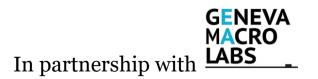# #Leplusimportant

"Taking back control of AI at work: 20 proposals to promote responsible algorithmic management"

WHITE PAPER

By Jérémie **Giniaux-Kats**, Director of the Algorithmic Management Unit at #Leplusimportant, with Erik **Campanini**, Mathias **Dufour**, Florian **Forestier**, Pauline **Maury**, Sylvie **Teynier**

GENEVA MACRO LABS

In partnership with

October 2021

Presentation of the White Paper

Algorithmic management, i.e. the use of AI to recruit and manage the workforce and organise human work, is in the process of disrupting work, its organisation and the relationship between employers and workers. The combination of two famous expressions: "*software is eating the world*" and "*code is law*" gives a glimpse of the exceptional scale of this upheaval, which potentially affects all companies and workers.

While labour law is still very much focused on the organisation of pre-digital labour relations, taking the measure of the challenges posed by algorithmic management on work and acting now to frame it is therefore absolutely critical. It is a question of allowing its potential to be exploited while guaranteeing working conditions that preserve the dignity, health, equal treatment and autonomy of workers.

To this end, this draft White Paper provides the first elements of an **analysis of the** issues and **guiding principles**, broken down into **4 areas of** concrete **proposals**
1. *PREVENT* **- Adopt safe design rules for responsible AI at work**
2. *PROTECT* **- Adapting individual workers' rights to the transformations of work brought about by algorithms**
3. *GUARANTEE* **- Ensure effective compliance with the principles of responsible AI and workers' rights**
4. *UPDATE* **- Allow the regulatory framework to adapt over time to developments in AI**

This White Paper is published in partnership with Geneva Macro Labs on the occasion of the international conference "***For a Web of Enlightenment***: *putting digital culture back at the service of the Enlightenment emancipation project*" organized on October 21, 2021.

It completes our first contributions on the subject:
- Article "Algorithmic management, a new paradigm of work organization ? in AOC in April 2021
- Article "Faced with platforms, we need a real social right for the algorithm", in Politis of January 2021
- Article "Counterbalancing the power asymmetries arising from the digital society: towards a social right of the algorithm" in the CIEP's Cahier n°29 *"Inégalités programmées. Capitalism, algorithms and democracy"*
- The video replay and podcast of our June 2021 webinar "*Work and algorithmic management: what's at stake? How to regain control? "*with speakers from the ILO, OECD, Stanford and Salesforce with our partner Geneva Macro Labs

We warmly thank Guillaume Bon, Anna Byhovskaya, Odile Chagny, Ekkehard Ernst, Clémentine Kleitz, Alexis Masse, Franck Morel, Louis-Laurent Preux, Eva Thoreau and Paul Welti for their contribution to this document, <u>which is only binding for #Leplusimportant</u>.

Summary of the White Paper and summary of proposals

Summary

While labour law is still very much focused on the organisation of pre-digital labour relations, it is absolutely critical to take the measure of the challenges posed by algorithmic management on work and to act now to frame it. It is a question of allowing its potential to be exploited while guaranteeing working and employment conditions that preserve the dignity, health, equal treatment and autonomy of workers.

The purpose of this White Paper is to provide initial analysis of the issues, concrete proposals for promoting responsible algorithmic management and, in the appendix, proposals for adjustments to the European Commission's draft Regulation.

I. Algorithmic management: a reality already present, major challenges

Algorithms have long been used to optimize the functioning of machines. Insurance, finance, weather, examples abound. What is new is their use to **organize human work** and **manage human resources** within companies.

In the [European Commission's draft regulation on AI](Annex III, Recital 36) two employment and work-related topics are identified as "high risk": (1) **recruitment** and (2) workforce **management**: decisions related to promotion, dismissal, **task allocation**, monitoring and evaluation of workers' performance and behaviour.

Indeed, among the use cases that are developing are the use of AI for recruitment (including screening applications, evaluating candidates through interviews, tests or automated video analysis) or for making decisions about the promotion and termination of contractual work relationships, for the allocation of tasks and for monitoring and evaluating the performance and behaviour of individuals in these relationships.

If these uses of AI in the workplace are developing today particularly in distribution, call centers, warehouses, they will become widespread in all sectors. According to a Mercer study of 7,500 HR managers worldwide, **in 2020** 39% of HR departments will be using predictive algorithms, compared to 10% in 2016, a **fourfold increase in 4 years**! In this same study, the generalization of AI appears as the third strategic lever to face the crisis.

In total, the ILO distinguishes **5 elements of algorithmic management**: constant monitoring, permanent evaluation of performance, automatic application of decisions without human intervention, interaction of workers with a system, low transparency of algorithms.

AI at work, like any major technological innovation, represents both **opportunities and threats for workers.**

While the use of AI by employers is primarily driven by its ability to improve productivity ([Accenture reports 40% productivity gains globally by 2035](#) thanks to AI), it can have several benefits for workers themselves. It has the potential to make their jobs more productive and easier, less subject to arbitrariness, and more interesting. A [study of 10,000 workers](#)shows an increase in job satisfaction for those who use algorithms in their work. More specifically, AI can help ensure that workers are treated with less bias and discrimination and are treated in a more personalized way in relation to their needs and aspirations.

Conversely, algorithmic management also opens up **4 main types of [negative risks](#)** for workers: (1) abusive surveillance, (2) discrimination, (3) loss of autonomy and psycho-social risks and (4) obsolescence of skills and loss of employability.

(1) First, the exceptional development of capabilities for machines to collect and process personal data about employees' work and lives - in other words, to **monitor workers** - is the backdrop to the development of AI at work (like *"[smart badges](#)"* that measure the speed at which workers perform tasks, track their interactions, and analyze their conversations, or AI tools that analyze workers' emails to monitor their productivity, spot the most innovative workers, or deviant behavior).

For the [European Commission's Joint Research Centre](#), algorithms thus function as "*surveillance instruments that replace direct supervision and create power asymmetries*".

(2) AI could also lead to **increased discrimination**, even if it is [unintentional](#)or even unwanted, but linked to the opacity of the algorithms' actual decision-making criteria. [Amazon, for example, had to abandon a recruitment algorithm program](#), having found that it disadvantaged female applicants for tech jobs.

Furthermore, the widespread use of algorithms in employee compensation increases the risk of "[algorithmic collusion](#)" across multiple employers, creating conditions for lowering working conditions.

(3) Algorithmic management will lead to a growing number of workers having their work influenced or even directed by processes that they do not control. The perception of a transfer of decision-making capacity from the worker to the machine feeds fears of **loss of autonomy**, including among skilled workers and managers. This can even lead to a **feeling of alienation**, linked to the perception of becoming an appendage of the machine.

(4) Finally, AI in the workplace, which is already the cause of salary disparities, poses a **risk of obsolescence of skills**, internally, and of **reduced employability**, externally, for a certain number of workers, which in turn calls for specific measures on the part of companies, or even public authorities, to train them in the various ways of working with AI.

More broadly, **algorithmic management presents 5 types of** issues**: explicability**, **objectivity**, **responsibility**, **adaptation of workers' rights** and finally, more broadly, **value sharing**.

1. How can we guarantee the **explicability of algorithms**, whose operating rules are currently opaque? This question directly engages, for example, the **regulators' capacity for certification**, to verify that the algorithms respect the normative framework. The tragic consequences of the validation by the American regulator of the MCAS software used to assist the piloting of Boeing 737 Max aircraft, without actually understanding it because it lacked the required expertise, show the extent and topicality of these issues.

2. How to guarantee the **objectivity of algorithms** and avoid **bias and discrimination**? As an example, the Bologna court recently ruled that Deliveroo's algorithm, which used a "reliability" index to rank delivery drivers and thus prioritize the allocation of deliveries, was discriminatory.

3. Who is **responsible for** the functioning of algorithms and their results? Algorithms blur the identity of the principal, and even more so by substituting "*nudge"* for injunction. These forms of influence do not fall under the traditional exercise of authority and pose real challenges to the law. What is left of human responsibility, how can it be attributed?

4. How to adapt **workers' rights to the era of algorithmic management (and maintain their employability)?** How to rebalance the relationship between the worker and the employer in the face of algorithmic management? How can we guarantee effective access to these rights, whose complexity risks making them ineffective?

5. *Last but not least*, the **sharing of value is** a key issue: how much of the economic value created is redistributed to the workers concerned? To what extent can workers claim and obtain their fair contribution to the success of the company?

II. Guidelines for promoting responsible algorithmic management

The challenge is none other than to **regain control of AI at work: to promote responsible algorithmic management**. To exploit its potential while ensuring working conditions that preserve the dignity, health, equal treatment and autonomy of workers, it is essential to define what responsible algorithmic management would consist of: to this end, we propose several **guiding principles**.

10 business guidelines for responsible algorithmic management

The use of AI by companies to organise human work is, first and foremost, a matter of entrepreneurial freedom: it is a way for them to improve their efficiency and better fulfil their missions. The constraints that should be imposed on them must be strictly necessary and proportionate to the goal they pursue (such as the protection of the fundamental rights and freedoms of individuals).

They are thus entitled to ensure that the regulatory framework for algorithmic management respects **intellectual property**, guarantees **fair competition** between players, preserves the capacity for **innovation** and guarantees their **legal security**.

On the other hand, the very high stakes linked to algorithmic management impose specific **responsibilities** on companies in this area: **guaranteeing the rights of their workers** (*see below. Guiding principles for workers)*, **informing workers** about their practices regarding the use of AI to organise work, in particular about the nature of the personal data collected, **training their workers** in algorithmic management practices, **investing in the employability of workers** if this is called into question by AI, **making AI and its uses at work a topic for social dialogue** and **lending themselves to external audits** to check compliance with the above points.

10 guiding principles to protect workers and enable them to get the most out of AI

To prevent the risks posed by algorithmic management for workers (abusive surveillance, discrimination, psycho-social risks, loss of autonomy, loss of employability), we propose **10 guiding principles**, relating to their **right to dignity** (right to privacy, right to information and consent, right to human interaction), their **right to equal treatment** (right to non-discrimination), their **right to health** and their **right to autonomy** (right to understanding or explicability of algorithms, right to human decision, right to challenge, right to training in the use of AI and right to negotiation).

Finally, it is clear that algorithmic management raises issues that call for specific intervention by **public authorities**. This is based on a **fourfold responsibility**: to have a **deep and autonomous expertise** to be able to produce relevant and independent regulation on AI at work, to guarantee **effective access to rights** (that the rules enacted are well respected and the guaranteed rights are effectively accessible, especially for individual workers), to ensure **support for employability** and finally to ensure the **adaptability of the regulatory framework**, which must not be defined once and for all, but must be highly evolving, at the risk of losing its relevance.

We have identified 21 proposals to translate these principles into concrete actions, organized into 4 main lines of action:

AXIS 1: PREVENT - Adopt safe design rules for responsible AI at work

<u>Proposal #1</u> - Introduce a **goal of algorithm explicability** and ethics by design

<u>Proposal #2</u> - Ensure the **traceability of algorithmic** decisions

<u>Proposal #3</u> - Make certain **mandatory certain sequences** within algorithms to guarantee workers' rights and conversely prohibit sequences with deleterious effects

<u>Proposal #4</u> - **Supervise the design** of algorithms, in particular through upstream control of the learning databases to ensure their diversity and minimize the biases of the algorithms

<u>Proposal #7</u> - Put in place a **framework** framework **of the model life cycle** to minimize bias and the risk of model quality drift over time

AXIS 2: PROTECT - Adapting workers' rights to the transformations of work brought about by algorithms

<u>Proposal #9</u> - Introduce a true **right to a human decision**

<u>Proposal #10</u> - Introduce a **right of communication** of an intelligible and opposable document on algorithms at work

AXIS 3: GUARANTEE - Ensure effective compliance with the principles of responsible AI and workers' rights

<u>Proposal #12</u> - Move from vendor self-assessment of regulatory compliance of algorithms to assessment by **assessment by third party organizations**

<u>Proposal #13</u> - Implement risk **procedures for estimating risks** algorithms, on the model of the banking stress test

<u>Proposal #18</u> - Build up a **public expertise capacity** and train the relevant administrations to enable public authorities to produce relevant and independent regulation on AI at work

FOCUS 4: UPDATE - Allow the regulatory framework to adapt over time to changes in AI

<u>Proposal #21</u> - To make algorithmic management a **topic of social dialogue** in its own right

Historically, as _____ Alain Supiot reminds us, trade unions have fought "[against exploitation at work, not against oppression at work](#)", which has led to the evacuation of "the content and meaning of work from the notion of social justice".

At a time when there is a growing awareness of a "loss" of meaning in work, linked in particular to a feeling of loss of autonomy on the part of workers, this is perhaps, in *the*

end, **the fundamental challenge posed by algorithmic management: to put the meaning and content of work back at the centre of the debate**.

It is only through the awareness of all actors - companies, workers, trade unions, public authorities - that we will find appropriate responses. But even though we are only at the beginning of the deployment of algorithmic management and its effects on work, the scale of the issues and their topicality make it necessary to act now.

In this respect, the draft Regulation on AI proposed by the European Commission represents a major step forward. It is the first attempt in the world to regulate AI systems horizontally and to prevent their possible misuse. It clearly states the prohibitions and the essential points of vigilance.

However, this text has significant room for improvement, particularly in relation to the gap between the aims pursued and the insufficient or ineffective nature of the measures put forward to achieve them. It also takes insufficient account of the specific challenges of AI at work. A summary of our proposals for adjustments to the European Commission's text, centred on 2 priority axes, can be found in the Annex:

1. adapt the draft regulation to enable it to go beyond the declaration of intent and enforce workers' rights,
2. ensure, given the singular complexity of AI-related subjects, the competence of all the actors involved in this regulation, at national and European level.

I. Algorithmic management: a reality already present, considerable potential and major challenges

1A. Algorithmic management: what are we talking about?

What do Boeing 737 Max aircraft grounded because of a failure of their flight assistance software, apps to get delivery at home in a few minutes and the growing use of 100% automated online recruitment solutions have in common?

Algorithmic management, i.e. the use of AI to organize human work.

Algorithms have long been used to optimize the functioning of machines. Insurance, finance, weather, examples abound. What is new is their use to organize work and manage human resources within companies.

In the European Commission's draft regulation on AI (Annex III, Recital 36) two employment and work-related topics are identified as "high risk": (1) **recruitment** and (2) workforce **management**: decisions related to promotion, dismissal, task allocation, monitoring and evaluation of workers' performance and behaviour.

Indeed, among the use cases that are developing are the use of AI for recruitment (including screening applications, evaluating candidates through interviews, tests or automated video analysis) or for making decisions about the promotion and termination of contractual work relationships, for the allocation of tasks and for monitoring and evaluating the performance and behaviour of individuals in those relationships.

For example, the widespread use of telecommuting, with Covid, has led companies to recruit candidates entirely remotely and to increase the use of AI-based recruitment software. L'Oréal, for example, uses AI to manage its 100,000+ annual internship applications. The writing of job offers, the search for candidates and even their selection can now be quasi-automated. This growing use of AI is fuelling debates on the reinforcement of discrimination in hiring. Even in the case of Amazon workers whose contracts were terminated by algorithms.

Beyond the application of AI for recruitment and to support HR management processes, the heart of algorithmic management is the organization of human work, which takes different forms. First, a hyper-standardization of tasks for workers, with AI leaving tasks not yet automated to humans. Second, a hyper-complexification of tasks for workers,

with AI supporting the simplest tasks. Finally, the development of relationships between humans and collaborative robots or cobots.

If these uses are developing today, particularly in distribution, call centres and warehouses, they will become widespread in all sectors. According to a Mercer study of 7,500 HR managers worldwide, **by 2020** 39% of HR departments will be using predictive algorithms, compared to 10% in 2016, a **fourfold increase in four years**! In this same study, the generalization of AI appears as the third strategic lever to face the crisis.

The concept of "algorithmic management" has recently been coined by researchers to describe the mode of work organization introduced by digital work platforms like Uber and Lyft. But far from being limited to these platforms, it actually concerns the entire economy. A certain number of features have actually been present in the United States since the late 1980s (particularly in the retail sector), but they have taken off with the deployment of the current wave of digital transformation and the technological disruptions that have accompanied it - cloud, big data, mobile applications, geolocation, the Internet of Things, learning machines, etc. Amazon, for example, is one of the most advanced companies in terms of algorithmic management of its employees.

In total, the ILO distinguishes **5 elements of algorithmic management**: constant monitoring, permanent evaluation of performance, automatic application of decisions without human intervention, interaction of workers with a system, low transparency of algorithms.

Some researchers **consider management by algorithm as a new paradigm of organization and control**. Indeed, it profoundly modifies each of the three dimensions of the "subordination" (in the legal sense of the term) of salaried workers, namely the mechanisms of direction (algorithmic directives replace hierarchical directives), of evaluation (based on continuous algorithmic monitoring), and of sanction and retribution (linked to compliance with the algorithm's directives).

Management takes the form of both algorithmic recommendations that encourage workers to adopt pre-determined behaviors, or the inability to modulate or customize the execution of their tasks (the driver must follow the route determined by the algorithm used by Uber) and sometimes a restriction of their access to data or customers (an Uber driver does not know what leads the algorithm to assign them this customer rather than another).

Evaluation or control, by means of algorithms, is based on permanent surveillance of the worker during the execution of his tasks and a notation, sometimes in real time, of his performance. This continuous algorithmic recording can be fed by both internal and external data. This control by means of the algorithm can, in fact, replace in whole or in part the employer's power of control.

Sanctioning, through the use of the algorithm, tends to automate the hierarchization of workers according to their performance, in particular through the reduction of access to customers for less well rated workers (up to the breaking of the relationship) or, on the contrary, through the automated granting of benefits for those who follow to the letter the recommendations of the algorithm in the execution of their tasks (more opportunities for remuneration, better financial conditions, etc).

1B. Opportunities and threats for workers

AI at work, like any major technological innovation, presents workers with both opportunities and threats.

First of all, the use of AI by employers is mainly driven by the desire to improve the efficiency of decision-making processes, coordination and learning capacity of their organization. Accenture reports that by 2035, AI could help increase global productivity by 40% compared to 2016.

In the HR field, for example, an IBM study indicates that two-thirds of the CEOs of the world's top ten economies believe that AI will generate value for HR. According to Deloitte, the use of "*people analytics"* tools (AI data analysis applied to workforce management) would be a high priority for 71% of global companies for their organization, with the benefits of gaining valuable business insights and taking care of "human problems".

Beyond the benefit to the company, AI can have several benefits for workers.

It has the potential to make workers' activities more productive and easier, less subject to arbitrariness, and more interesting. A study of 10,000workers in Japan (admittedly the 2nd most robotized country in the world) shows an increase in job satisfaction for those who use algorithms in their work.

Furthermore, the use of algorithms, particularly in production lines, can help reduce the risk of accidents, particularly due to poor handling and fatigue. Thus, the monitoring of production steps or safety instructions reduces the risk of mishandling and work accidents.

The OECD thus cites among the AI-related opportunities for workers:

- The increase in human capacity and the improvement in productivity and creativity, freed up by AI taking over the most repetitive tasks. According to a

2017 study, within 10 years, more than 80% of these human tasks could be replaced by the algorithm. Paradoxically, the deployment of AI in the enterprise could make work more human, focusing it on its highest added value.

- promoting the inclusion of under-represented populations,
- the reduction of economic, social, gender and other inequalities,
- and the protection of natural environments, thereby stimulating inclusive growth, sustainable development and well-being.

More specifically, AI can help ensure that workers are treated with less bias and discrimination and are treated in a more personalised way in relation to their needs and aspirations. In particular, this can allow for a better assessment of a worker's entire career path, beyond the certificates and diplomas obtained, thus providing a holistic picture of the person. Professional reorientation is then improved with the help of an algorithmic support.

Conversely, algorithmic management also opens up **four main types of negative risks** for workers: (1) abusive surveillance, (2) discrimination, (3) loss of autonomy and psycho-social risks, (4) skill obsolescence and loss of employability.

Logically, we will dwell more on the analysis of these risks and how to prevent them than on that of the opportunities, because it is these that call for proactive action as a matter of priority, both by companies and by public authorities.

(1) First, the exceptional development of the capacity for machines to collect and process personal data on the work and lives of employees - in other words, to **monitor workers** - is the backdrop to the development of AI at work. Valério de Stefano cites the growing use of "*connected accessories, such as 'smart badges', that measure the speed at which workers move and perform tasks, track their movements in the workplace and their interactions, and even assess the quality of their conversations*", or "*AI tools that analyse workers' emails and messages to monitor their productivity, identify the most innovative workers, but also deviant behaviours*".

These machine data collection and processing capabilities, he concludes, "*supersede any human monitoring system, past or present.* For example, data on the mental health and emotional state of workers is "*increasingly accessible through the use of voice recognition software installed in microphones, monitoring the brain activity of workers*. Some companies award bonuses to employees who agree to wear connected bracelets that monitor their sleep and even their morale.

As the European Commission's Joint Research Centre notes, algorithms thus function as "*instruments of surveillance that replace direct supervision and create power asymmetries*". The capture of relational, health and emotional data, etc. could lead to the implementation of very pernicious forms of surveillance and influence.

Reducing these asymmetries and especially preventing abusive surveillance practices is a priority for workers.

(2) The second risk, which has already been mentioned, concerns possible **increased discrimination**, even if it is involuntary or even unwanted, but linked to the opacity of the actual decision-making criteria of the algorithms. Amazon, for example, had to abandon a recruitment algorithm programme after it was found to disadvantage female applicants for tech jobs.

In this respect, a risk of discrimination that is not yet very present in the public debate concerns "algorithmic collusion", a form of implicit cooperation between several (large) employers with the objective of lowering pay rates. Thus, the widespread use of algorithms in the remuneration of employees increases or creates the conditions for lowering working conditions.

(3) Algorithmic management will lead to a growing number of workers having their work influenced or even directed by processes that they do not control. The perception of a transfer of decision-making capacity from the worker to the machine feeds fears of **loss of autonomy**, including among skilled workers and managers. As Franck Pasquale points out, managers and recruiters are generally unaware of how and why the decisions proposed by algorithms are designed, especially when companies use commercial software and programs.

This loss of autonomy can even lead to a **feeling of alienation**, linked to the perception of becoming an appendage of the machine. In e-commerce logistics chains, order pickers can spend their days with a headset on their head telling them what to do. In retail, the Percolata algorithm allows salespeople to be reassigned directly and without managerial intervention, depending on the flow of customers and the performance of each salesperson. We are thus witnessing a **disintermediation of the manager**.

It is therefore important that AI enhances human capabilities without replacing them. As the European Economic and Social Committee has written, **humans must remain in the driving seat**, as "it is *not ethically acceptable for a human being to be constrained by AI or to be considered as an executor of the machine which would dictate the tasks to be performed*".

Faced with these changes, workers are not inactive, and in fact various avoidance or resistance strategies have been observed: bypassing the use of machines, informal organization charts, alternative procedures, setting up an alternative sociality, etc. Observing, for example, the way in which judges and prosecutors used predictive algorithms introduced in various American courts, Angèle Christin explains that "*what*

struck me was that they "especially don't use them". There is a certain resistance from the judiciary to use these predictive algorithms."

(4) Finally, as with any technological change in the world of work, algorithmic management will change the scope and skills required for a number of jobs. Clearly, these transformations, which are already the source of salary disparities, pose a **risk of skill obsolescence** for a number of workers, internally, and of **reduced employability**, externally, which in turn call for specific measures on the part of companies, and even public authorities, to train them in the various ways of working with AI.

According to the US Bureau of Labor, the most affected workers are expected to be blue collar workers (assembly line workers, machinists, forklift operators, transportation workers, etc.).

But as the scope of AI affects the entire company, all professions are affected by the obsolescence of skills, from the trader whose buy and sell orders are carried out by an algorithm, to the account manager whose banking offers are presented by a chatbot, to the lawyer whoserecommendations are based on an algorithm that analyzes case law for him.

According to the French HRDs surveyed in 2020 by the Cegos Group, 47% of jobs are at risk of skillsobsolescence in the next three years.

1C. Issues for businesses and workers

More broadly, **algorithmic management presents five types of challenges for workers and service providers: explicability**, **objectivity**, **responsibility**, **adaptation of workers' rights** and finally, more broadly, **sharing of value**.

1. First, how can we guarantee **the explicability of algorithms**, whose operating rules are currently opaque to workers and even to regulators?

This question is far from theoretical. For example, it directly involves the **regulators' ability to certify** that the algorithms comply with the normative framework in force. The tragic consequences of the US regulator's validation of the MCAS software used to assist the piloting of Boeing 737 Max aircraft, without actually understanding it, due to the lack of the required expertise - which only Boeing had - shows the extent and topicality of these issues.

2. Then come the **questions about the objectivity of the algorithms** and the **risks of bias and discrimination**, which are moreover those that come to mind most often when it comes to AI at work. As an example, the Bologna court has just judged as discriminatory the Deliveroo algorithm, which used a "reliability" index to classify the delivery drivers and thus prioritize the allocation of races.

Deliveroo's algorithm did not distinguish according to the causes of shift cancellations and therefore potentially discriminated against workers who were sick or on strike. This decision is a reminder that an algorithm should not be expected not to discriminate - that is its very purpose. The requirement that algorithms be objective must therefore be put into context. However, the algorithm must discriminate *fairly*, i.e. with knowledge of all the parameters and all the possible causes of a phenomenon. The decision of the Bologna court thus reminds us not only that an algorithm can be discriminatory but also that its learning set must be exhaustive and itself unbiased.

The risk is not specific to the employee: a poorly specified recruitment algorithm can pass the company by potential candidates. Thus, a recent study shows the extent to which a company relying solely on algorithmic recruitment can find itself at an increasing disadvantage in a hyper-competitive and talent-starved job market.

This in turn raises the question of the ethical standards that the company-supplier-worker ecosystem must set up: how can we guarantee the objectivity of the engineers-suppliers who created the algorithm? How can we avoid that the data feeding the "machine-learning" are a source of bias? What ethical guidelines have been applied within the supplier?

3. **Accountability** issues are equally important.

Algorithmic processing conceals in a "black box" a set of decisions and instructions taken upstream. Indeed, as Antonio Casilli points out, algorithms are not neutral, rational, objective calculation processes, but instructions resulting from a programming made with precise objectives, based on a certain number of principles. Algorithms blur the identity of the principal, especially when they replace injunction with suggestion ("*nudge")*.

These forms of influence pose real challenges to the law: they do not fall within the classical exercise of authority, contractualized. Taking up what Michel Foucault called governmentality (a neologism associating government and rationality), Antoinette Rouvroy uses the term algorithmic governmentality to designate this new normativity, which surrounds each individual with an individualized bubble of influences and incentives (control is exercised, but gently, by orienting behaviour through "nudges"). Especially in case of litigation, it becomes more convenient to hide behind an algorithmic decision rather than to develop an alternative argument.

In the end, who is responsible for the operation of the algorithms and their results? The coder? The manager? The worker who acts according to what the algorithm tells him (but who keeps the possibility of not following it)? Whatever its autonomy, the algorithm cannot be responsible in place of man. What is left of human responsibility, and how can it be attributed?

More specifically, **four questions seem essential to us**:

- How to allow companies to keep control of the algorithms they use, in particular to avoid unintentional biases introduced by the algorithms (as was in fact the case in the [Amazon example already cited](#) where the algorithm for recruiting assistance disadvantaged women).
- How can we manage and support the redefinition of decision-making processes within the company, particularly in the face of the risk of disintermediation and disempowerment of managers?
- More broadly, how can we guarantee, in a context of continuous improvement in performance and no doubt in the place of algorithms, that the human being remains in charge and that each worker preserves a margin of autonomy in relation to the machines?
- What are the responsibilities of employers towards their employees in terms of algorithmic management (e.g. to support them in the face of new uses, maintain their employability)?

More broadly, **won't algorithms help redefine the very contours of business?** Already, digital platforms, like Uber, rely on algorithms to create "[market-company hybrid models](#)". As noted by O. Chagny and Florian Forestier, by using algorithms to replace the usual coordination mechanisms within firms, and thus reducing the raison d'être of firms as a [place to reduce transaction costs](#), platforms are substituting a commercial relationship for an employment relationship, a service contract for an employment contract.

4. The fourth type of issue concerns **the adaptation of workers' rights (and the maintenance of their employability)**.

Historically, labour law was built in part to rebalance workers' rights in the face of power asymmetries with their employers. However, today it makes little provision for the regulation of algorithms, which nonetheless introduce new asymmetries with workers, particularly in terms of information and decision-making capacity.

For example, since digital surveillance tools and algorithms provide employers with almost total transparency on workers and their practices, how can the collection and use

of this data be regulated to avoid abusive surveillance (knowing that it is not only personal data)?

How can we ensure that workers are given the same level of guarantees in the face of the control of their actions by algorithms as they have today in terms of managerial control by employers (of which they are informed of the procedures and which is governed by the fundamental rights and individual freedoms of employees)?

What concrete measures should be taken to avoid the other risks mentioned for workers: discrimination, psychosocial risks, risks of loss of autonomy by workers, obsolescence of skills and loss of employability?

How to rebalance the relationship between the worker and the employer in the face of algorithmic management? How can we guarantee effective access to these rights (the RGPD already offers them, but they are very complex to implement, which makes them *de facto* partly ineffective)?

5. *Last but not least*, the **sharing of value** is a key issue. The fact that machines are now capable of learning could lead to a capture of human expertise that is not only accelerated but also unbalanced in the remuneration of the value created. AIs that model an expertise (medical, legal, etc.) do so from data produced by real work. How much of the economic value created is redistributed to the workers involved? For example, will VTC platforms that try to evolve towards models with autonomous vehicles pay back a share of their profits to the drivers who will have fed the AIs behind these autonomous vehicles?

These questions should lead to a better evaluation of these value creations, and to think about renewed mechanisms of redistribution (towards individuals or ubiquitous professions for example).

II. Guiding principles for promoting responsible algorithmic management

While labour law is still very much focused on the organisation of pre-digital labour relations, taking the measure of the challenges posed by algorithmic management on work and acting now to frame it is absolutely critical.

As we have seen, algorithmic management is a source of both opportunities and threats for workers. It has the potential to make workers' activities more productive, less subject to arbitrariness, and more interesting. On the other hand, it also opens up risks of abusive surveillance, discrimination, loss of autonomy and even increased psycho-social risks, and loss of employability.

However, as Jeremias Adams-Prassl points out, we must not fall into the trap of technological determinism: none of these trends is inexorable and history has shown how much regulation by public authorities and business practices can guide and frame technological developments.

As indeed the European Parliament writes in 2019 in its resolution on a comprehensive European industrial policy on artificial intelligence and robotics: '*developments in the field of AI can and should be designed in a way that preserves the dignity, autonomy and self-determination of individuals*'.

The challenge is none other than to **regain control of AI at work: to promote responsible algorithmic management** to exploit its potential while ensuring working conditions that preserve the dignity, health, equal treatment and autonomy of workers.

However, if we want to exploit its potential while guaranteeing working conditions that preserve the dignity, health, equal treatment and autonomy of workers, it is essential to define what responsible algorithmic management would consist of. To this end, we propose several **guiding principles**, inspired in particular by those of the ILO, the OECD and the TUC.

We have chosen to propose guiding principles for the three main types of actors concerned by algorithmic management: companies, workers and public authorities, starting with companies because they are the ones who implement it and with, for them, a reflection on both their rights and their duties in this area.

2A. 10 business guidelines

The use of AI by companies to organize human work is, first and foremost, a matter of entrepreneurial freedom: it is a way for them to improve their efficiency and better fulfill their missions.

Before discussing their responsibilities in this area, and the constraints that the public authorities could impose on them, it is therefore important to recall their **rights**, and the general principle that any constraints on their freedom to undertake activities must be strictly necessary and proportionate to the aim they pursue (such as the protection of the fundamental rights and freedoms of individuals).

It seems to us that companies are entitled to a regulatory framework for algorithmic management:

1. Respects **intellectual property, for** example in terms of "transparency" of algorithms, as these are often at the heart of their intellectual property,
2. Ensures a **level playing field** between players,
3. Preserves the capacity for **innovation**, by imposing only strictly necessary limits, for example based on the fundamental rights and freedoms of individuals
4. guarantees their **legal security**, despite the need for this framework to evolve given the speed of technological developments underlying algorithmic management, in particular by being clear enough to avoid technical and legal uncertainties and by being easily enforceable against providers.

The very specific issues related to algorithmic management place specific **responsibilities on** companies in this area:

5. **guarantee the rights of their workers** *(see below, Guidelines for workers)*, in internal procedures and upstream in the relationship with suppliers,
6. **informing workers** about their practices in terms of the use of AI to organise work, in particular about the nature of the personal data collected, the use made of it, the criteria and parameters for the operation of the algorithms, the modalities of human control over the operation of the algorithms,
7. **train their workers** in algorithmic management practices,
8. **invest in the employability of workers** if it is challenged by AI,
9. **make AI and its uses in the workplace a topic for social dialogue**,
10. **be subject to external audits** to verify compliance with the above points

2B. 10 guidelines for workers

For workers, who are already "protected" by labour law, it is logical that the guiding principles relating to algorithmic management should be linked to the prevention of the specific risks posed by it (abusive surveillance, discrimination, loss of autonomy, loss of employability).

We propose **10 guiding principles for workers**, relating to their right to dignity, their right to equal treatment, their right to health and their right to autonomy.

Preserving the **dignity of** workers

1. **Right to privacy**: algorithms must not give rise to abuses in the surveillance of workers or the collection of data about them. Since their implementation is not apparent to the worker in his or her daily tasks, algorithms make it easier, whether intentionally or not, to infringe on workers' personal freedoms and contribute to blurring the already blurred boundary between professional and private life.
2. **Right to information and consent**: workers must be informed of the nature of the data and information collected about them and the uses to which it is put, and the data collected must be done with their informed consent,
3. **Right to human interaction**: workers should not have to work only with algorithms but should be able to have human interaction in their work.

Preserving **equal treatment of** workers

4. **Right to non-discrimination**: algorithms at work must not discriminate among workers (in the company but also at the hiring stage)

Preserve the **health and psychological well-being of** workers

5. **Right to health**: algorithms must not be a source of health risks, especially psycho-social risks

Preserving **the autonomy of** workers

6. **Right to understanding (explicability of algorithms)**: workers must not only be informed that they are working with an algorithm, but understand how this interaction works (principle of intelligibility of algorithms),
7. **The right to human decision-making**: workers have the right to have decisions that affect them made by individuals,
8. workers must be able to challenge or appeal any decision made on the basis of an algorithm,
9. **Right to training in the use of AI**: workers have the right to be trained in the use of algorithms,
10. **Right to negotiate**: workers have the right, through their representatives, to be informed and to negotiate the conditions of use of algorithms at work in the company

2C. 4 guiding principles for public authorities

Clearly, algorithmic management raises issues that call for specific intervention by public authorities. The current regulatory framework is not sufficient, it must be adapted to these new challenges.

Public authorities thus have a fourfold responsibility in terms of algorithmic management:

1. **Depth and autonomy of expertise**: public authorities must provide themselves with the means to produce **relevant regulations** on AI at work, balanced between respect for entrepreneurial freedom and protection of workers' rights. This consideration might seem obvious, but the radical novelty, technicality and speed of evolution of the subject require public authorities to make a specific investment in order to have a detailed understanding of the realities and specificities of algorithmic management and an expertise that is truly independent of the actors.

2. **Effective access to rights**: given the very great complexity of the subject, even its opacity, and the risk that any abuses may be largely invisible to the actors themselves (for example, in the case of abusive surveillance or discrimination), promoting relevant regulation will not be enough; the public authorities must ensure that the **rules laid down are properly respected and the rights guaranteed are effectively accessible**, particularly for individual workers,

3. **Supporting employability** - since the increasing use of AI in the workplace will have considerable effects on the working conditions of workers, and in particular on their required skills, with a risk of "drop-out" or obsolescence of skills on not only an individual but also a collective scale, it is up to the public authorities to **promote awareness-raising and training schemes** to complement the actions carried out by companies, to equip individuals with the skills they need in the field of AI and thus support their employability in the long term,

4. **Adaptability of the regulatory framework** - the nature and uses of AI in the workplace are changing all the time, and very quickly, so its regulatory framework must not be defined or fixed once and for all but must be highly evolving, at the risk of losing all relevance, with more frequent review clauses to adapt the law and the regulatory framework, and greater openness to experimentation

III. 20 concrete proposals to promote responsible algorithmic management

How can we ensure that these guidelines are applied in practice?

We propose to give a real legal scope to these principles by adding them among the recitals to the draft European Regulation on AI, thus giving them a "preamble" status to this Regulation. We also propose to include them in the guidelines already defined by the OECD on AI and adopted by its member states.

In addition, in order to give them an even more solid legal basis, we propose that they be broken down into **four areas of concrete recommendations**, all of which are fully compatible with the promotion of innovation, respect for intellectual property and competition rules, and respect for the legal certainty that economic players are legitimately entitled to expect.

   I.   **PREVENT - Adopt safe design rules for responsible AI at work**

  II.   **PROTECT - Adapting individual workers' rights to the transformations of work brought about by algorithms**

 III.   **GUARANTEE - Ensure effective compliance with the principles of responsible AI and workers' rights**

 IV.   **UPDATE - Allow the regulatory framework to adapt over time to developments in AI**

*

AXIS 1: *PREVENT* - Adopt safe design rules for responsible AI at work

Proposal #1 - Introduce an objective of explicability and ethics "*by design*" of algorithms

We propose to **enshrine a code or algorithm explicability objective in European law, via the draft AI Regulation**.

There are several arguments for this proposal:

1. The explicability of the algorithm is essential for workers and service providers to be able to enter into contracts with full knowledge of the facts. The use of algorithms, as an indirect contractual instrument, creates rights and obligations. However, the operating rules of algorithms are currently opaque to workers and even to employers. It is not uncommon for only the author of the code, an external service provider, to be able to explain it in full. Without transparency or readability, the decision resulting from algorithmic management cannot be motivated for its recipient: it is therefore... arbitrary.

2. Explainability is an essential prerequisite to circumscribe the risks of bias, discrimination, or other infringements of the rights of workers and service providers, to make the content of the algorithm an object of negotiation and to allow the audit of the algorithm by a third party and in particular a judge in the event of a dispute;

3. There is already a major legal precedent, the Constitutional Council having recognised the constitutional value of the objective of accessibility and intelligibility of the law. This objective of intelligibility now constitutes a new component of legal certainty. Now, if we consider, in a figurative sense, that "*Code is law*", it is legitimate, similarly, to consecrate an objective of intelligibility of the algorithm

4. Conversely, this objective of explicability will not lead to any harm to the promotion of innovation, nor to the intellectual property rights attached to algorithms.

An algorithm must be interpretable, and a result explicable. Only this intelligibility, established as a principle from the design of the algorithm, will allow the parties - workers, employers or principals - to access the reasons for a decision in order to understand it and correct it if necessary, which will preserve the responsibility of each party.

Admittedly, the European Platform to Business (P2B) Regulation already introduces the notion of transparency in the relationship between workers and platforms, but this transparency is limited to platforms and is only considered downstream of the algorithm, in the effects of its implementation. However, we should intervene upstream, by **introducing a requirement of explicability of the algorithm to be implemented from its conception in the European Regulation on AI**.

This objective of explicability must be translated into **criteria of accessibility, auditability, and even the involvement of different stakeholders in the design of algorithms in order to appropriately** manage their development and thus become "*ethical by design*".

The European Regulation on AI could order *at least* **the explicability of the object code** -and not the source code- of the algorithm, as well as the **communication on request** of any worker or service provider concerned by its implementation, of an

intelligible document opposable to its issuer, explaining in a clear and complete way the conditions and the effects of the implementation of the algorithm on the management of which it is the object.

This explicability objective must be extended to the results of tests and audits of the algorithms.

Article 13 of the draft European Regulation states:

> "*The design and development of high-risk AI systems shall be such that the operation of such systems is sufficiently transparent to allow users to interpret the results of the system and to use it appropriately. An adequate type and level of transparency ensures compliance with the relevant user and provider obligations set out in Chapter 3 of this Title.*"

Transparency is therefore only defined by certain purposes (allowing users to interpret the results of the system) and its scope is not circumscribed. Moreover, only the employer or the principal can benefit from this transparency as the text currently stands.

It is therefore possible to clarify the provisions of this article 13 in order to

- introduce the notion of explicability of the algorithm, which is more precise than the notion of transparency, which is envisaged in the draft regulation to different degrees (thus the simple fact of informing people of the implementation of an algorithm is, according to certain provisions, sufficient transparency),
- endorse the fact that this explicability must benefit not only the user of the algorithm but also the person who is the subject of a decision dictated or assisted by the algorithm,
- specify the scope of this explicability objective, which must be applied from the design of the algorithm and downstream, to ensure the readability of audit and test results.

Proposal #2 - Ensure traceability of algorithmic decisions

The implementation of the algorithm in the managerial context impacts the future of a contractual relationship.

Indeed, the algorithm will sometimes intervene as an aid to the managerial decision, and sometimes take the decision alone (*de facto*). Its intervention must be known to the worker, as must any managerial decision impacting the execution of the contract (in terms of work organization, for example).

It is essential to ensure the traceability of algorithmic decisions affecting contract performance for two reasons:

First of all, a worker or service provider whose contract is affected by the implementation of an algorithm must not only be informed, but **must also know** in detail, upon first request, **what decision was taken by or through the algorithm concerning him or her**: nature, meaning, date, time, impact on the performance of his or her contract.

Then, **the worker must be able to question this algorithmic decision** in terms of its motivation, its validity, and to know precisely the place of the algorithmic decision in the decision finally adopted (decision aid for the superior or principal, or autonomous decision of the algorithm).

Only this traceability will make it possible to trace the reality of a bias, of a discrimination and possibly to find its source (purely human bias, code of the algorithm itself or more frequently, weakness in the selection of the training data of the algorithmic model).

Traceability therefore appears, in this contractual framework, as one of the **protections of the worker or service provider against arbitrariness** and an imperative to make the employer or principal who implements the algorithm accountable.

This traceability, which is already *minimally* envisaged by the draft European Regulation in its Article 20 ("Automatically generated logs"), must allow not only for the systematic creation of the log relating to the operations of the algorithm but also - and effectively - for the intelligibility and accessibility of these logs for the workers or service providers who have been subject to algorithmic decisions or who have been assisted by the algorithm.

Finally, since the algorithmic decision is based on the objective pursued by the client, as well as on the algorithmic work of analysis and sometimes learning and comparison, this traceability must not be limited to the decision itself. Such a limit would not allow to keep track of the motivation of the decision - independently of its readability.

Article 20 of the draft European regulation could therefore be completed to specify the activities of the algorithm subject to the automatic generation of a log and to affirm the purpose of the log thus produced: to ensure the traceability of decisions and algorithmic work to preserve the rights of workers whose activity is affected.

The use of blockchains to operationalize this traceability deserves to be studied, even if it poses several important questions, including that of its articulation with the right to be forgotten guaranteed by the RGPD.

Proposal #3 - Make certain sequences mandatory within algorithms to guarantee workers' rights and conversely prohibit sequences with deleterious effects

Algorithmic management affects relationships that are primarily contractual (between principals and service providers or between employers and employees). It should therefore be subject to the same legislative framework and judicial review as the contract whose performance it affects.

This is especially important because a study by consulting firm BCG shows that less than half of the companies that have deployed AI in their organizations have done so with responsible AI.

- Just as there are compulsory clauses in certain types of employment contracts, it is possible to provide for **blocks of mandatory codes**. They would be intended to enshrine certain guarantees for workers: the legislative enshrinement of rights, more complete information, effective and systematic access to data, etc.

- Conversely, the judiciary may exclude the application of certain clauses and deem them unwritten, such as "ineffective clauses". In the same way, **certain blocks of code with deleterious effects should be able to be submitted to the judge in order to have their application excluded from the employment relationship** (if they are not already excluded upstream by the double effect of the normative framework and collective bargaining).

Proposal #4 - Supervise the design of algorithms, in particular through upstream control of learning databases to ensure their diversity and minimize algorithmic bias

Biases in self-learning algorithms can come directly from the database used during the learning phase of the algorithms.

Indeed, the work of data preparation is very often neglected in favour of model generation. However, this phase of construction of the training databases is totally decisive to guarantee the quality of the generated model and thus minimize the algorithmic biases.

Two points of attention are therefore essential:
- The first is part of the **data quality** phase in order to make them usable by the algorithms. A quality control must therefore be necessary with access to the

metadata in order to know the characteristics of the training data (types, origin, statistics, etc.) in order to measure the applicability levels of the trained models.

- The second concerns the **enrichment of databases**. We can talk about (i) simple enrichment with the addition of external public data and (ii) complex enrichment which comes from a more or less automated generation of new information by crossing the initial data. The purpose of this data enrichment phase is to increase the amount of information available for training the algorithms and thus minimize learning biases.

In the same way that an algorithm must be explained on request by a team of data experts (such as a data scientist or data engineer), the metadata of the learning databases as well as the characteristics and conditions of enrichment of these same databases must also be explained on request by a team of data experts.

Proposal #5 - Provide for the pooling of learning data with third party institutions, either governmental or in the form of data trusts, to minimize bias

In order to limit the biases linked to databases, they must be enriched with a sufficient quantity of well-described, calibrated and sufficiently diversified data. In its report "Algorithms: Bias Control", the Institut Montaigne mentions the emblematic issue of facial recognition, for which developers do not have a database that is sufficiently representative of the French population to test the absence of bias in their algorithm.

However, this expansion of data collection will only happen if a framework of trust is established.

The Institut Montaigne proposes (proposal 5 of the report "Algorithms: Bias Control") the creation of one or more public databases containing variables protected under the 25 discrimination criteria. These would be used to test the methodology, before or after training on the own data. According to the Institut Montaigne, the Laboratoire National de Métrologie et d'Essais (LNE) could assume this responsibility. This laboratory is indeed *"already* involved in *AI application areas such as autonomous vehicles, smart medical devices (especially diagnostic aids), collaborative industrial robots or drones. It is responsible for a die designed to evaluate agricultural robots, and has already formed partnerships to develop methods for evaluating automatic speech transcription algorithms."*

An alternative solution that we favour because of its greater capacity to ensure a framework of trust would be to develop "data trusts" dedicated to algorithmic

management tools. These "data trusts" are independent intermediaries with a fiduciary duty to act on behalf of the data subjects whose individual data they collect and manage. They have a twofold interest. 1) to simplify and make more transparent the collection and use of users' personal data, 2) to allow actors to benefit from diversified data pools, while guaranteeing that they are used in a way that complies with the wishes of the stakeholders, and 3) to complete, if necessary, via synthetic data approaches, possible under-representations of a part of the sample that will allow the protection of personal data to be preserved

The Open Data Institute (ODI), a research institute working on the issue of the data trust, proposes six pillars around which to organize such an organization: a precise goal, a decision-making process defined upstream of its creation, a profit-sharing model, a real legal structure, user rights to the data they share, and sustainable funding. A data trust can take different forms, from a company to a non-profit association, including an administrative authority, as long as its structure is defined beforehand.

Data trusts are currently used in the context of sectoral policies (in the field of health, the best known example is the Health Data Hub) or territorial policies, in particular for "smart cities" or "intelligent territories" projects, according to Luc Bélot's formula, for example to coordinate new intelligent multimodal mobilities.

We propose to create such structures for the data used for algorithmic management solutions. Provided they reach a critical size, these would allow to :
- reduce the risk of bias and make the algorithms more sensitive to the facts, thanks to the multiplication and diversification of the learning modes of the algorithms;
- propose a model that can act as a filter for content generated by the linguistic models used to train algorithms on a large scale at a lower cost for the operators who will use them, for example for SMEs, while offering, thanks to the intervention of the third party, guarantees as to their quality and the prevention of the biases that they could incorporate.

Proposal #6 - Place certain types of data under the status of "data of general interest

The ability to access a large amount of data (sourcing, CV sorting, interviewing and assessment) is also a significant barrier to entry for solution design, reinforcing the position of a few companies and putting them in a position of strength, allowing them to impose standards without client companies and other stakeholders having sufficient decision-making power or control.

This situation is obviously not limited to management issues and concerns many areas, such as energy, water and mobility. In these areas, however, many milestones have already been set in order to build a framework to secure the access of public actors to these data. Thus, the 2016 Law for a Digital Republic made it mandatory to make available data "whose publication is of economic, social, health or environmental interest". As such, it obliges the delegates of public services that collect water, energy or waste management data to open access to these data.

A broadening of the possibilities for collecting "private data of general interest" is also being discussed, giving public actors the means to apply it, for example by making the delivery of activity permits conditional on the obligation to transmit certain data. For now, the focus is on data from platforms such as waze, Uber or AirBnB, which are essential for the digital transition of territories. In 2017, the Belot report proposed a status of "data of territorial interest" which would allow to engage at the local level "a census of available data of territorial interest", produced by companies or associations.

We propose to include certain data concerning managerial practices in this reflection.

First, by identifying precisely the types of data concerned.

Then, by proposing modalities to facilitate their availability and sharing.
- Either by obliging private sector actors to practice *open data* concerning these data (like what has been planned for the availability of local energy data by the law of energy transition).
- Either by allowing a third party to use the data while leaving their processing to private actors[1].
- Or, finally, in a more ambitious way, by extending to the end the logic of "data of general interest" by making an inventory of all the data falling under this issue and by giving them the status of "commons". The status of data as a "commons" is defined on the basis of a bundle of rights that breaks down ("dismember" in the legal sense of the term) the ownership of data into a set of distinct rights.

---

[1] "Engaging in a private/public rebalancing through data," *Innovation & Foresight Papers* No. 05, Commission nationale de l'informatique et des libertés, 2017.

Proposal #7 - Implement a specific Model Life Cycle framework to minimize bias and the risk of model quality drift over time

In the same way that data governance methods aim to ensure that the quality of data is maintained during its ingestion and redistribution, the model lifecycle ensures that the quality of models remains stable over time. Indeed, a model in production is "connected" to data in perpetual evolution which are emitted for example daily.

Thus the predictive model is effectively generated from past data (controlled, identified, validated ...). However, it feeds on recent data to propagate information towards the future. Thus the assumption that data are independent and identically distributed during the learning phase becomes erroneous in the real world. For example, let's think about a fraud detection system where the adaptation of fraudsters makes the model built obsolete. We talk about model drift when in reality it is the data itself that has changed. It becomes essential to build and control the model life cycle.

The model life cycle therefore encompasses the events that occur during the existence of a model, i.e. once it has been put into service.

The first phase lies in the change management defined when the model is inserted into the existing process by defining, for example, access and usage rights (etc.). Secondly, the data life cycle ensures that the data served to the model obeys the standards defined during its learning phase.

It is only during the third stage - that of model monitoring - that the quality of the model is monitored and controlled in terms of correct operation, performance, feedback from use, etc.

In a project involving AI, we can distinguish three levels of intervention on the life cycle of the models:

- <u>Business organization</u> - ensures the business impact of a model by ensuring the applicability of the model over time.
- <u>Scientific</u> - to evaluate the robustness of the models to the effectiveness of a drift detection system and to the different solutions of performance adjustment: re-training of the models.
- <u>Technological</u> - managing the system architecture in production, interfacing the data life cycle with the model and ensuring the rapid deployment of solutions recommended by the scientific level.

We therefore recommend setting up a team of experts responsible for the model life cycle to ensure that the model does not drift over time:
- Data Quality Manager - Ensuring the sustainability of the data repository

- responsible for the robustness of the model - Check not only the performance of the model but also its robustness by monitoring its uncertainty: Model monitoring
- Production Manager - Ensure the orchestration and implementation of the processing chain on the production platform
- Monitoring & Alerting Manager - Monitor performance indicators and manage alerts in case of problems.

Proposal #8 - Supervise the production of algorithms with different stakeholders

In the case of algorithms, the law cannot be satisfied with framing technical possibilities *a posteriori* and seeking to rebalance after the fact an asymmetry of power inscribed in the very design of the algorithms: it must construct a framework to guide their development, limiting from the outset the establishment of such asymmetries and the way in which they are reinforced.

As the CNnum points out in its [report on work in the age of platforms](), one of the most controversial tools of algorithmic management is abusive design. This is the result of choices made by the application's architects and designers, who mobilise the user's cognitive biases in order to influence his or her behaviour.

The notion of deceptive design refers to practices aimed at representing elements in such a way that they could mislead the individual. Dangerous designs are methods that can make the worker directly or indirectly vulnerable. These designs may encourage the worker to commit offences. For example, timers were automatically triggered on some bike delivery apps when a ride was accepted. This could induce couriers to commit traffic violations by playing on cognitive biases.

It is therefore necessary to train designers in ethical issues and, more specifically, in labour law, for example by creating cross-training courses within law and design courses.

Deceptive designs could also be condemned on the basis of unfair commercial practices. For example, on March 15, 2021, the California [Consumer Privacy Act](https://...) (CCPA) was updated to prohibit companies from using certain "dark patterns" (techniques used by platforms to dissuade an Internet user from opting out of a service to which he or she has subscribed or to get him or her to involuntarily share his or her private data).

The term "*dark pattern", which* first appeared in 1994 in *Wired* magazine, refers to interfaces (considered unethical) designed to obtain the desired action from the user. As

explained by the CNIL in its report on "[The shape of](#)choices [2]the question of indication by design is however vast: beyond dark patterns strictly speaking, many designs are conceived to numb the user's vigilance, to elicit automatic behaviors, and, as Dominique Boullier points out, "to *reduce to the extreme conscious hesitations and arbitrations, to create a form of naturalness (...) that will seem very economical on the cognitive level.* " (p.13)

This dimension would require a broader regulatory framework.

___

---

[2] Woodrow Harztog distinguishes between abusive design, deceptive design and dangerous design.

AXIS 2: PROTECT - Adapting workers' rights to the transformations of work brought about by algorithms

Proposal #9 - Introduce a true right to human decision

The GDPR has created rights for workers in algorithmic management[3]:
- the right of workers not to be subject to decisions based solely on automatic processing (Article 22),
- the supervision of profiling, i.e. the automatic processing of data to evaluate individuals or predict their behaviour (Article 4).

The Digital Service Act, for its part, should by 2022 govern the responsibility of intermediaries -and more specifically platforms-, subject to increased transparency (especially algorithmic) and accountability obligations. The European Platform to Business (P2B) Regulation, which came into force on 12 July 2020, already lays the foundations for these transparency, fairness and accountability requirements.

But **these rights enshrined in the RGPD and the future Digital Service Act are above all formal. All the** employer or the principal has to do is to provide an equally formal response to the rights holder's request, in order to render them **ineffective**: the employer can confirm that it has checked the algorithm's decision, the platform can assert that its code is intelligible to a competent data scientist or that its complexity is incompressible.

The CNIL recommended that the right not to be subject to decisions based exclusively on automated processing (Article 22 RGPD) should be considered more broadly as a **right to supervision** of the algorithmic decision by the employer or the principal, over more or less numerous series of decisions. This would make it possible to introduce "*forms of human and contradictory deliberation framing and accompanying the use of algorithms by examining and questioning the parameterization but also all the effects - direct and indirect - of the system*".

The EU AI Regulation should introduce a **positive right to a human decision** (with the possible implementation of an algorithmic process) vs. currently the "hollow" right not to be subject to an algorithmic decision:

---

[3] The Digital Service Act should also govern the responsibility of intermediaries - and more specifically platforms - by 2022, subject to increased transparency (especially algorithmic) and accountability obligations. The European Platform to Business (P2B) Regulation already laid the foundations for these transparency, fairness and accountability requirements in July 2020.

- This human control of the algorithmic decision must be defined by the public authorities, according to criteria that make it possible to make the employer or the principal responsible, while making this right effective for the worker.
- This right to a human decision should be put **forward by the principal or the platform**, both at the time of the conclusion of the contract and during its execution, with the establishment of **internal procedures facilitating its implementation**.

Article 14 of the draft Regulation intends to impose a design and a mode of operation of the algorithm that allows its user to retain control over the work and decisions of the algorithm. However, it does not create any rights for the worker or service provider affected by the decision dictated or assisted by the algorithm.

As explained, this right can take the form of an evolution of Article 22 of the GDPR. However, it can also find its place in Article 14 of the draft European regulation, in that the algorithm also uses data that are not personal and by nature fall outside the scope of the RGPD to reach its decisions.

It is therefore possible to complete this article 14 to introduce, following the means made available to the user of the algorithm to maintain control over the AI, a real right for the worker to solicit the user and impose the effective use of the means that the latter has under article 14.

This right to human decision making could also be included in the "Charter of Digital Rights" that the European Commission will soon propose.

Proposal #10 - Introduce a right of communication of an intelligible and opposable document on algorithms at work

The communication of an intelligible document on the algorithm and opposable to its issuer, should be a right for each worker and for the institutions representing the personnel.

This document should **clearly and comprehensively explain the effects of the implementation of the algorithm** on the management of the workers.

The emergence of a formal obligation to provide information - contractual or even pre-contractual - on the part of the employer or principal is comparable to the duty to provide pre-contractual information already in force in consumer law or in ordinary contract law.

The algorithm must be explained on request, by a team of data experts (such as a data scientist or data engineer, whose services may be requested by employee representative institutions, as they are already for other experts).

Proposal #11 - Ensure a better sharing of the value created by the use of AI in the workplace between the company and the workers who produce the data

A misleading perception of the creation of value linked to algorithmic management could lead to the idea that it is the AI alone that creates the value and that it is therefore legitimate that this value be returned, in full, to the company.

In reality, no matter how relevant its contributions, AI only intervenes at the end of the value creation chain, which could not exist in any case without the data itself on which it relies and therefore the workers who are at its origin.

Even if this data production does not result from a specific additional work of the worker compared to his normal activity within the company, he plays an essential role in the creation of value related to algorithmic management, which justifies that he receives a part of it.

This is all the more true when the data captured by the machine is part of the worker's personal expertise and the machine learns from this expertise and models it (e.g. in the medical or legal field), without this "transfer of knowledge and skills" being provided for as such in the employment contract and therefore in the worker's remuneration.

For example, will VTC platforms that are trying to move towards models with autonomous vehicles give a share of their profits to the drivers who will have powered the AIs behind these autonomous vehicles?

Of course, one could argue that workers are already likely to benefit from this value through potentially easier, more interesting and more productive working conditions. The fact remains that, *in the end,* it is the company that is capturing the extra financial value creation, even though it is partly due to the workers themselves.

Moreover, the data thus created have an economic value that goes beyond their original purpose. When this data has a second life, notably through its aggregation and commercialization for purposes other than those of the initial business activity, it is legitimate that the value received by the company be directly shared with those who contributed to its creation.

In the case of the capture of purely contingent data in the context of the worker's activity, as in the case of the worker's expertise, the company manages to create value indirectly from the worker's activity without this creation being a contractual task. Since it is not linked to the employment contract, it is not part of the economic equilibrium of the contract, to the detriment of the worker's interests.

It is fully legitimate for the worker to receive a share of the value that his activity has contributed to create, even if this was not the direct object of his work, as long as the company itself benefits from it.

That is why we propose to introduce into labour law a provision for value-sharing mechanisms for the benefit of workers in the event of the capture and exploitation of data produced by the worker that give rise to the creation of value for the company without this constituting a contractual task.

Of course, the simplest way to deal with this issue would be to include in the terms of the employment contract value-sharing arrangements that cover all the value created by the company from the worker's activity.

Of course, it is not up to the Labour Code to regulate this question in detail, but at least to affirm the principle, so that the social partners can, through social dialogue, find solutions adapted to the different situations.

However, this social dialogue risks coming up against the technical nature of the subject and the difficulty of accurately assessing the value created and tracing the methods of its creation, with the risk that it will not be very effective.

However, given the magnitude of the issues at stake and the exponential growth in the use of data produced incidentally by workers, we feel it is essential to set a first milestone and launch a social dialogue process.

AXIS 3: GUARANTEE - Ensure effective compliance with the principles of responsible AI and workers' rights

Proposal #12 - Move from vendor self-assessment of regulatory compliance of algorithms to assessment by third party organizations

The Commission's draft Regulation on AI does not depart from the EU's "new legislative framework", which aims to establish the European legal basis for the conformity of all

non-food products. It therefore takes up the concepts of *ex ante* and *ex post* control, here applied to the algorithm, just as it enshrines in advance the intervention of the standards harmonisation bodies at European level, namely the European Committee for Standardisation (CEN) and the European Committee for Standardisation in Electronics and Electrotechnics (CENELEC).

Indeed, beyond the principles that this draft Regulation intends to establish, CEN and CENELEC are called upon to draw up harmonised standards that algorithm suppliers will be able to follow, failing which they will attempt to interpret the future Regulation themselves.

However, compliance by a product - here by an algorithm - with these "EC standards", i.e. the standards issued by these committees, is equivalent to a presumption of conformity.

Continuing with its "new legislative framework", the EU distinguishes between AI systems intended for use as safety components of otherwise regulated products, such as toys or medical devices, and high-risk autonomous AI systems as listed in Annex III of the draft Regulation.

The former will have to be subject to the *ex-ante* and *ex-post* control procedures already governing the products of which they are components, while the latter will have to be subject to a new compliance system, which in fact relies almost exclusively on internal ex-ante control by the supplier of the artificial intelligence system.

However, as [Valério De Stefano](#) believes, this self-assessment by the supplier is insufficient to provide protections equivalent to those put in place for other high-risk systems "*which must be subject to stricter conformity assessment procedures, with the involvement of a notified body*". Especially since this provision seems not to have been subject to any form of social dialogue at EU level.

It is furthermore notable that several national laws within the EU go further than the regulation and "prohibit or significantly limit the use of technological tools to monitor workers (Aloisi and Gramano, 2019)."

Private players and national public decision-makers therefore seem ready to raise the level of requirements and transparency applicable to the compliance of artificial intelligence systems, by acknowledging the reality of the risks and the need for specific third-party certifiers.

A fortiori, we propose that this conformity assessment of high-risk algorithms be established, *ex post* and/or *ex ante*, under the responsibility of third-party bodies such as notified bodies - already known actors of conformity under the EU "new legislative

framework", and not only through the supplier's self-assessment, which could have the effect of rendering ineffective the conformity sought, together with the presumption of conformity attached to compliance with the standards enacted by CEN and CENELEC

Proposal #13 - Implement procedures for estimating the risks of algorithms, based on the model of the banking stress test

More and more companies and institutions are confronting the problem of *"algorithmic risk"* and resorting to new experts, specialists in algorithmic auditing, capable of auditing and interpreting the results of a model and assessing its regulatory compliance.

But what does this expression "algorithmic risk" cover? We need to take a broad view here and not limit ourselves to risks related to the reliability of systems. As is already the case in certain industrial sectors (e.g. aerospace, healthcare, and especially environmental impact studies), the question of the social harm that algorithms may cause must be included in the risks.

Companies are de facto already sensitive to this dimension because of the reputational risks to which they may be exposed by certain undesirable effects of algorithms: for example, there has been a lot of talk since 2018 about the problem of the recruitment tooldeveloped by Amazon, which eliminated women. For all that, there is still no systematic framework for addressing the issue of algorithmic risks, especially in management.

The evaluation of algorithms involves finding a compromise between three requirements: the level of precision and reproducibility of a decision resulting from an AIS by learning, the explicability of a decision according to the level (designer, user, user) and the domain concerned, and the risks according to the domain envisaged. For this, it is also necessary to clearly distinguish the risks related to the data from the risks related to the algorithms. If the use of deterministic symbolic AI (expert analysis of the decision rules encoded in the algorithm) indeed allows an independent evaluation of the algorithms, machine learning and even more so deep learning generate de facto a black box effect intimately linked to the data used to exercise the AI solution.

The next step is to set up control and audit protocols.

As far as discrimination is concerned, we can refer to some existing protocols, for example the guide established by Philippe Besse aiming at better detecting and evaluating the risks ofdiscriminatory impacts of AI algorithms. This one draws up a protocol passing by a set of questions allowing to accompany an organization, concerning the design of the solutions, the processes to test and control the possible biases during

the phase of development, deployment and use of the system, the analysis or the indicators used.

The evaluation can then take several forms.

- Internal, which implies training employees, providing tools (i.e. mainly algorithmic audit libraries, i.e. algorithms capable of analysing other algorithms). Internal control can address the most frequent and best identified biases: **data bias** (a recruitment algorithm trained on a database in which men are over-represented will exclude women), **stereotype bias** (a woman tends to click on job offers she thinks are easier to get as a woman, **omitted variable bias, selection bias** (a bank will use internal data to determine a credit score, focusing on people who have or have not obtained a loan, but ignoring those who have never needed to borrow, etc).

- An external control, which can be exercised by third-party structures, as recommended by the Institut Montaigne. The creation of a twenty-strong cluster of expertise on algorithms at the Directorate General for Business (Ministry of the Economy) capable of auditing algorithms, initiated in 2019, sets milestones in this direction, even if it would be desirable to broaden its scope and involve experts and stakeholders from various backgrounds (worker collectives, sociologists, lawyers, etc.) to identify sensitive management issues. But it would also be necessary to provide for the intervention of other actors, similar to auditing companies, and to consider algorithmic certification procedures, as is already the case for accounting issues. Another possibility is that of accountability, leaving organizations free to choose the form of control they choose to implement, but strongly sanctioning the realization of the risk after the fact.

- A sectoral evaluation to prevent any implicit or explicit collusion between algorithms. This can be done through benchmarking or inter-sectoral comparison procedures, for example, in order to detect a possible convergence of working conditions and remuneration towards the lowest common denominator. This cannot be done without collaboration between different regulators, in particular the competition authorities, which should be made aware of the impact that such collusion could have on the labour market.

These control bodies, both internal and external, must also be able to benefit from adapted and efficient tools, which implies supporting the research currently underway in the field of algorithmic evaluation. One example is the "Matilda" algorithm stress-testing project at the University of Melbourne, which provides public tools for a more refined and multifactorial evaluation of algorithmic performance, so as to adapt the test procedure to the criteria and purposes of a solution. Indeed, underlines Professor Kate

Smith-Miles of the University of Melbourne, who is in charge of the project, algorithms are currently tested according to very standard procedures, on pre-established samples, well marked out by the scientific literature, poorly adapted to the contexts of use, which limits in particular the possibility of anticipating and evaluating the human effects of their implementation.

Proposal #14 - Integrate the measurement of risks related to algorithmic management into the risks within companies, particularly psychosocial risks

Studies show that the deployment of AI in the enterprise is increasing the pay gap between those who are proficient and those who are not proficient in the use of algorithms.

Beyond that, the use of AI also impacts the way workers understand well-being, stress and health at work.

By the depth of its professional impacts and the diversity of tasks involved for the worker, AI multiplies psychosocial risks at three levels:

1. **hyper-standardization of tasks for workers**, with AI leaving the not yet automated tasks to humans;
2. a **hyper-complexification of tasks for workers**, with AI supporting the simplest tasks. A Japanese study conducted in 2017 with 10,000 workers showed that while the use of AI contributed to job satisfaction by focusing some of the workers on less repetitive tasks, AI introduces in these cases an additional level of stress related to the complexity of the tasks remaining to be performed by the human.
3. Finally, **the relationship between humans and collaborative robots or cobots** leads to three risks[4] in terms of Occupational Health and Safety (OHS): i) risks of robot-human collision, due to unpredictable robot behaviour following automatic learning, ii) security risks, when the Internet connection of robots affects the integrity of the software programming, leading to the appearance of vulnerabilities, iii) environmental risks, in the event of degradation of sensors or unforeseen human behaviour in unstructured environments

Note that these "cobots" are already a well-established reality: the Dutch organization for applied scientific research wrote back in 2018 that "Amazon thus has 100,000 AI augmented cobots, which have reduced employee training time to less than two days.

---

[4] https://osha.europa.eu/en/publications/osh-and-future-work-benefits-and-risks-artificial-intelligence-tools-workplaces/view

Airbus and Nissan, meanwhile, are using cobots to speed up production and improve efficiency."

The prevention of psychosocial risks is all the more important as it cannot be based exclusively on generic models and requires, on the contrary, a proactive audit and analysis approach specific to the company in order to be as effective as possible.

This is especially true when implementing an algorithm in the employment relationship.

We recommend integrating the measurement of risks related to algorithmic management into the framework of psychosocial risk prevention by meeting the following prerequisites:

- the involvement of the company's management,
- Massive training of all employees in the uses, opportunities and risks of AI. A study conducted by IBM indicates that within the world's top 10 economies, some 120 million workers will need to be re-trained and their skills retrained to support AI and intelligent automation,
- the development of a culture of prevention of psychosocial risks linked to the implementation of an algorithm in work relations,
- in-depth reflection on how best to combine the skills of humans and robots in the specific context of the company concerned[5],
- the involvement of staff representatives and their training accordingly,
- the possibility for staff representatives to be assisted by third parties in identifying and assessing these specific risk factors,
- if necessary, support and training for occupational health services on the specific problems of implementing algorithms within the company,
- informing employees about preventive measures and giving them access to the tools for this prevention,
- the possibility for employees and staff representatives to initiate alerts on psychosocial risks linked to the implementation of algorithms, without any consequence on their job or their mandate,
- the development of a specific section on the implementation of algorithms within the company, in the Single Risk Assessment Document,
- the development of action plans after identification of the risks and the budgeting of their implementation, followed at regular but realistic intervals by a new audit to take into account the evolving nature of algorithms, their uses and therefore their impact in terms of PSR.

---

[5] For example, the German Federal Institute for Occupational Safety and Health (BAuA) holds annual workshops on the subject of "human-robot collaboration". https://osha.europa.eu/en/publications/future-work-robotics/view

Proposal #15 - Extend the responsibilities of the DPO to the field of algorithmic management

The RGPD already requires some companies and institutions to create a position of DPO, for Data Protection Officer (Délégué à la Protection des Données).

It would be possible to broaden the responsibilities of the latter, or to impose the creation of the **position of Algorithmic Management Officer** on organizations that make massive use of algorithmic management processes (delegate for the protection of responsible algorithmic management).

For companies, appointing a manager and integrating him or her into an organisation chart is often the first way to initiate a policy, processes, standards and frameworks that will enable it to be deployed. His/her mission would be to inform and advise his/her organization on the one hand, and to control the development of applications and interfaces for a responsible use of algorithmic management tools on the other. This person could be an employee of the company or an external service provider and would report to the highest level of the organization. He or she would lead a network of relays within the organization's subsidiaries and/or a team of in-house experts (design, legal, IT, human resources) and, like the DPO, would be the natural contact for the authorities as well as for employee representatives for any problem related to algorithmic management.

Proposal #16 - Integrate individual and collective rights relating to algorithms into corporate codes of conduct

These codes of conduct are annexes to the internal regulations. They are therefore subject to the procedures for informing and consulting employee representative bodies.

We propose to include in the regulatory part of the labour code, the obligation for companies to mention individual and collective rights related to the implementation of algorithms by the employer, in the codes of conduct and IT charters.

We also propose to integrate in the CSR declaration obligations of companies a "responsible AI" section describing the transparency practices in terms of algorithmic management and the measures taken by the company to maintain it over time.

Among these good practices, the mobilization of buyers is essential to ensure a code of transparency with suppliers and employees. The training and support of these buyers allow the deployment of responsible AI and avoid the "black box" effects of algorithms purchased and used within the company without knowing their real mode of operation.

Proposal #17 - Train social partners and employees on the issues and practices of AI at work

The use of AI in human resources has developed at a considerable speed in recent years. Indeed, half of the AI applications developed between 2010 and 2020 are aimed at automating business management processes. Already, nearly 40% of international companies - a fourfold increase in 4 years - use predictive algorithms in their HR activities: recruitment, career management, GPEC in particular.

The integration of AI is profoundly transforming HR processes: recruitment and integration of employees, management of their career development, performance evaluation and management, active and forward-looking management of jobs and skills, and training plans that are appropriate to the company's future skill needs. At the very least, the machine becomes an aid to human decision making through its ability to process quantities of data on workers: skills, positions held, performance, etc.

As highlighted in the CNNum report on work in the age of platforms, digital transformation poses new challenges to social dialogue. This is even more true when it comes to algorithmic management.

It is therefore essential to train **staff representatives** on all these issues in order to give them the means to carry out their negotiation mission on all HR topics impacted by AI (recruitment, career management, GPEC, training ...).

What types of data are collected by the company via these HR software and algorithms, for what purposes, what data are fed into the algorithms, how is the absence of bias ensured, how much of the decisions are made by humans and how much by algorithms, etc.?

The employer and the trade unions must have this exchange as soon as the HR software and algorithms are set up in the company in order to explain the objective, the functionalities and the expected results of the software in HR management activities.

Better knowledge and training in these tools should allow for an informed and balanced discussion between the social partners on key human issues.

This training on the use of algorithms in the organisation of human work and the management of "human resources" could be part of the right to economic, social and trade union training leave and the right to training on health, safety and working

conditions provided for in Articles L2145-1 and L.2315-18 of the Labour Code respectively.

Similarly, it seems desirable that **all employees** working with algorithms are trained not only on the specific functioning of these algorithms but more broadly on the issues of AI at work and algorithmic management.

In particular, they should be informed about the nature of the personal data collected and how it is collected and processed, and how it is used by the algorithms.

The introduction of AI in the organization of human work profoundly modifies the activities and jobs of the workers concerned: hyper-standardization of certain tasks for the workers, AI leaving to the human the tasks not yet automated, on the contrary great complexity of the tasks for the workers, AI supporting the simplest tasks, management of the interactions between the human and the collaborative robots or cobot.

For workers who will be required to develop machine supervision, analysis and problem solving activities and more generally activities with higher added value, having a thorough knowledge of AI systems and their functionalities is essential in order not to lose control of the job and the ability to react in case of failure.

Proposal #18 - Build a deep public expertise capacity and train the relevant administrations to enable public authorities to produce relevant and independent regulation on AI at work

It is essential that the public authorities have the means, or more precisely give themselves the means, to produce **relevant regulations** on AI at work and to ensure both compliance with them and their updating in line with the evolution of techniques and uses of AI applied to human work.

This consideration might seem obvious, but the radical novelty, the very technical nature and the speed of evolution of the subject require a specific investment from the public authorities to have a detailed understanding of the realities and issues.

Producing regulation that balances respect for entrepreneurial freedom and the protection of workers' rights, and more broadly between the interests of business and workers, will in itself be a particularly complex task.

But the stakes go much higher. When it comes to issues as sensitive as personal data and freedoms and the possibility of human beings being manipulated by machines, it is the **basis of the pact of trust** between workers and employers that is at stake.

This is why we recommend the **constitution of an in-depth public expertise capacity on AI applied to human work**, which is independent from that of the operators. This is indeed the necessary condition for the elaboration of a policy and a regulation truly adapted to the specificities of the subject, which is neither inspired by the solutions "proposed" by the economic operators, nor copied from solutions adopted for different problems.

This expertise could be included in PEReN: Pôle d'expertise de la Régulation numérique.

Who today, at Bercy or the Ministry of Labour for example, can claim to be familiar with the issues of algorithmic management, nudge, deceptive designs, explainability and auditability of algorithms...? Building up this capacity of expertise alone is not enough: it is important to **train the administrations concerned** by the regulations on these issues.

Proposal #19 - Set up incentive systems, such as labelling, to mobilise the whole company

Making virtuous practices visible in the public arena could result in market-based punishments and incentives. This implies setting up tools to list, categorize, inventory and analyze good and bad practices (...) on the model of GoodUI for example.

The National Digital Council thus recommends setting up a digiscore that would apply to employment platforms like Uber or Deliveroo. This one, thought on the model of the "nutriscore", which would allow to educate citizens consumers by giving them a clear indication, before they use an operator, of the way this one respects certain criteria, in particular social, but also technical, in particular the way algorithms are used as management tools.

In the same perspective, but inspired by the Yuka application, the Institut Montaigne recommends making public a list of the type of data used to train the algorithm and the objectives that the algorithm must achieve.

Finally, labelling could also prove to be an appropriate tool to promote the responsible use of algorithms in management.

A first option would be to label technical tools, imposing a set of transparency and explanation criteria. This certification could be entrusted to AFNOR, under the cover of the public authorities, and involve authorities such as the Defender of Rights (for discrimination issues). It could focus "on the auditability of algorithms, on the quality of data or on the presence of a process for assessing the risks of bias within the company", as the Institut Montaigne also suggests, which also stresses that setting up this type of label would also encourage the development of more skills and methods for auditing algorithms. The same institute also recommends that the labeling process should focus on "the team developing the algorithm (its internal processes, practices, composition) rather than on the algorithms themselves. [Fair data use](#) is an example of such a label.

A second option would be to include the issue of algorithmic management more strongly in some existing labels concerning HR functions. The **Label Diversité**, created in 2008 and owned by the State, seems for example very well suited to this type of issue. Aimed at preventing discrimination and promoting **diversity** in the public and private sectors, it focuses on the human resources processes of a structure (recruitment, career management, training, communication), and more generally on all of its activities (relations with users and customers, with subcontractors, and other stakeholders), and encompasses a set of fields of an evolving nature, to which the issue of algorithms used for managerial purposes could be explicitly included.

FOCUS 4: UPDATE - Allow the regulatory framework to adapt over time to changes in AI

Proposal #20 - Include in the scope of activities considered "high risk" by the draft EU AI Regulation decisions related to compensation

The European Commission's approach to artificial intelligence is based firstly on identifying and managing the risks according to the areas of application - public or private - concerned:

- **unacceptable risks** to security, livelihoods and human rights;
- **high risks** in areas affecting critical infrastructure (energy, transport...), education, vocational training, employment, human resources, essential private and public services, law enforcement, justice, democratic processes...

Annex III of the European Commission's draft Regulation on AI (Annex III) already provides for the inclusion of recruitment, promotion and dismissal decisions, task allocation and performance evaluation in the high risk activities. However, decisions related to remuneration are formally absent, yet crucial. The text should leave no doubt that they are indeed covered by the Regulation, among the high-risk activities.

Proposal #21 - Make algorithmic management a topic of social dialogue in its own right

For Valério de Stefano, the negotiation of algorithms could, or even should, become a crucial objective for all trade union organisations in the years to come.

Indeed, it stresses that collective bargaining remains the best tool to ensure that workers' interests are respected in algorithmic management and to avoid negative impacts on their physical, mental and safety. "*Workers' representatives and trade unions must be involved in the decision-making processes that lead to the definition and use of algorithms. It is undeniably more effective to address these issues upstream than to have to do damage control afterwards.*"

As Odile Chagny and Franck Bonot, co-leaders of the Sharers & Workers network, point out, this is in principle already the case, as staff representatives have a right to be consulted in all situations where new technologies are introduced. However, they are very rarely consulted on these issues.

Yet social dialogue is crucial to implement a "*people first"* approach in the company and to regulate such sensitive issues as the scope, use and ownership of data collected on workers.

Collective agreements could set limits on the surveillance of workers through AI. They could also set criteria to increase the transparency of AI-based decision-making processes (Dagnino and Armanoli 2019), to allow for a better understanding of how these devices reach their conclusions, or more broadly address the impact of algorithms on recruitment and workforce management practices and the organisation of human work, as well as the content of jobs and the skills required.

Valerio de Stefano recalls in this regard that Article 88 of the GDPR, dealing with data processing in the context of employment relationships, recognises the importance of collective agreements for the protection of data processing rights and for the regulation of algorithmic decision-making processes.

This article defines collective agreements as important instruments to ensure fair and lawful processing of data in employment relationships. It explicitly mentions data processing for recruitment and management purposes, which means that collective agreements could offer appropriate safeguards when AI-based tools and algorithmic management practices are used in the workplace (Dagnino and Armaroli 2019; Hendrickx 2018). They could, for example, require information on how employers use workers' personal data as well as how this data is processed by AI systems (TUC 2021). They could also ban the most intrusive uses of these technologies, such as neuro-surveillance (De Stefano 2020).

Furthermore, algorithmic management makes extensive use of new forms of influence, which are not the classic exercise of authority, contractualized and based on instruction, but more or less surreptitious forms of influence. Here again, it is not so much the use of these forms that dilutes and invisibilizes authority by pushing each person to internalize the constraint that is new, but rather the way in which they are systematically implemented by certain companies.

In order to address these practices, which directly affect the individual's ability to exercise autonomy, it is necessary to make these issues the subject of collective discussions in the context of full-fledged social dialogue.

Algorithmic management could thus contribute to a significant overhaul of labour law by bringing the "content" of work back to the heart of negotiations[6]. Born from the industrial revolution, labour law indeed frames the working conditions (hours, remuneration) while leaving to a certain extent the content and the way of accomplishing the task to the discretion of the employer.

However, the relevance of this partition is challenged by the increasingly shared observation of a "loss" of meaning in work, linked to the loss of autonomy of workers who feel increasingly dispossessed of control over their expertise. However, by structuring working conditions and work content in a way that is difficult to dissociate, algorithmic management *de facto* leads to the content of work becoming a subject of collective deliberation.

The development of a social law of the algorithm is also a historic opportunity to provide a real legal framework to this movement of reappropriation.

---

[6] Chagny O. and Forestier F. (2021), "Contrebalancer les asymétries de pouvoir nées de la société numérique : vers un droit social de l'algorithme ?", Cahier n°29 du CIEP *"Inégalités programmées. Capitalism, algorithms and democracy"*

Conclusion and next steps

Historically, as Alain Supiot reminds us, trade unions have fought "against exploitation at work, not against oppression at work", which has led to the evacuation of "the content and meaning of work from the notion of social justice".

At a time when there is a growing awareness of a "loss" of meaning in work, linked in particular to a feeling of loss of autonomy on the part of workers, this is perhaps, in *the end,* **the fundamental challenge posed by algorithmic management: to put the meaning and content of work back at the centre of the debate**.

It is only through the awareness of all actors, companies, workers, trade unions, public authorities and international bodies, that we will find the appropriate responses.

All in all, while it is clear that we are still only at the beginning of the deployment of algorithmic management and its effects on work, the scale of the issues and their topicality make it necessary to act now.

To enshrine the principle of intelligibility of the algorithm, to introduce a legislative framework for certain elements of the content of the algorithms to reinforce their objectivity, to guarantee the effective access of workers to the new rights guaranteed by the European Regulations, to introduce new individual and collective rights: there is clearly a case for France to take a strong initiative within the framework of its presidency of the EU, to promote responsible AI reconciling respect for entrepreneurial freedom and the protection of workers' rights, and to set an example by changing our national legislative and regulatory framework now.

Annex 1 - Summary of proposed adjustments to the Draft EU Regulation on AI

The draft Regulation on AI proposed by the European Commission represents a major step forward. It is indeed the first global attempt to horizontally regulate AI systems and to prevent their potential misuse. In particular, its approach by categorization according to risk assessment is very relevant and allows to state in a clear and pedagogical way the prohibitions and the indispensable points of vigilance.

However, this text has significant room for improvement, particularly in relation to the gap between the aims pursued and the insufficient or ineffective nature of the measures put forward to achieve them. It also takes insufficient account of the specific challenges of AI at work.

The primary reason for this shortcoming is the inclusion of this text in the new European legislative framework which regulates non-food products in the EU, from toys to electric radiators, with the usual mechanisms of conformity, intervention by notified bodies, liability and consumer protection. This regulatory framework is known to be open to improvement. In this case, it does not effectively address the specific risks posed by AI.

In the context of our recommendations on the control of AI at work, we propose several adjustments to the European Commission's text, with 2 priority axes, :
1. adapt the draft regulation to enable it to go beyond the declaration of intent and enforce workers' rights,
2. ensure, given the singular complexity of AI-related topics, the competence of all actors in this regulation, at national and European level (authorities, CEN, CENELEC, notified bodies, data trustees, etc.).

1. Incorporate the guiding principles of this White Paper into the recitals of the draft Commission Regulation

2. Expressly prohibit algorithms that influence behaviour through the use of cognitive biases

Include in **Article 5**, which lists prohibited AI practices, AI systems that mobilize the user's cognitive biases in order to influence his or her behaviour (deceptive design that may also constitute a deceptive commercial practice).

3. Introduce a principle of explicability of algorithms

To clarify the provisions of **Article 13** in order to, inter alia :
○ introduce the notion of explicability of the algorithm, which is more precise than the notion of transparency that is envisaged in the draft regulation to varying degrees,
○ endorse the fact that this explicability must benefit not only the user of the algorithm but also the person who is the subject of a decision dictated or assisted by the algorithm,
○ specify the scope of this explicability objective, which must be applied from the design of the algorithm and downstream, to ensure the readability of audit and test results.

4. Raise the level of transparency on the use of algorithms by imposing the communication to the persons concerned of a complete and opposable document on the AI system implemented

Complete **Article 13** on transparency by adding the formalization of an obligation of information - contractual or even pre-contractual - on the part of the employer or the principal, going further than the simple information of the implementation of an algorithmic processing and imposing on the employer or the principal to provide the person subject to the algorithmic processing with a complete and opposable document on the data used, the nature, the purpose, and the consequences for him/her of this algorithmic processing.

5. Establish a right to a human decision, more effective and broader than the right not to be subject to a purely algorithmic decision

Complete **Article 14** to introduce a real right for the worker to human decision, imposing on the company implementing the algorithm a right for the worker to human decision, broader and more effective than the right that workers already have from Article 22 of the GDPR.

6. Strengthen the compliance obligation by involving Notified Bodies systematically for high risk AI systems

Complete **Article 17** by requiring for high-risk AI systems an *ex-ante* or *ex-post* assessment by notified bodies, in parallel to the already foreseen obligation of self-assessment by AI system providers.

7. Add the construction phase of the algorithm's training databases to the compliance obligation

In **Article 17**, on the quality management system for artificial intelligence systems, insert the construction phase of the learning databases (quality setting and enrichment) in the scope of the compliance obligation foreseen by the Regulation for high-risk AI (*ex-ante* and *ex-post* compliance with self-assessment by the provider and assessment by a notified body).

8. Strengthen traceability mechanisms for algorithms to include all decisions made or recommended

Clarify the terms of **Article 20** of the draft Regulation by defining the algorithmic activities subject to the automatic generation of a log and affirm the purpose of the log thus produced: to ensure the traceability of decisions and algorithmic work in order to preserve the rights of workers whose activity is affected.

This clarification should be linked to the one concerning explicability, in that the logs generated in this way should also be subject to the requirement of explicability in order to ensure that the information is legible for those to whom it is addressed, which should include the workers who request it.

9. Introduce data trustees to pool and improve the quality of training data for algorithms

Introduce data trusts in **Chapter 4 of Title III** (*Articles 30 to 39)* on notified bodies. These bodies can be set up at national level and possibly at sectoral level, but their inclusion in the Regulation would be a significant step forward.

10. Add to the requirements for high-risk AI systems an assessment of the algorithms on the model of the banking stress test

Add to **Chapter 5 of Title III** (Articles 40 to 51) procedures for estimating algorithms based on the model of banking stress tests.

11. Complete the list of high-risk activities related to AI in the workplace (activities affecting workers' compensation)

Broaden the scope of activities considered as "high risk". **Annex III** already includes AI systems implemented in recruitment to assist in the recruitment decisions themselves, and in workforce management. However, decisions related to remuneration are not expressly provided for and should in our view be explicitly mentioned.

Appendix 2 - Bibliographic References

- Adam Patrice, Le Friant Martine, Tarasewicz Yasmine (dir.), Association Française du Droit du Travail et de la Sécurité sociale (2020), Intelligence artificielle, gestion du personnel et droit du travail, Dalloz
- Adams-Prassl J. (2019) What if your boss was an algorithm? Economic incentives, legal challenges, and the rise of artificial intelligence at work, Comparative Labor Law & Policy Journal, 41 (1), 123-146.
- Adler-Bell Sam and Miller Michelle (2018), The Datafication of Employment. How Surveillance and Capitalism Are Shaping Workers' Futures without Their Knowledge, The Century Foundation
- Ajunwa I. (2019) The auditing imperative for automated hiring.
- Ajunwa I. and Greene D. (2019) Platforms at work: automated hiring platforms and other new intermediaries in the organization of work, in Vallas S.P. and Kovalainen A. (eds.) Work and labor in the digital age, Bingley, Emerald Publishing, 61-91.
- Ajunwa I., Crawford K. and Schultz J. (2017) Limitless worker surveillance, California Law Review, 105 (3), 735-776.
- Aloisi A. (2016) Commoditized workers: case study research on labour law issues arising from a set of 'on-demand/gig economy' platforms, Comparative Labor Law & Policy Journal, 37 (3), 653-690.
- Aloisi A. and De Stefano V. (2021) Frankly, my rider, I don't give a damn, Rivista il Mulino, 7 January 2021.
- Aloisi A. and Gramano E. (2019) Artificial intelligence is watching you at work: digital surveillance, employee monitoring, and regulatory issues in the EU Context, Comparative Labor Law & Policy Journal, 41 (1), 95-121.
- Aranguiz A. (2021) Spain's platform workers win algorithm transparency, Social Europe, 18 March 2021.
- Armstrong J., Eyres K. (2021), Italian DPA fines Deliveroo & food delivery start-up over AI algorithm use, Cordery Client alert
- Bertail P., Bounie D., Clémençon S., Waelbroeck P. (2019), "Algorithms: bias, discrimination and equity", TelecomParisTech and Abéona Foundation
- Bodie M.T., Cherry M.A., McCormick M.L. and Tang J. (2017) The law and policy of people analytics, University of Colorado Law Review, 88 (4), 961-964.
- Boot R. (2019) UK businesses using artificial intelligence to monitor staff activity, The Guardian
- Bouverot A., Delaporte T. (2020), "Algorithms: Bias control", Institut Montaigne
- Cappelli, Peter and Tambe, Prasanna and Yakubovich, Valery, Artificial Intelligence inHuman Resources Management: Challenges and a Path Forward (April 8, 2019)
- Chagny O. and Forestier F. (2021), "Contrebalancer les asymétries de pouvoir nées de la société numérique : vers un droit social de l'algorithme ?", Cahier n°29 du CIEP *"Inégalités programmées. Capitalism, algorithms and democracy"*
- Economic and Social Committee (2017) Opinion of the European Economic and Social Committee on 'Artificial intelligence: the impact of artificial intelligence on the (digital) single market, production, consumption, employment and society' (own-initiative opinion), Official Journal of the European Union, C 288
- Dagnino E. (2017) People analytics: lavoro e tutele al tempo del management tramite big data, Labour & Law Issues, 3 (1), I.1-31.

- Dagnino E. and Armanoli I. (2019) A seat at the table: negotiating data processing in the workplace, Comparative Labor Law & Policy Journal, 41 (1), 173-196.
- Dastin, J. (2018), "Amazon scraps secret AI recruiting tool that showed bias against women", Reuters, Reuters, (accessed on 13 August 2020).
- De Stefano V. (2016) The rise of the 'just-in-time workforce': on-demand work, crowdwork and labour protection in the 'gig-economy, Comparative Labor Law & Policy Journal, 37 (3), 471-504.
- De Stefano V. (2019) 'Negotiating the algorithm': automation, artificial intelligence, and labor protection, Comparative Labor Law & Policy Journal, 41 (1), 1-32.
- De Stefano V. (2020) 'Masters and servants': collective labour rights and private government in the contemporary world of work, International Journal of Comparative Labour Law and Industrial Relations, 36 (4), 425-444
- De Stefano V. , Taes S. (2021), "Algorithmic management and collective bargaining", Perspectives Note, ETUI
- Estlund C. (2018) Book review: Rethinking autocracy at work, Harvard Law Review, 175, 795-826.
- Eubanks V. (2018) Automating inequality: how high-tech tools profile, police, and punish the poor, New-York, St. Martin's Press. FEPS (2017) Work in the European gig-economy, Brussels, Foundation for European Progressive Studies.
- Fillon W. (2018) How rating your server is making their life miserable, Eater, 22 June 2018.
- Fischbach K., Gloor P.A., Lassenius C., Olguin Olguin D., Pentland A.S., Putzke J. and Schoder D. (2010) Analyzing the flow of knowledge with sociometric badges, Procedia - Social and Behavioral Sciences, 2 (4), 6389-6397.
- Forestier F., Bonot F., Giniaux-Kats J., Chagny O. (2021), "Faced with platforms, we need a real social right for the algorithm", Politis
- Giniaux-Kats J., Campanini E., Dufour M., Forestier F. (2021), ''Le management algorithmique, nouveau paradigme d'organisation dutravail ? '', AOC
- Gonfalonieri A. (2020) What brain-computer interfaces could mean for the future of work, Harvard Business Review
- Guszcza, J. et al. (2018), "Why We Need to Audit Algorithms", Harvard Business Review
- Harwell D. (2019) A face-scanning algorithm increasingly decides whether you deserve the job, The Washington Post
- Hendrickx F. (2018) Privacy, data protection and measuring employee performance: the triggers from technology and smart work, European Labour Law Journal, 9 (2), 99-100
- Kellogg Katherine C., Valentine Melissa A., Christin Angèle (2020), Algorithms at work: the new contested terrain of control, Academy of Management Annals 2020, Vol. 14, No. 1, 366-410
- Lee S.M. (2017) Workplaces are tracking their employees' sleep, Buzzfeed.News.
- Li Lan, Lassiter Tina, Lee Min Kyung (2021), Algorithmic Hiring in Practice: Recruiter and HR Professional's Perspectives on AI Use in Hiring, Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society
- Moore P.V. (2019) OSH and the future of work: benefits and risks of artificial intelligence tools in workplace, EU-OSHA Discussion Paper.

- Moore P.V., Akhtar P. and Upchurch M. (2018) Digitalisation of work and resistance, in Moore P.V., Upchurch M. and Whittaker X. (Eds.) Humans and machines at work, Cham, Palgrave Macmillan, 2018, 17-44.
- OECD (2019), Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449 (May 12, 2019),
- OECD (2021), The impact of AI on the labour market: what do we know so far? , January 2021
- O'Donovan C. (2018) An invisible rating system at your favorite chain restaurant is costing your server, BuzzFeed.News, 21 June 2018
- O'Neill C. (2016) Weapons of math destruction: how big data increases inequality and threatens democracy, London, Penguin.
- Pasquale F. (2015) The black box society: the secret algorithms that control money and information, Boston, Harvard University Press.
- Samuel S. (2019) Brain-reading tech is coming: the law is not ready to protect us, Vox, 20 December 2019.
- Satariano S. (2020) How my boss monitors me while I work from home, The New York Times, 7 May 2020.
- Schubert C. and Hütt M.-T. (2019) Economy-on-demand and the fairness of algorithms, European Labour Law Journal, 10 (1), 3-16.
- The Economist (2018) AI-spy. The workplace of the future, 28 March 2018.
- TUC (2021) Dignity at work and the AI revolution: a TUC manifesto, London, Trade Union Congress.
- Zuboff S. (2019), The age of surveillance capitalism: the fight for a human future at the new frontier of power, London, Profile Books.

Leplusimportant

in partnership with

**GENEVA
MACRO
LABS**