co-responsibility in their use.
- Facilitating access to services that facilitate caring and work-life balance for company employees (nurseries, canteens, day centres, subsidies, grants).
- Promoting the visibility of female managers and women in positions of responsibility, providing them with resources and support to promote their participation in the media, events, professional networks, etc.
- Raising awareness of the importance and nuances of truly inclusive language.
- Strengthening the visibility of the organisation's female talent.
- Offering women leadership of high visibility projects and recognising their role.
- Implementing the 2012 Directive on improving gender balance on boards of directors, which sets a minimum target of 40% of non-executive board members of the under-represented sex in the governing bodies of the administration and metropolitan public sector bodies.
- Influencing training and education in order to overcome social stereotypes.
- Encouraging greater participation of women in scientific publications and, especially, in the registration of patents, for example, by introducing a clause to improve the evaluation of a project if the principal investigator is a woman.
- Monitoring the situation of women in society and the economy with key indicators and, above all, publicising results to raise awareness of the existence of discrimination and its cost to society.

5. Identifying the most common cases of legal non-compliance in terms of gender equality and good governance in the organisation's work areas.

6. Developing protocols, procedures and training to act in accordance with the organisation's gender equality and good governance regulations and instruments and promote positive action in the organisations.

7. Providing quality services and accountability to citizens in the context of good governance, gender equality and transparency:
   - Drawing up citizen satisfaction questionnaires.
   - Establishing a citizens' complaints and advice service.
   - Establishing management-by-objectives models and providing regular personalised assessment of their achievement.

8. Reporting annually on public management and submitting it to public scrutiny. In particular:
   - Using annual reporting models that include indicators and metrics providing economic-financial, sector-based, social impact and environmental impact information.

9. Adopting ethical principles within the framework of equality and good governance policies in organisations. Ensuring commitment, knowledge and promotion of due diligence in regulatory compliance; taking on the obligations and challenges of implementing transparency and good governance and promoting professionalism and service quality; acting with independence, honesty, loyalty, respect and firmness; ensuring the release of information on public management and submitting it to public scrutiny; recognising the legitimate right of citizens to information on the operation and activity of government agencies and ensuring good practices in organisations and in effective gender equality.

Results of the Demeter Cycle

Based on the conclusions of the SWOT analysis in the four workshops held between October and November 2020, the 50a50 Association proposed that the Transparency Agency implement a number of initiatives for the Demeter Programme to continue progressing toward its objectives:

— Awareness-raising campaign aimed at Catalan society, seeking the cooperation of people already actively working and promoting shared leadership to achieve a fairer and more balanced society.

— Training programme aimed at senior management and officials in public organisations. The involvement of organisation decision-makers in the training is essential.

— Training in diagnosis and design of equality plans, accompanying public organisations in this project.

— Mentoring programmes for organisations having difficulty implementing their plans.

Guide to gender mainstreaming in participatory processes

The guide is the result of concern for gender equality in the Government of Catalonia Directorate-General for Public Participation and Electoral Processes and is part of the project of gender mainstreaming analysis in public participation. The Surt Foundation also took part in 2019.

The guide consists of two blocks, on:

1. Basic concepts, such as participatory processes, gender inequalities and intersectionality, to ensure that we mean the same thing when using these words.
2. The elements needed to achieve fairer and more equal participatory processes, which become a tool for promoting gender equality and respect for diversity in public policy.

Worth highlighting in the guide are the gender clauses in the procurement of dissemination, promotion or synthesising information that facilitate innovation,

given that they prioritise the technical team's capacity and experience in gender equality in participation.

The *Guide to Incorporating Gender Mainstreaming in Public Procurement* provides an in-depth look at the legal basis: Law 17/2015, of 21 July, on effective equality between women and men. Article 10: 'Contracting authorities should endeavour to include social clauses in order to promote equal treatment and equal opportunities for women and men.'

Requirement clauses

Among the technical or professional solvency criteria:

— Ensure that, under equal conditions in terms of the required professional profiles, at least 50% of team members are women.

— Demonstrate gender training, knowledge and skills. In particular, the capacity to prevent and address sexist, LGBTI-phobic or racist aggression.

Tender assessment clauses

— Assess methodologies for achieving equal, inclusive and diverse participation.

— Assess the inclusion of women in situations of social vulnerability in professional teams.

— Assess the training actions planned for professionals assigned to the contract.

— Assess whether the tenderer has a plan for preventing and dealing with all kinds of harassment affecting professionals, participants or other agents involved in the participatory process.

— Consider appointing a person to ensure the implementation and monitoring of gender equality measures.

— Assess the implementation of responsible work-life balance measures beyond legal requirements.

In this context, it is therefore considered appropriate to incorporate the aforementioned gender clauses that reinforce current ones in the procurement procedures carried out by AMB and to assess their impact.

Recommendations for the promotion of effective equality between men and women as a good governance policy in the context of the Demeter Programme

In accordance with the above, the Transparency Agency's powers, the aforementioned legal framework and the outcome of the workshops in the Demeter Cycle carried out from 2018- 2020, the following recommendations are proposed:

1. Approve the AMB Equality Plan and carry out an annual assessment of the impact of the measures established using quantitative and qualitative indicators that facilitate monitoring of effective equality.
2. Create the Demeter women's metropolitan leadership network as an integrating space for debate, with a special presence of metropolitan municipalities that have signed up to the Demeter Programme, to promote and exchange good practices, pilot projects for the socio-economic promotion of women and the development of Sustainable Development Goal Five on the United Nations 2030 Agenda.
3. Develop a training programme on equality policy in relation to good governance, aimed at AMB senior management and officials. This training can be extended to related bodies, service providers and metropolitan municipalities.
4. Draw up protocols and procedures, in coordination with the AMB Human Resources Department, to act in accordance with the organisation's gender equality and good governance regulations and instruments, and to promote positive actions to eradicate discrimination in organisations.
5. Effectively implement gender clauses in AMB procurement, to promote good governance and metropolitan equality, in coordination with the AMB Procurement Service.
6. Carry out awareness-raising campaigns on gender-based violence and equality aimed at the metropolitan population, in coordination with metropolitan municipalities, other government bodies and AMB's communication services.
7. Include gender mainstreaming in actions by the different services and areas, carried out at AMB. This recommendation is proposed in coordination with the relevant AMB services.
8. Hold an annual meeting of the Metropolitan Council that includes accountability on the impact of internal and external effective metropolitan equality policies.
9. Establish stable coordination channels with civil society organisations and different government bodies and institutions on effective policies for equality between men and women.
10. Include the gender dimension in active publicity and the right of access to information and prioritise including the effective equality dimension and promotion of women in the design of metropolitan Next Generation projects.

Ripollet. Crops on the Ripoll riverbed
© KIM MANRESA

CONTRIBUTORS

Sara Berbel Sánchez is Municipal Manager of Barcelona City Council and President of the BCN Vocational Training Foundation. PhD in Social Psychology and postgraduate in Group Analysis and Management from the University of Barcelona. She has held the positions of Director-General of Barcelona Activa, Manager of Barcelona City Council Economic Policy and Local Development, Director-General for Equal Opportunities in the Catalan Ministry of Employment and President of the Catalan Women's Institute. In 2015 she founded Empowerment Hub, a project for personal, professional and social empowerment.

Gemma Calvet Barot is Director of the AMB Transparency Agency. Lawyer and consultant in law and social policy. She has held the positions of Director of the Basque Government Drug Addiction Unit, Chair of the Justice and Human Rights Committee in the Catalan Parliament and member of the Catalan Police Ethics Committee. In the field of equality policy, she has advised public and private bodies such as the Government of Andorra, the Dominican Republic, town councils and the CCOO and UGT trade unions, and has worked on drawing up equality plans for companies such as the Puig Group and Boehringer-Ingelheim.

Àngel Castiñeira Fernández is Director of the Chair in Leaderships and Democratic Governance and Professor in the Department of Society, Politics and Sustainability at ESADE (Ramon Llull University). He is also the Academic Director of the La Caixa Foundation Observatory for the Sustainable Development Goals (SDG). PhD in Philosophy and Educational Sciences. Between 1998 and 2004, he was Director of the Centre for Studies on Contemporary Issues.

Jordi Duró Trouillet is a graphic designer. He has worked in the USA, at Louise Fili's prestigious studios, and later on as a contributor to Pentagram in Paula Scher's team. He is co-founder of the magazine *Scope*, and founder and creative director of the design and branding studio Duró. At present he also teaches at Pompeu Fabra University and EINA School of Design. He has been vice-president of the Art Directors and Graphic Designers Association (ADG-FAD). He has had his own graphic opinion section in the newspaper *Ara* since it was founded.

Oriol Estela Barnet is General Coordinator of the Barcelona Metropolitan Strategic Plan (PEMB). Economist and geographer. A professional in local development and strategic planning since 1995, first as a consultant for government agencies (1995-2005) and later with Barcelona Provincial Council (2005-2016), he frequently gives lectures and training activities and has written various publications. He monitors fulfilment of the PEMB goals and manages the coordination office.

Natza Farré Maduell is a journalist. She works in radio, television and newspapers. She is the author of the books *Curs de Feminisme per microones* and *Que no t'expliquin contes!* (Ara Llibres). She writes a weekly column for the newspaper *Ara* and contributes to the programme *Els Matins* on the TV channel TV3. She also curated the exhibition 'You had to be a feminist', which opened at the Palau Robert in Barcelona, in 2020.

Vittorio Galletto is Head of the Regional and Urban Economy Department of the Institute of Regional and Metropolitan Studies of Barcelona (IERMB). PhD in Economics. Researcher in the IERMB Economics Department since 2009, first as project manager and then as head of department since 2010. He previously worked in the Department of Applied Economics at the UAB as an associate professor and researcher. He has also worked with private sector consultancy firms, specialising in local development projects. His current research at the IERMB focuses on inclusive urban growth, productivity and innovation.

Jaume Garcia Soler is Director of the Legal and Good Governance Department at TMB (Barcelona public transport operator). Law graduate from Pompeu Fabra University. He completed his academic training with several master's degrees to specialise in company, corporate, administrative and contract law. He has held a variety of posts in both the public and private sectors. He previously held the position of Head of the Legal Service and Secretary of the Board of Directors of the Autoritat del Transport Metropolità (Metropolitan Transport Authority, ATM) and the company in which the ATM has a shareholding, Sermetra SL. He also worked in the legal departments of AGBAR, Zurich Spain and Tarragona City Council, among others.

Maria Eugènia Gay Rosell is delegate of the Spanish Government in Catalonia. She holds a degree in Law from the University of Barcelona - Abat Oliba CEU, a master's degree in Mediation from the Barcelona Bar Association and is a graduate from the Harvard Negotiation Institute of the Harvard Law School. She is a lawyer and founding partner of the Gay-Rosell & Solano law firm. She was dean of the Barcelona Bar Association from 2017 to 2022. During this period, she was also president of the Council of Catalan Bar Associations, vice-president of the General Council of Spanish Bar Associations and vice-president of the World Jurist Association.

Anna Gener Surrell is President and CEO of Savills Aguirre Newman Barcelona. Committed to economic and social affairs in Barcelona, she is an active member of a number of organisations. She is a consultant member of the Barcelona Chamber of Commerce, member of the International Advisory Board of the UPF Barcelona School of Management, member of the Social Council of the Ateneu Barcelonès, trustee of the Cares Foundation and trus-

tee of the Museu Picasso Foundation. She is also a member of the *El Periódico* Editorial Committee, the Board of Directors of Mutual Tamaño Cyclops, the RICS Council in Spain and the TechEstate Senior Advisory Board. She writes regularly for various media outlets: *Eje Prime*, *Vía Empresa* and *El Periódico*.

Kim Manresa Mirabet is a European photographer. He has used photography as a tool for social criticism. He has won numerous awards, and his photostories on child prostitution, poverty and female genital mutilation have generated great interest and a new vision of these problems. His photostory 'The day Kadi lost part of her life' has been selected by the agency Associated Press as one of the 100 best photostories of the 20th century. He has produced more than 30 books, some of which have been published in several languages (Catalan, Spanish, Basque, Portuguese, French, English, Arabic, Turkish and Korean).

Guillem Martí Soler holds a PhD in Philosophy from the University of Barcelona. He is a professor of Ethics Applied to Social Education at the UOC (Catalonia Open University) and a contributing researcher at the Ramon Llull University Ethos Chair. He is also a consultant in accessibility and an educationalist specialising in educational activities and projects in contemporary art and disabilities.

Pilar Molina Mesa is Director of Services of the Metropolitan Taxi Institute (IMET). Administrative assistant for 20 years, senior official for another 19 years and currently director of services.

Carme Poveda Martínez is Director of Economic Analysis at the Barcelona Chamber of Commerce and Associate Professor at the UAB (Autonomous University of Barcelona) . With over 20 years' experience in economic research and gender economics.

Mar Rosàs Tosas holds a PhD in Humanities from Pompeu Fabra University. She is currently a lecturer in Anthropology and Bioethics at Ramon Llull University; she coordinates research in Applied Ethics for the Ethos Chair and is editor of the academic journal *Ramon Llull Journal of Applied Ethics*. Previously, she was a lecturer in Catalan Studies at the University of Chicago. Her research has allowed her to author several publications in academic journals and the book *Mesianismo en la filosofía contemporánea* (Herder Editorial).

Maria Teixidor Jufresa holds a degree in Law from Pompeu Fabra University and is a practising lawyer specialising in intellectual property at Vuca Solutions. The first woman to hold the post of Secretary of the Barcelona Football Club Board of Directors, where she was a member from 2015 to 2020 and head of Women's Football (2018-2020), she is currently president of

the Edelmira Calvetó Group for raising the visibility of women. She is also a member of the Board and Executive Committee of PIMEC (Catalan SMEs Employers' Association), where she chairs the Woman and Business Committee.

Francesc Torralba Roselló holds a PhD in Philosophy from the University of Barcelona (1992), in Theology from the Faculty of Theology of Catalonia (1997), in Pedagogy from Ramon Llull University (2018) and in Christian History, Art and Archaeology from the Facultat Antoni Gaudí (2022). He is currently director of the Ethos Chair at Ramon Llull University and chairman of a number of ethics committees. He is a member of the Board of Trustees of the Aldees Infantils SOS children's homes and a full member of the Royal European Academy of Doctors. He has written more than 100 philosophical works, translated into over seven languages. His latest book is *El hermano del hijo pródigo. Del resentimiento a la reconciliación* (Ediciones El Gallo de Oro).

Núria Tria Paradeda holds a PhD in Social Anthropology from the University of Barcelona and is a lecturer in Bioethics at the University of Vic-Central University of Catalonia and in Cultural Transmission and Education at the University of Barcelona. She is a contributing researcher with both the Ethos Chair and the ICT-supported Pedagogy, Society and Innovation research group at Ramon Llull University. Her research on cultural transmission and education processes, medical anthropology, and issues of gender and social exclusion has been published in several works.

Juliana Vilert Barnet is Director of Organisation and People at Ferrocarrils de la Generalitat de Catalunya (FGC). Psychologist, Master's Degree in Senior Management from the Public Administration School of Catalonia. Member of the Board of Directors of the 50a50 Association and former President of the Factor Humà Foundation (2016-2019).

The 50a50 Association is a non-profit organisation that arose out of the Barcelona Chamber of Commerce's Women, Business and Economy Observatory in 2017. With a membership consisting of over a hundred businesswomen and female executives and entrepreneurs, it serves as a lobby for progress toward Sustainable Development Goal 5 of the United Nations 2030 Agenda, on the full and effective participation of women and equal opportunities in leadership in all fields of political, economic and public life.

Sant Feliu de Llobregat. Torreblanca Park
© KIM MANRESA

SUMMARY

LUX BARCELONA

Light (Lux, lucis in Latin). «Physical agent, one of the forms of energy to which eyes react, making things visible».

This magazine, supported by the Barcelona Metropolitan Area, aims to be a converging point for brainstorming and generating useful materials for good governance. It is well aware of the unique opportunity to work for a contemporary humanism which adopts the challenges that the 21st century democracies are facing. Public policies must dig down to the roots for them to guarantee the future of an ethical and fair society. In this matter, the metropolitan administration is a governance tool for cities, and thus it cannot disregard the democratic transformation. This series of texts intend to lead this transformation and shed light on the idea of good governance that had illuminated the world during the Enlightenment.

The Director of the AMB Transparency Agency, **Gemma Calvet**, introduces this fourth volume of the LUX BARCELONA magazine, a monographic issue on gender equity, with a presentation of the Demeter Programme, which wants to incorporate the gender perspective into transparency and data visualisation.

The research work "The ethics of algorithms and the gender perspective: from opacity to transparency", by **Francesc Torralba**, **Núria Tria**, **Mar Rosàs** and **Guillem Martí**, constitutes one of the pillars of this volume and analyses gender biases in the artificial intelligence algorithms. The other three pillars are documents linked to the deployment of the Demeter Programme: "The role of leadership for egalitarian and ethical organisations"; "The challenge of equality in organisations", drawn up by the **50a50 Association**, and the "Recommendation to promote effective equality between men and women as a metropolitan good governance policy", prepared by the Transparency Agency.

Sara Berbel, **Maria Teixidor** and **Natza Farré** reflect, under the "Lighthouse" heading, on ideas relating to the practical application of gender equality policies and the prevention of discrimination against women. As always, the magazine includes several snapshots of the metropolitan landscape captured by the renowned photographer **Kim Manresa** and illustrations on feminism with the particular graphic opinion of **Jordi Duró**.

AMB Transparency Agency