BLACK ON WHITE

(volume 1)

THE IMPACT OF BIASES IN HUMAN RESOURCES THROUGH AI

Index

Artificial Intelligence

Introduction: why the Cambrian influences talent management

Everything is interconnected. Nothing happens by chance. What we are today as a society, the way we live, is the result of millions of evolutionary changes, transformations, and decisions that, in many cases, have occurred randomly, shaping us into who we are. Understanding this reality, delving into it, and analyzing it help us to know ourselves better, to understand ourselves deeply. And that is the key to exercising, in the best possible way, the limited control we have over ourselves and our surroundings.

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

If you have read this far and think that all of this has nothing to do with you, your organization, and especially not with the Human Resources function, you are mistaken. We are talking about people, individuals, and how the heritage of hundreds of thousands of years imprinted in our DNA conditions the way we understand reality and, therefore, defines how we act in our daily lives, both inside and outside of work. And most importantly for this analysis: how what we are has always been transmitted to what we do, and today, more than ever, to Artificial Intelligence.

Let's begin this journey towards understanding in the Cambrian period, about 500,000 years ago. At that time, the most intense burst of life ever known occurred. The Cambrian explosion led to the emergence of an incredible diversity of life on Earth, including many of the major animal groups present today, known through the large number of fossils found worldwide, especially in China, Canada, and Greenland.

From the moment diverse life is generated, identity emerges as a primary defense tool: what I am, against what I am not; what I eat, against what eats me. This fact is not trivial. Over the subsequent thousands of years in which evolution followed its relentless path, "identity" remained imprinted in the primordial DNA of amphibians, reptiles, birds, and mammals.

Now, let's think about the brain. That complex control organ can manage vital functions, making decisions, or calculating risks while processing the thousands of stimuli per second it receives from the five senses. Its sole purpose: to stay alive. From the simplest reptilian brain to the complete central nervous system of the human being, the brain is programmed to survive. In its sophistication, it gained complexity, and our brain is still a black box from which we have not been able to extract many of its operational secrets.

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

But some things we do know. For example, an evolutionary resource of the brain is the ability to "simplify" information to detect dangers as quickly as possible. To do this, it selects information by looking for everything that is "familiar" and categorizes it as "dangerous" or "not dangerous." This is what, thousands of years ago, helped us quickly identify a "saber-toothed tiger" to run away. This instinctive operation has side effects, such as "pareidolia."

Technically, pareidolia occurs when the brain assigns a known identity to any object, area, or landscape in nature. The most recognizable, since we live in society, is the human face. This is why we can see something resembling a face in abstract shapes: eyes, nose, mouth... Think, for example, of the doorknob that Alice opens to enter Wonderland in Disney's famous movie. In the film, pareidolia comes to life: the knob is the nose; the keyhole, the mouth; and the screws holding the handle at the top, the eyes.

Our brain seeks ways to interpret information and adjust it to what it knows, and today, with no wild predators lurking, its main function is to recognize faces and people. Additionally, it promptly identifies the emotions they are expressing through facial expressions and gestures, which can reveal their intentions towards us.

Cognitive biases underlie this biological survival tool. The term "bias" has become newsworthy due to its impact on the functioning of modern Artificial Intelligence (AI). Could we be experiencing a new Cambrian era, a moment of technological explosion where new applications of Artificial Intelligence emerge daily? Let's set aside the answer for now and talk about biases.

Current status: the universality of biases

A cognitive bias is a systematic misinterpretation of available information that influences the way thoughts are processed, judgments are made, and decisions

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

are taken. The concept of cognitive bias was introduced by Israeli psychologists Kahneman and Tversky in 1972. As mentioned before, the brain performs millions of mental processes daily, and biases are the shortcuts it takes to be more efficient. These shortcuts influence how we perceive the world and are determined by cultural implications, social influence, emotional or ethical motivations, information reduction, or distortions in memory retrieval and recall, among many others. Biases are, therefore, inherent to the human being.

There are numerous types of biases, but in the context of the labor market and human resources, we need to focus on the so-called "unconscious biases". Unconscious biases are assumptions, beliefs, or attitudes that we hold but may not necessarily be aware of. While biases are a normal part of human brain function, they often can reinforce certain stereotypes, causing more harm than benefit to companies in terms of hiring, people management, and decision-making.

Unconscious biases develop over time as we accumulate life experiences and are exposed to different stereotypes. These biases include both favorable and unfavorable assessments and are activated involuntarily and without the individual's awareness or intentional control. As a result, they influence our beliefs and behaviors, and when transferred to our professional lives, they affect how we hire, interact, and make business decisions.
We identify 5 major groups of unconscious biases that particularly impact workplace relationships.

1.- Gender bias. It is the unconscious association of certain stereotypes with gender and sexual orientation. This type of unconscious bias affects hiring policies, dynamics of relationships within the company, job opportunities, leadership, and compensation policies. In the realm of Corporate Social Responsibility (CSR), it impacts diversity goals. From a regulatory perspective, it becomes a legal issue for non-compliance, as in Spain, for example, with the

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

Equality Law. It's crucial to note that gender bias is not specifically "feminine." In feminized professions (such as nursing, childcare, and elderly care), an algorithm could bias male applicants.

2.- Age bias. Ageism refers to stereotypes based on age. This bias affects workplace relationships, talent management, and employment policies. The main impact is on hiring policies and contract terminations; it influences organizational diversity and undermines Sustainable Development Goals (SDGs), as well as affecting CSR policies. In terms of public employment, it may hinder older individuals' access to training programs, weakening employability.

3.- Appearance bias. Lookism is based on favorable treatment toward people who fit established beauty prototypes in each culture, involving height, size, hair color, and different abilities. It directly affects hiring decisions, potentially impacting business outcomes. It also influences how we work in teams, professional promotion decisions, and can lead to harassment issues. Studies reveal, for example, that tall and blonde individuals are potentially more "acceptable" in hiring processes.

4.- Intuition bias. Occurs when we let ourselves be guided by first impressions, whether positive or negative. We construct the "whole" of a person with limited information. It affects hiring and promotion decisions, as well as decisions when letting go of personnel. This is the most subjective bias and is based on the experience and learning of each individual.

5.- Confirmation bias. Selects and uses information that confirms one's own viewpoints or expectations. Anything different from me is a potential threat. It impacts race, culture, religious beliefs, origin, social class, or disability. It can affect business decision-making because anything different, distant, or unknown produces an unconscious adverse effect that leads us to discard it. Internally, it causes flaws in the organization's culture.

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

At this point, before proceeding further in our analysis, it is important to differentiate between **prejudice, stereotype**, and **bias**.

- A prejudice is a generally negative opinion towards a person or group, formed without reason, lacking necessary knowledge. It generates a conscious and hostile attitude with the intention of acting accordingly.

- A stereotype is a simplified and broadly generalized image of a group of people who share certain characteristics. It is assumed to be true collectively by a group that often identifies itself as "similar," and it usually lacks nuanced details.

Prejudices and stereotypes form the basis of discriminatory attitudes that impact coexistence and can lead to mistreatment and the limitation of rights.

Biases, on the other hand, underlie much more deeply in our subconscious and, therefore, impact our actions unconsciously and unintentionally. Additionally, they do not identify us with any particular group and can provoke either hostile or cordial attitudes.

The impact of biases in Generative Artificial Intelligence (GenAI) applied to Human Resources

Compiling the expressed thesis so far, biases are inherent to the human condition and influence our decision-making. Starting from this premise, how do they impact the development of AI? We are dealing with a technology based on advanced computational and statistical mathematics. If we add the 'G' for "generative," we endow it with learning capabilities. Generative Artificial Intelligence (GenAI) is based on machine learning models to learn patterns and relationships from a dataset of content created by people. And it is at that point

-"created by people"- where the crux of the matter lies because their biases unconsciously transfer to the algorithm.

Algorithmic bias occurs in cases where a particular algorithmic model based on data consistently produces undesired results for the people developing, creating, and training the system. Often, but not always, this is due to biased collection and use of training data. In other instances, the cause lies in interaction issues between an algorithm and other processes in a specific context. Furthermore, the lack of data may lead the designer to assign more weight to a variable based on intuition, market perception, or previous experience.

In most of these cases, biases have unconsciously become embedded in the algorithmic model. Consequently, its operation yields undesired outcomes and triggers a form of systematic discrimination that typically affects protected or vulnerable social groups.

But how do biases end up there? To comprehend the process, we must consider three fundamental elements in the functioning of GenAI development.

1.- Data. Behind GenAI tools, as we have seen, are algorithms of complex, computational, and statistical mathematics. To achieve the desired "predictions" – for example, identifying the best candidate for my organization – these algorithms must work with millions of data points. Not thousands, nor hundreds of thousands: millions. These data are crucial; it is in this big data where information is naturally biased. Consider a database collecting professional data of the best nursing professionals: in the case of Spain, purely statistically, the most efficient profile would be that of a woman, nursing graduate, and of Spanish origin. In other words: a white, heterosexual, Catholic woman. Although it may seem peculiar, if these data are not considered in programming, anything deviating from this prototype defined as "excellent" will be dismissed.

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

A multinational beverage company, with global expansion, uses a GenAI tool to design training itineraries for employees across all subsidiaries. The tool, subcontracted to a British provider and claimed to be supported by machine learning, is trained with 300,000 data points. The reliability of results from a tool trained with only 300,000 data points for a multi-country project seems questionable.

2.- Global perspective. Considering that developers of GenAI applied to HR are mainly from the United States, we must consider that if the database used corresponds to the U.S. labor market, it will likely have biases that, in Spain, could lead to regulatory non-compliance. In matters of the labor market and human resources, the information used for its application is particularly sensitive. Labor regulations vary by country, as do social diversity and disability protection, for example. Data protection laws also influence this issue, affecting the use of information.

3.- Model adjustments. Consequently, if there is awareness of potential inaccuracies in the database, decisions are sometimes made to emphasize certain aspects in the algorithmic model as defined by the developers or the company using the tool. This delicate process can, once again, cause problems in functionality by unconsciously introducing biases that distort the model's outcome.

Despite this, human resources processes and the labor market, in general, are environments where the application of GenAI could enhance and optimize results. Imagine, for example, a public employment system equipped with an algorithmic tool fed with data from individuals actively seeking employment (profile, education, experience, skills, time unemployed, benefits received, etc.), current job market offerings (experience requirements, skills, public job announcements, etc.), and existing training opportunities for professional

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

development, ensuring optimal system functioning by "predicting" optimal job options for each candidate.

But how do we ensure that a tool with such high potential functions correctly? How do we define its learning model to include former inmates, for example? How can it be trained to see only the talent of each person and not be influenced by biases related to disability, gender, or race?

The first step has been the European strategy on AI, aiming to become a global reference for this technology with a humanistic approach. In the last semester of 2023, during Spain's presidency of the EU Council, provisional agreement was reached on the European regulation on AI use, focusing on safeguarding citizens' rights while not losing development, innovation, and competitiveness opportunities. After three days of negotiations, an interim agreement was reached on the proposed harmonized rules on artificial intelligence, expected to become law in 2026.

This future AI law is a pioneering initiative worldwide seeking to strike a balance between the potential development of secure and reliable AI in the single market by the public and private sectors and, at the same time, stimulate investment and innovation in AI in Europe.

The priority of the European Parliament is to ensure that AI systems used in the European Union are safe, transparent, traceable, non-discriminatory, and environmentally friendly. The challenge is to establish a legal framework that finds the balance between this guarantee of ethical, safe, and reliable AI use and the drive for development, innovation, and competitiveness that AI use entails. Especially considering that outside the EU, the normalization of AI use is not part of the roadmap for countries like the U.S., China, or India, and therefore, its development and application are far from ensuring the goal set by the EU.

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

In our case study, the most important aspect of this draft is the classification of AI into four risk levels:

1.- Unacceptable risk. This category includes all artificial intelligence systems whose use poses an unacceptable risk to security, life, and fundamental rights. Therefore, these developments are PROHIBITED within the EU framework. It includes systems capable of manipulating human behavior, classifying individuals based on their behavior, socioeconomic status, or characteristics, or discriminating through real-time and remote biometric identification systems, such as facial recognition.

2.- High risk. This group encompasses all systems that negatively impact security or fundamental rights and is divided into two categories:

- ➢ AI systems used in products subject to European legislation on product safety. This includes toys, aviation, automobiles, medical devices, and elevators.
- ➢ AI systems falling within eight specific areas that must be registered in a EU database:
    - ✓ Biometric identification and categorization of individuals
    - ✓ Management and operation of critical infrastructure
    - ✓ Education and vocational training
    - ✓ Employment, worker management, and access to self-employment
    - ✓ Access to and enjoyment of essential private services and public services and benefits
    - ✓ Law enforcement
    - ✓ Migration management, asylum, and border control
    - ✓ Assistance in legal interpretation and law enforcement application

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

All "high-risk" systems will be assessed before their commercialization and throughout their lifecycle. This European recommendation gives rise to the AI+Equal project, proposing a model for algorithmic analysis of AI tools applied to human resources processes and the labor market.

3.- Limited risk. These are systems that do not pose a high risk to rights and freedoms. They must be transparent, and users must be aware that they are interacting with a machine so that they can make an informed decision about whether to proceed or retreat, as in the case of chatbots.

4.- Minimal risk. These are systems that do not have implications for user rights and do not fit into the other classifications. Examples include spam filters or AI applied to video games. The vast majority of AI systems currently used in the EU fall into this category.

As we have seen, all AI tools applied to human resources processes and the labor market are considered high risk. The impact of biases in AI applied to these areas represents a setback in the progress achieved in terms of equal rights and diversity management.

Moreover, it violates fundamental rights because it goes against the SDGs in Gender Equality (5); Decent Work and Economic Growth (8); and Reduced Inequalities (10). On the other hand, it constitutes a serious breach of labor regulations, which can have civil and criminal consequences for offenders.

Fundamentals of AI+Equal in the Human Resources field

In 2015, Amazon implemented an AI-based tool for the selection of specialized personnel in software and other technical areas. Due to certain erroneous biases ("black boxes") in its algorithm, the tool preferred and chose male candidates over female candidates. In 2018, Google's AI system identified Black men as gorillas. In 2020, an Italian court ruling determined that the algorithm of the

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

delivery management platform "Frank," used by Deliveroo, penalized absenteeism among delivery riders without differentiating between minor infractions and absences due to illness, disability, childcare, or the exercise of their right to strike (the latter being legitimate and fundamental rights of citizens).

These are just three examples, but reality shows that decisions that were once made by humans are now made by GenAI algorithms, such as those related to hiring or performance evaluation.

As we've explained, the use of these systems carries risks because the data used to train the algorithms are influenced by our knowledge and biases. For this reason, to prevent an algorithm from discriminating against certain groups, it is necessary to ensure that its learning data does not contain any bias.

The Royal Spanish Academy defines an **algorithm** as software code that processes a limited set of instructions. Creating an algorithm requires a very complex statistical, mathematical, and human process, including data collection, preparation, and analysis in various stages, influenced by the decisions of developers and executives. Algorithms can include **unconscious** biases due to poor choices of data, biased data against a group, poor, incomplete, incorrect, outdated, poorly collected, or low-quality data.

Unconscious biases in GenAI applied to tools used in human resources processes -such as personnel selection (e.g., CV categorization), performance evaluation, career development, talent leakage prediction, etc.- have been classified by the EU as high risk. Their inappropriate use negatively impacts equal opportunity policies, diversity management, and access to employment in member states, violating the fundamental rights of individuals, especially those belonging to the most disadvantaged groups.

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

This is because unconscious biases shape the learning of algorithms and end up discriminating against job candidates from the beginning of the selection processes based on their gender, origin, age, background, or disability. All these issues have nothing to do with talent and, besides violating people's rights, prevent companies from incorporating or promoting the best talent, which in the medium term impacts the business results.

If we determine that biases in AI applied to human resources processes are not only affecting business outcomes but also may result in serious violations of labor regulations and considering that the EU has already identified them as high-risk, we conclude that it is imperative to launch projects aimed at finding efficient solutions to this reality.

One line of work is the creation of a normalization and certification model that allows auditing these tools, identifying potential biases, and implementing their removal from the algorithm. Carrying out an empirical test of how algorithms are functioning allows us to take a snapshot of reality to work on experience rather than intuition.

The AI+Equal project, funded by Next Generation funds through the Government of Spain's Transformation and Resilience Plan, is a social innovation initiative in the Community of Madrid that is conducting an audit of AI applied to human resources processes. The initial evidence from the research reveals a lack of technological knowledge on the part of HR leaders in organizations, a reluctance to invest in process improvement in this area—supported by the reputational risk it entails—and the need to enhance internal communication when these tools are applied to counteract the negative image that AI has in society.

Facing ignorance, misinformation, and fear are three basic challenges that must be addressed to promote the implementation of AI in the workplace. This,

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

undoubtedly, would optimize resources, manage talent more efficiently, and could signify a definitive advancement in equal opportunities.

Challenges for Human Resources

Human resources faces the challenge of incorporating GenAI tools into its processes that optimize results. The competitiveness of organizations relies on efficient talent management to attract professionals from the global market and retain them with state-of-the-art evaluation, promotion, and compensation systems. Employee experience has become the key to success for many organizations, and applied GenAI provides the opportunity to analyze data and generate valuable information to achieve it.

However, why do only 14% of professionals in the field work with this technology, even though 76% believe that GenAI will be essential in the future? While many consider it crucial, few organizations have ventured to introduce it into their HR processes, despite successful implementations in other business processes. The challenges are evident:

- Reputational impact. The use of GenAI in HR processes, without the guarantee of ethical and reliable operation, carries the risk of labor lawsuits for normative non-compliance related to equality and diversity. This could negatively impact the brand's reputation. Some companies opt for outsourcing to avoid civil and criminal liabilities, but this doesn't save them from reputation damage.
- Expand knowledge. HR departments lack technological profiles to oversee the acquisition and operation of GenAI tools applied to their processes. Each tool must be trained with appropriate data tailored to the company's objectives and professional profiles. Advisory multidisciplinary boards should supervise the selection of data used to train the algorithm and the design of the learning model.
- Focused communication. Controversies generated by media reports lead to public rejection. The "irrational" fear of being dismissed by an

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

algorithm prevails in organizations, and the implementation of GenAI tools must be accompanied by an appropriate information and awareness campaign to ensure usability.

- Regulatory guarantee. Labor legislation must accompany the framework for the use of this technology. Despite new standards, such as the UNE19602 Labor Compliance approved in July of this year, there is still much work to be done.

Conclusion

In the last two decades, significant progress has been made in the field of equality and diversity in the workplace. Regulations have been developed to promote progress toward equal opportunities, with organizational models that advocate for the inclusion and diversity of all groups, recognizing the value of diverse talent and different capabilities. All these advances, achieved not only through legislative and administrative efforts but also from organizations through Corporate Social Responsibility plans, compliance with the Sustainable Development Goals (SDGs), and the 2030 Agenda, are at risk of being lost if we do not establish monitoring systems to ensure the ethical and reliable functioning of GenAI.

There is no doubt that GenAI has come to stay, becoming an essential tool in the field of HR soon, and companies can only be competitive if they can attract the best talent, wherever it may be.

To achieve this, and following the EU recommendations, which are expected to become law in 2026, it is essential to establish a working model that allows for systematic auditing and certification of algorithms applied in GenAI, starting with those identified as **high risk**. It must be a scalable and personalized algorithmic certification model for each area of AI use, conducted by external certifying entities working with international ethical and reliable standards.

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

HR processes and those related to the labor market are a priority area to establish these mechanisms of supervision, analysis, and algorithmic control. Biases represent the focal point on which we must focus our efforts. The goal is twofold: on the one hand, to encourage companies to hire the best talent to boost their competitiveness, and on the other hand, to ensure equal opportunities for all groups, even those identified as "vulnerable".

GenAI has no consciousness. It doesn't know what is true or false. It doesn't know what is right or wrong. And it does not question itself for the results it obtains when analyzing data. Not yet, at least. That's why it is essential to establish "humanistic" work systems around purely "technical" processes. The AI+Equal project aims to create an inclusive GenAI that ensures access to the job market is managed in an equitable and fair manner, based solely on the candidate's talent and suitability for the position. The AI+Equal management platform allows configuring the analysis model to examine unconscious biases in algorithms not only from a gender perspective but also from discrimination based on age, sexual orientation, degree of disability, origin, religion, etc.

To achieve this, in addition to technology, it is necessary to innovate in decision-making processes when acquiring this technology; update the knowledge that HR departments have on the subject; educate and raise awareness among professionals about the potential negative repercussions of their work due to these unconscious biases; and accompany society in its own digitization process, overcoming the fear of using AI but understanding the risks of its inappropriate use. So far, we have entrusted the development of technology to technicians: the result has been increasingly advanced, sophisticated, and efficient technology. In this new Cambrian period of the explosion of new forms of GenAI, we are not considering the transmission of our "human flaw" to its learning model. The time has come to make it **black on white**.

Marisa Cruzado
CEO of AI+Equal

Financiado por la Unión Europea NextGenerationEU
GOBIERNO DE ESPAÑA
Plan de Recuperación, Transformación y Resiliencia
Comunidad de Madrid
Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES

Bibliography

Díaz Mairal, Francisco (2023): *Cultura del dato, inteligencia artificial y RRHH*. Capital Humano, 386.

Fuller, P., Chow, A., Murphy, M. (2023): *Sesgos inconscientes: cómo reformularlos, cultivar conexiones y crear equipos de alto rendimiento*. EE.UU.: Conecta

Raghavan, M., Barocas, S., Kleinberg, J., Levy, K. (2020): *Mitigating bias in algorithmic hiring: evaluating claims and practices*. FAT* 2020: pp. 469-481. https://doi.org/10.1145/3351095.3372828

Rueda, Francisco J. (2023): *Retos pendientes en el uso de la Inteligencia Artificial en el sector de los Recursos Humanos*. CEPAL.

Financiado por la Unión Europea NextGenerationEU

GOBIERNO DE ESPAÑA

Plan de Recuperación, Transformación y Resiliencia

Comunidad de Madrid

Dirección General de Evaluación, Calidad e Innovación
CONSEJERÍA DE FAMILIA, JUVENTUD Y ASUNTOS SOCIALES