

#Leplusimportant

“Reprendre le contrôle de l’IA au travail : 20 propositions pour promouvoir un management algorithmique responsable”



LIVRE BLANC

Par Jérémie **Giniaux-Kats**, directeur du Pôle Management algorithmique de #Leplusimportant, avec Erik **Campanini**, Mathias **Dufour**, Florian **Forestier**, Pauline **Maury**, Sylvie **Teynier**

En partenariat avec **GENEVA
MACRO
LABS**

Octobre 2021

Présentation du Livre Blanc

Le management algorithmique, c'est-à-dire l'usage de l'IA pour recruter et gérer la main d'œuvre et organiser le travail humain, est en train de bouleverser le travail, son organisation et les relations entre employeurs et travailleurs. La combinaison de deux expressions célèbres : "[*software is eating the world*](#)" et "[*code is law*](#)" permet d'entrevoir l'ampleur exceptionnelle de ce bouleversement, qui touche potentiellement l'ensemble des entreprises et travailleurs.

Alors que le droit du travail est encore très centré sur l'organisation des relations de travail pré-numérique, prendre la mesure des enjeux posés par le management algorithmique sur le travail et agir dès maintenant pour l'encadrer est donc absolument critique. Il s'agit de permettre d'exploiter son potentiel tout en garantissant des conditions de travail qui préservent la dignité, la santé, l'égalité de traitement et l'autonomie des travailleurs.

Ce projet de Livre Blanc fournit à cette fin des premiers éléments d'**analyse des enjeux** et des **principes directeurs**, déclinés en **4 axes de propositions** concrètes

1. **PRÉVENIR - Adopter des règles de conception sécurisant une IA au travail responsable**
2. **PROTÉGER - Adapter les droits individuels des travailleurs aux transformations du travail portées par les algorithmes**
3. **GARANTIR - Assurer le respect effectif des principes de l'IA responsable et des droits travailleurs**
4. **ACTUALISER - Permettre l'adaptation du cadre réglementaire dans le temps aux évolutions de l'IA**

Ce Livre Blanc est publié en [partenariat avec Geneva Macro Labs](#) à l'occasion du colloque international "**Pour un Web des Lumières : remettre la culture numérique au service du projet d'émancipation des Lumières**" organisé le 21 octobre 2021.

Il vient compléter nos premières contributions sur le sujet :

- Article "[Le management algorithmique, nouveau paradigme d'organisation du travail ?](#)" dans AOC en avril 2021
- Article "[Face aux plateformes, il faut un véritable droit social de l'algorithme](#)", dans Politis de janvier 2021
- Article "[Contrebalancer les asymétries de pouvoir nées de la société numérique : vers un droit social de l'algorithme ?](#)" dans le Cahier n°29 du CIEP "*Inégalités programmées. Capitalisme, algorithmes et démocratie*"
- Le [replay vidéo](#) et [podcast](#) de notre webinaire de juin 2021 "*Travail et management algorithmique : quels enjeux ? comment reprendre le contrôle ?*"

avec des intervenants de l'OIT, de l'OCDE, de Stanford et de Salesforce avec notre partenaire [Geneva Macro Labs](#)

Nous remercions chaleureusement Guillaume Bon, Anna Byhovskaya, Odile Chagny, Ekkehard Ernst, Clémentine Kleitz, Alexis Masse, Franck Morel, Louis-Laurent Preux, Eva Thoreau et Paul Welti pour leur contribution à ce document, qui n'engage que #Leplusimportant.

Sommaire du Livre Blanc et récapitulatif des propositions

Présentation du Livre Blanc	2
Synthèse	6
I. Management algorithmique : une réalité déjà présente, un potentiel considérable et des enjeux majeurs	13
1A. Management algorithmique : de quoi parle-t-on ?	13
1B. Opportunités et menaces pour les travailleurs	15
1C. Enjeux pour les entreprises et les travailleurs	19
II. Principes directeurs pour promouvoir un management algorithmique responsable	23
2A. 10 principes directeurs relatifs aux entreprises	24
2B. 10 principes directeurs relatifs aux travailleurs	25
2C. 4 principes directeurs relatifs aux pouvoirs publics	26
III. 20 propositions concrètes pour promouvoir un management algorithmique responsable	28
AXE 1 : PRÉVENIR - Adopter des règles de conception sécurisant une IA au travail responsable	28
Proposition #1 - Introduire un objectif d’explicabilité et d’éthique “by design” des algorithmes	28
Proposition #2 - Assurer la traçabilité des décisions algorithmiques	30
Proposition #3 - Rendre obligatoires certaines séquences au sein des algorithmes pour garantir les droits des travailleurs et inversement interdire des séquences aux effets délétères	32
Proposition #4 - Encadrer la conception des algorithmes, notamment par un contrôle en amont des bases de données d’apprentissage pour garantir leur diversité et minimiser les biais des algorithmes	32
Proposition #5 - Prévoir une mutualisation des données d’apprentissage auprès d’institutions tierces, relevant de l’Etat, ou prenant la forme de fiduciaires de données (“data trusts”) pour minimiser les biais	33
Proposition #6 - Placer certains types de données sous le statut de “données d’intérêt général”	35
Proposition #7 - Mettre en place un encadrement spécifique du Cycle de Vie des Modèles pour minimiser les biais et le risque de dérive en qualité de modèles au cours du temps	36
Proposition #8 - Encadrer la mise en production des algorithmes avec différentes parties prenantes	38

AXE 2 : PROTÉGER - Adapter les droits des travailleurs aux transformations du travail portées par les algorithmes	40
Proposition #9 - Introduire un véritable droit à la décision humaine	40
Proposition #10 - Introduire un droit de communication d'un document intelligible et opposable sur les algorithmes au travail	41
Proposition #11 - Assurer un meilleur partage de valeur créée liée à l'usage de l'IA au travail entre l'entreprise et les travailleurs producteurs de données	42
AXE 3 : GARANTIR - Garantir le respect effectif des principes de l'IA responsable et des droits travailleurs	44
Proposition #12 - Passer de l'autoévaluation de la conformité réglementaire des algorithmes par des fournisseurs à une évaluation par des organismes tiers	44
Proposition #13 - Mettre en place des procédures d'estimation des risques des algorithmes, sur le modèle du stress test bancaire	45
Proposition #14 - Intégrer la mesure des risques liés au management algorithmique dans les risques au sein des entreprises, notamment psychosociaux	47
Proposition #15 - Élargir les responsabilités du DPO au champ du management algorithmique	49
Proposition #16 - Intégrer les droits individuels et collectifs relatifs aux algorithmes dans les codes de conduite des entreprises	50
Proposition #17 - Former partenaires sociaux et salariés aux enjeux et aux pratiques de l'IA au travail	50
Proposition #18 - Constituer une capacité d'expertise publique approfondie et former les administrations concernées pour permettre aux pouvoirs publics de produire une réglementation pertinente et indépendante sur l'IA au travail	52
Proposition #19 - Mettre en place des dispositifs incitatifs, de type labellisation, pour mobiliser l'ensemble de l'entreprise	53
AXE 4 : ACTUALISER - Permettre l'adaptation du cadre réglementaire dans le temps aux évolutions de l'IA	55
Proposition #20 - Inclure dans le périmètre des activités considérées comme "à haut risque" par le projet de Règlement européen sur l'IA les décisions liées à la rémunération	55
Proposition #21 - Faire du management algorithmique un thème du dialogue social à part entière	56
Conclusion et prochaines étapes	59
Annexe 1 - Récapitulatif des ajustements proposés au Projet de Règlement européen sur l'IA	60
Annexe 2 - Références bibliographiques	63

Synthèse

Alors que le droit du travail est encore très centré sur l'organisation des relations de travail pré-numérique, prendre la mesure des enjeux posés par le management algorithmique sur le travail et agir dès maintenant pour l'encadrer est absolument critique. Il s'agit de permettre d'exploiter son potentiel tout en garantissant des conditions de travail et d'emploi qui préservent la dignité, la santé, l'égalité de traitement et l'autonomie des travailleurs.

Ce Livre Blanc a ainsi pour objet de fournir des premiers éléments d'analyse des enjeux, des propositions concrètes pour promouvoir un management algorithmique responsable ainsi, en annexe, que des propositions d'ajustement au projet de Règlement de la Commission européenne.

I. Management algorithmique : une réalité déjà présente, des enjeux majeurs

Les algorithmes sont utilisés depuis longtemps pour optimiser le fonctionnement des machines. Assurance, finance, météo, les exemples abondent. La nouveauté est leur utilisation pour **organiser le travail humain** et **gérer les ressources humaines** au sein des entreprises.

Dans le [projet de règlement de la Commission européenne sur l'IA](#) (Annexe III, considérant 36) deux thèmes liés à l'emploi et au travail sont identifiés comme "à haut risque" : (1) le **recrutement** et (2) la **gestion de la main d'œuvre** : les décisions liées à la promotion, au licenciement, à l'**attribution des tâches**, au suivi et à l'évaluation des performances et du comportement des travailleurs.

De fait, parmi les cas d'usages qui se développent figurent l'utilisation de l'IA pour le recrutement (notamment pour le filtrage des candidatures, l'évaluation des candidats au cours d'entretiens, de tests ou d'[analyses automatisées de vidéos](#)) ou pour prendre des décisions concernant la promotion et la cessation de relations contractuelles liées au travail, pour la répartition des tâches et pour le suivi et l'évaluation des performances et du comportement des personnes dans le cadre de ces relations.

Si ces usages de l'IA au travail se développent aujourd'hui en particulier dans la distribution, les centres d'appels, les entrepôts, ils vont se généraliser dans tous les secteurs. Selon une étude du cabinet Mercer menée auprès de 7 500 managers RH à travers le monde, **en 2020 39 % des services RH utilisent des algorithmes prédictifs**, contre 10% en 2016, soit un **quadruplement en 4 ans** ! Dans cette même étude, la généralisation de [l'IA apparaît comme le troisième levier stratégique pour faire face à la crise](#).

Au total, le BIT distingue **5 éléments constitutifs du management algorithmique** : la [surveillance constante](#), l'évaluation permanente des performances, l'application automatique des décisions sans intervention humaine, l'interaction des travailleurs avec un système, la faible transparence des algorithmes.

L'IA au travail, comme toute grande innovation technologique, représente à la fois des **opportunités et des menaces pour les travailleurs**.

Si l'usage de l'IA par les employeurs est d'abord motivé par sa capacité à améliorer la productivité ([Accenture annonce 40% de gains de productivité à l'échelle mondiale d'ici 2035](#) grâce à l'IA), elle peut présenter plusieurs avantages pour les travailleurs eux-mêmes. Elle a le potentiel de rendre leur activité plus productive et plus facile, moins soumise à l'arbitraire, et plus intéressante. Une [étude menée auprès de 10 000 travailleurs](#) montre ainsi une augmentation de la satisfaction au travail pour ceux qui utilisent des algorithmes dans leur travail. Plus spécifiquement, l'IA peut contribuer à ce que les travailleurs soient traités avec moins de biais et de discriminations et qu'ils soient traités de façon plus personnalisée par rapport à leurs besoins et aspirations.

Le management algorithmique ouvre à l'inverse aussi pour les travailleurs **4 principaux types de risques négatifs** : (1) surveillance abusive, (2) discriminations, (3) perte d'autonomie et risques psycho-sociaux et (4) obsolescence de compétences et perte d'employabilité.

(1) En premier lieu, l'exceptionnel développement des capacités de collecte et de traitement des données personnelles sur le travail et la vie des salariés par les machines - en d'autres termes, de **surveillance des travailleurs** - constitue la toile de fond du développement de l'IA au travail (à l'image des «[badges intelligents](#)» qui mesurent la vitesse à laquelle les travailleurs effectuent des tâches, suivent leurs interactions et analysent leurs conversations ou encore des outils d'IA qui analysent les courriels des travailleurs pour contrôler leur productivité, repérer les travailleurs les plus innovants, ou les comportements déviants).

Pour le [Centre commun de recherche de la Commission européenne](#), les algorithmes fonctionnent ainsi comme des « *instruments de surveillance qui remplacent l'encadrement direct et créent des asymétries de puissance* ».

(2) L'IA pourrait aussi être porteuse de **discriminations accrues**, fussent-elles [involontaires](#) voire même non désirées mais liées à l'opacité des critères effectifs de décision des algorithmes. [Amazon a ainsi dû renoncer à un programme d'algorithme de recrutement](#), ayant constaté qu'il défavorisait les candidatures féminines pour les emplois tech.

Par ailleurs, l'utilisation généralisée des algorithmes dans la rémunération des employés augmente le risque d'une « [collusion algorithmique](#) » à travers plusieurs employeurs, créant ainsi des conditions d'abaissement des conditions de travail.

(3) Le management algorithmique va conduire un nombre croissant de travailleurs à voir leur travail influencé voire dirigé par des processus que les travailleurs ne maîtrisent pas. La perception d'un transfert de la capacité de décision du travailleur à la machine nourrit des craintes de **perte d'autonomie**, y compris chez les travailleurs qualifiés et les managers. Celle-ci peut même aller jusqu'à un **sentiment d'aliénation**, lié à la perception de devenir un appendice de la machine.

(4) Enfin l'IA au travail, [d'ores et déjà à l'origine de disparités salariales](#), fait peser sur un certain nombre de travailleurs un **risque d'obsolescence de compétences**, en interne, et de **moins employabilité**, en externe, qui en retour appellent des mesures spécifiques de la part des entreprises, voire des pouvoirs publics, pour les former aux différentes modalités de travail avec l'IA.

Plus largement, **le management algorithmique présente 5 types d'enjeux** : des enjeux d'**explicabilité**, d'**objectivité**, de **responsabilité**, d'**adaptation des droits des travailleurs** et enfin plus largement de **partage de la valeur**.

1. Comment garantir l'**explicabilité des algorithmes**, dont les règles de fonctionnement sont aujourd'hui opaques ? Cette question engage par exemple directement la **capacité de certification des régulateurs**, pour vérifier que les algorithmes respectent le cadre normatif. Les conséquences tragiques de la validation par le régulateur américain du logiciel MCAS d'aide au pilotage des Boeing 737 Max, sans en réalité le comprendre faute d'avoir l'expertise requise, montrent l'ampleur et l'actualité de ces enjeux.

2. Comment garantir l'**objectivité des algorithmes** et éviter les **biais et discriminations** ? A titre illustratif, le tribunal de Bologne vient de [juger discriminatoire l'algorithme de Deliveroo](#), qui utilisait un indice de "fiabilité" pour classer les livreurs et ainsi prioriser l'attribution des courses.

3. Qui est **responsable** du fonctionnement des algorithmes et de leurs résultats ? L'algorithme floute l'identité du donneur d'ordre, et d'autant plus en substituant la suggestion ("*nudge*") à l'injonction. Ne relevant pas de l'exercice classique de l'autorité, ces formes d'influence posent de véritables défis au droit. Que reste-t-il de la responsabilité humaine, comment l'attribuer ?

4. Comment adapter les **droits des travailleurs à l'ère du management algorithmique (et maintenir leur employabilité)** ? Comment plus largement rééquilibrer les rapports entre le travailleur et le donneur d'ordre face au management

algorithmique ? Comment garantir l'accès effectif à ces droits, dont la complexité risque de les rendre inopérants ?

5. *Last but not least*, le **partage de la valeur** constitue un enjeu essentiel : quelle part de la valeur économique créée est redistribuée aux travailleurs concernés ? Dans quelle mesure les travailleurs peuvent-ils pouvoir revendiquer et obtenir leur juste contribution au succès de l'entreprise?

II. Principes directeurs pour promouvoir un management algorithmique responsable

L'enjeu n'est autre que de **reprenre le contrôle de l'IA au travail : promouvoir un management algorithmique responsable**. Pour exploiter son potentiel tout en garantissant des conditions de travail qui préservent la dignité, la santé, l'égalité de traitement et l'autonomie des travailleurs, il est essentiel de définir en quoi consisterait un management algorithmique responsable : nous proposons à cette fin plusieurs **principes directeurs**.

10 principes directeurs relatifs aux entreprises pour favoriser le déploiement d'un management algorithmique responsable

L'usage de l'IA par les entreprises pour organiser le travail humain relève, d'abord et avant tout, de la liberté d'entreprendre : c'est un moyen pour elles d'améliorer leur efficacité et de mieux remplir les missions qui sont les leurs. Les contraintes qui devraient leur être imposées doivent être strictement nécessaires et proportionnées au but qu'elles poursuivent (comme la protection des droits et libertés fondamentales des personnes).

Elles ont ainsi droit à ce que le cadre réglementaire relatif au management algorithmique respecte la **propriété intellectuelle**, garantisse un **jeu concurrentiel équitable** entre les acteurs, préserve la capacité d'**innovation** et garantisse leur **sécurité juridique**.

A l'inverse, les enjeux très forts liés au management algorithmique imposent aux entreprises des **responsabilités** spécifiques en la matière : **garantir les droits de leurs travailleurs** (*cf. infra. Principes directeurs pour les travailleurs*), **informer les travailleurs** sur leurs pratiques en matière d'usages de l'IA pour organiser le travail, en particulier sur la nature des données personnelles collectées, **former leurs travailleurs** aux pratiques de management algorithmique, **investir dans l'employabilité des travailleurs** si celle-ci est remise en cause par l'IA, **faire de l'IA et de ses usages au travail un thème de dialogue social** et **se prêter à des audits externes** pour vérifier le respect des points ci-dessus.

10 principes directeurs pour protéger les travailleurs et leur permettre de tirer le meilleur de l'IA

Pour prévenir les risques posés par le management algorithmique pour les travailleurs (surveillance abusive, discriminations, risques psycho-sociaux, perte d'autonomie, perte d'employabilité), nous proposons **10 principes directeurs**, relatifs à leur **droit à la dignité** (droit à la vie privée, droit d'information et de consentement, droit aux interactions humaines), leur **droit à l'égalité de traitement** (droit à la non-discrimination), leur **droit à la santé** et leur **droit à l'autonomie** (droit à la compréhension ou à l'explicabilité des algorithmes, droit à la décision humaine, droit à la contestation, droit à la formation à l'usage de l'IA et droit de négociation).

Enfin, à l'évidence, le management algorithmique soulève des enjeux qui appellent une intervention spécifique de la part des **pouvoirs publics**. Celle-ci se fonde sur une **quadruple responsabilité** : disposer d'une **expertise profonde et autonome** pour pouvoir produire une réglementation pertinente et indépendante sur l'IA au travail, garantir un **accès effectif aux droits** (que les règles édictées soient bien respectées et les droits garantis effectivement accessibles, en particulier pour les travailleurs individuels), assurer un **soutien à l'employabilité** et enfin assurer **l'adaptabilité du cadre réglementaire**, qui ne doit pas être défini une fois pour toutes mais être très évolutif, au risque sinon de perdre sa pertinence.

Nous avons identifié 21 propositions afin de décliner ces principes en actions concrètes, organisées selon 4 axes d'action, dont voici les principales :

AXE 1 : PRÉVENIR - Adopter des règles de conception sécurisant une IA au travail responsable

Proposition #1 - Introduire un **objectif d'explicabilité des algorithmes** et d'éthique "by design"

Proposition #2 - Assurer la **traçabilité des décisions** algorithmiques

Proposition #3 - Rendre **obligatoires certaines séquences** au sein des algorithmes pour garantir les droits des travailleurs et inversement interdire des séquences aux effets délétères

Proposition #4 - **Encadrer la conception** des algorithmes, notamment par un contrôle en amont des bases de données d'apprentissage pour garantir leur diversité et minimiser les biais des algorithmes

Proposition #7 - Mettre en place un **encadrement** spécifique **du cycle de vie des modèles** pour minimiser les biais et le risque de dérive en qualité de modèles au cours du temps

AXE 2 : PROTÉGER - Adapter les droits des travailleurs aux transformations du travail portées par les algorithmes

Proposition #9 - Introduire un véritable **droit à la décision humaine**

Proposition #10 - Introduire un **droit de communication** d'un document intelligible et opposable sur les algorithmes au travail

AXE 3 : GARANTIR - Garantir le respect effectif des principes de l'IA responsable et des droits travailleurs

Proposition #12 - Passer de l'autoévaluation de la conformité réglementaire des algorithmes par des fournisseurs à une **évaluation par des organismes tiers**

Proposition #13 - Mettre en place des **procédures d'estimation des risques** des algorithmes, sur le modèle du stress test bancaire

Proposition #18 - Constituer une **capacité d'expertise publique** approfondie et former les administrations concernées pour permettre aux pouvoirs publics de produire une réglementation pertinente et indépendante sur l'IA au travail

AXE 4 : ACTUALISER - Permettre l'adaptation du cadre réglementaire dans le temps aux évolutions de l'IA

Proposition #21 - Faire du management algorithmique un **thème du dialogue social** à part entière

Historiquement, comme le rappelle Alain Supiot, les syndicats se sont battus "[contre l'exploitation au travail, pas contre l'oppression dans le travail](#)", ce qui a conduit à évacuer "le contenu et le sens du travail de la notion de justice sociale".

A l'heure du constat de plus en plus partagé d'une « perte » de sens du travail, liée notamment à un sentiment de perte d'autonomie de la part des travailleurs, c'est peut-être là, *in fine*, **l'enjeu fondamental posé par le management algorithmique : remettre le sens et le contenu du travail au centre du débat.**

Ce n'est que par la prise de conscience de tous les acteurs -entreprises, travailleurs, syndicats, pouvoirs publics- que nous trouverons les réponses adaptées. Or même en n'étant encore qu'à l'orée du déploiement du management algorithmique et de ses effets sur le travail, l'ampleur des enjeux et leur actualité imposent d'agir dès maintenant.

A cet égard, le projet de Règlement sur l'IA proposé par la Commission européenne représente une avancée majeure. Il constitue la première tentative mondiale de réglementation horizontale des systèmes d'IA et de prévention des usages abusifs qui pourraient en être faits. Il permet d'affirmer de manière claire les interdits et les points de vigilance indispensables.

Toutefois ce texte présente des marges d'amélioration importantes, liées notamment à l'écart entre les finalités poursuivies et le caractère insuffisant ou ineffectif des mesures avancées pour les atteindre. Il prend aussi insuffisamment en compte les enjeux

spécifiques de l'IA au travail. Une synthèse de nos propositions d'ajustements au texte de la Commission européenne, centrés sur 2 axes prioritaires, figure en Annexe :

1. adapter le projet de règlement pour lui permettre de dépasser la déclaration d'intention et de faire respecter les droits des travailleurs,
2. assurer, compte tenu de la complexité singulière des sujets liés à l'IA, la compétence de tous les acteurs de cette réglementation, au niveau national et européen.

I. Management algorithmique : une réalité déjà présente, un potentiel considérable et des enjeux majeurs

1A. Management algorithmique : de quoi parle-t-on ?

Quel point commun entre les Boeing 737 Max cloués au sol à cause d'une défaillance de leur logiciel d'aide au pilotage, les applis pour se faire livrer chez soi en quelques minutes et le recours croissant d'entreprises à des solutions de recrutement 100% automatisé en ligne ?

Le management algorithmique, c'est-à-dire l'usage de l'IA pour organiser le travail humain.

Les algorithmes sont utilisés depuis longtemps pour optimiser le fonctionnement des machines. Assurance, finance, météo, les exemples abondent. Ce qui est nouveau, c'est leur utilisation pour organiser le travail et gérer les ressources humaines au sein des entreprises.

Dans le [projet de règlement de la Commission européenne sur l'IA](#) (Annexe III, considérant 36) deux thèmes liés à l'emploi et au travail sont identifiés comme "à haut risque" : (1) le **recrutement** et (2) la **gestion de la main d'œuvre** : les décisions liées à la promotion, au licenciement, à l'attribution des tâches, au suivi et à l'évaluation des performances et du comportement des travailleurs.

De fait, parmi les cas d'usages qui se développent figurent l'utilisation de l'IA pour le recrutement (notamment pour le filtrage des candidatures, l'évaluation des candidats au cours d'entretiens, de tests ou d'[analyses automatisées de vidéos](#)) ou pour prendre des décisions concernant la promotion et la cessation de relations contractuelles liées au travail, pour la répartition des tâches et pour le suivi et l'évaluation des performances et du comportement des personnes dans le cadre de ces relations.

La généralisation du télétravail, avec le Covid, a par exemple conduit les entreprises à recruter des candidats entièrement à distance et à démultiplier le recours aux logiciels de recrutement fondés sur l'IA. L'Oréal utilise ainsi l'IA pour sa gestion des plus de 100.000 candidatures annuelles de stages. La rédaction des offres d'emploi, la recherche de candidats et même leur sélection peuvent désormais être quasi-automatisées. Ce recours croissant à l'IA nourrit des [débats sur le renforcement des discriminations à l'embauche](#). Voire à la débauche, dans le cas de [travailleurs d'Amazon dont les ruptures de contrat ont été décidées par des algorithmes](#).

Au-delà de l'application de l'IA pour le recrutement et pour l'appui aux processus de gestion RH, le cœur du management algorithmique, c'est l'organisation du travail humain, qui prend différentes formes. D'abord, une hyper-standardisation des tâches pour les travailleurs, l'IA laissant à l'humain des tâches non encore automatisées. Ensuite, une hyper-complexification des tâches pour les travailleurs, l'IA supportant les tâches les plus simples. Enfin, le développement des relations entre l'humain et les robots collaboratifs ou cobots.

Si ces usages se développent aujourd'hui en particulier dans la distribution, les centres d'appels, les entrepôts, ils vont se généraliser dans tous les secteurs. Selon une étude du cabinet Mercer menée auprès de 7 500 managers RH à travers le monde, **en 2020 39 % des services RH utilisent des algorithmes prédictifs**, contre 10% en 2016, soit un **quadruplement en 4 ans** ! Dans cette même étude, la généralisation de **l'IA apparaît comme le troisième levier stratégique pour faire face à la crise.**

Le concept de “management algorithmique” a été forgé récemment par les chercheurs pour décrire le mode d'organisation du travail introduit par les plateformes numériques de travail comme Uber et Lyft. Mais loin de se limiter à ces plateformes, il concerne en réalité l'ensemble de l'économie. Un certain nombre de traits sont d'ailleurs en réalité présents aux Etats-Unis dès la fin des années 80 (en particulier dans la grande distribution), mais ont connu un net essor avec le déploiement de la vague actuelle de transformation numérique et les ruptures technologiques qui l'ont accompagnée – cloud, big data, applications mobiles, géolocalisation, internet des objets, machines apprenantes, etc. Amazon par exemple, parmi les enseignes qui traitent massivement de la donnée, fait partie des entreprises les plus avancées en matière de management algorithmique à l'égard de ses salariés.

Au total, le **BIT** distingue **5 éléments constitutifs du management algorithmique** : la **surveillance constante**, l'évaluation permanente des performances, l'application automatique des décisions sans intervention humaine, l'interaction des travailleurs avec un système, la faible transparence des algorithmes.

Certains chercheurs **considèrent le management par algorithme comme un nouveau paradigme d'organisation et de contrôle**. Il modifie en effet en profondeur chacune des trois dimensions de la “subordination” (au sens juridique du terme) des travailleurs salariés, à savoir les mécanismes de direction (les directives algorithmiques remplacent les directives hiérarchiques), d'évaluation (fondée sur une surveillance algorithmique continue), et de sanction et rétribution (liées au respect des directives de l'algorithme).

La **direction** prend à la fois la forme de recommandations algorithmiques incitant les travailleurs à adopter des comportements préétablis, ou l'impossibilité de moduler ou de personnaliser l'exécution de ses tâches (le chauffeur doit suivre le trajet déterminé

par l'algorithme utilisé par Uber) et parfois d'une restriction de son accès aux données ou à la clientèle (un chauffeur Uber ne sait pas ce qui conduit l'algorithme à lui affecter tel client plutôt que tel autre).

L'**évaluation** ou le contrôle, par le biais algorithmique, repose sur une surveillance permanente du travailleur pendant l'exécution de ses tâches et une notation, parfois en temps réel, de ses performances. Cet enregistrement algorithmique en continu peut d'ailleurs se nourrir tant de données internes que de données externes à la relation de travail. Ce contrôle au moyen de l'algorithme peut, de fait, se substituer en tout ou en partie au pouvoir de contrôle de l'employeur.

La **sanction**, par l'usage de l'algorithme, tend à automatiser la hiérarchisation des travailleurs selon leurs performances, notamment par la réduction de l'accès à la clientèle pour les travailleurs moins bien notés (jusqu'à la rupture de la relation) ou au contraire, par l'octroi automatisé d'avantages pour ceux qui suivent à la lettre les recommandations de l'algorithme dans l'exécution de leurs tâches (davantage d'opportunités de rémunération, de meilleures conditions financières, etc.).

1B. Opportunités et menaces pour les travailleurs

L'IA au travail, comme toute grande innovation technologique, représente pour les travailleurs à la fois des opportunités et des menaces.

Tout d'abord, l'usage de l'IA par les employeurs est principalement motivé par la volonté d'améliorer l'efficacité des processus décisionnels, la coordination et la capacité d'apprentissage de leur organisation. Le cabinet [Accenture annonce que d'ici 2035](#), l'IA pourrait contribuer à augmenter la productivité mondiale de 40 % par rapport à 2016.

Dans le domaine RH par exemple, une étude d'IBM indique que les deux tiers des PDG des dix premières économies mondiales estiment que l'IA sera génératrice de valeur pour les RH. Selon le cabinet Deloitte, l'usage des outils de "*people analytics*" (analyse de données par l'IA appliquée à la gestion de la main d'oeuvre) serait pour 71 % des compagnies internationales une priorité élevée pour leur organisation, avec pour bénéfiques l'obtention d'informations métier précieuses et la prise en charge des « problèmes humains ».

Au-delà de l'intérêt pour l'entreprise, l'IA peut présenter plusieurs avantages pour les travailleurs.

Elle a le potentiel de rendre l'activité des travailleurs plus productive et plus facile, moins soumise à l'arbitraire, et plus intéressante. Une [étude menée auprès de 10 000](#)

[travailleurs](#) au Japon (certes le 2ème pays le plus robotisé au monde) montre ainsi une augmentation de la satisfaction au travail pour ceux qui utilisent des algorithmes dans leur travail.

Par ailleurs, l'utilisation des algorithmes, notamment dans les chaînes de production, peut contribuer à diminuer les risques d'accident, suite à une mauvaise manipulation et à la fatigue, notamment. Ainsi, la surveillance des étapes de production ou des consignes de sécurité diminue le risque de mauvaise manipulation et des accidents de travail.

[L'OCDE](#) cite ainsi parmi les opportunités liées à l'IA pour les travailleurs :

- l'augmentation des capacités humaines et l'amélioration de la productivité et de la créativité, libérées par la prise en charge par l'IA des tâches les plus répétitives. [Selon une étude de 2017](#), d'ici 10 ans, plus de 80% de ces tâches humaines pourraient être remplacées par l'algorithme. Paradoxalement le déploiement de l'IA dans l'entreprise pourrait rendre le travail plus humain, concentrant ce dernier sur sa plus forte valeur ajoutée.
- la promotion de l'inclusion des populations sous-représentées,
- la réduction des inégalités économiques, sociales, de genre et autres,
- et la protection des environnements naturels, ce qui stimule la croissance inclusive, le développement durable et le bien-être.

Plus spécifiquement, l'IA peut contribuer à ce que les travailleurs soient traités avec moins de biais et de discriminations et qu'ils soient traités de façon plus personnalisée par rapport à leurs besoins et aspirations. Ceci peut, notamment, permettre une meilleure évaluation de l'ensemble du parcours professionnel d'un travailleur, au-delà des certificats et diplômes obtenus, fournissant ainsi une image holistique de la personne. La réorientation professionnelle se voit alors améliorée à l'aide d'un [support algorithmique](#).

Le management algorithmique ouvre à l'inverse aussi pour les travailleurs **quatre principaux types de [risques négatifs](#)** : (1) surveillance abusive, (2) discriminations, (3) perte d'autonomie et risques psycho-sociaux, (4) obsolescence de compétences et perte d'employabilité.

Logiquement, nous nous étendrons plus sur l'analyse de ces risques et de la manière de les prévenir que sur celle des opportunités, car ce sont eux qui appellent en priorité une action proactive, à la fois des entreprises et des pouvoirs publics.

(1) En premier lieu, l'exceptionnel développement des capacités de collecte et de traitement des données personnelles sur le travail et la vie des salariés par les machines - en d'autres termes, de **surveillance des travailleurs** - constitue la toile de fond du

développement de l'IA au travail. [Valerio de Stefano cite](#) ainsi le recours croissant à “des accessoires connectés, à l'image des «badges intelligents», qui mesurent la vitesse à laquelle les travailleurs se déplacent et effectuent des tâches, qui suivent leurs déplacements sur le lieu de travail ainsi que leurs interactions et qui évaluent même la qualité de leurs conversations”, ou encore à des “outils d'IA qui analysent les courriels et les messages des travailleurs pour contrôler leur productivité, repérer les travailleurs les plus innovants, mais aussi les comportements déviants”.

Ces capacités de collecte et de traitement de données par des machines, conclut-il, “supplacent celles de n'importe quel système de supervision humaine, passé ou présent”. Ainsi les données sur la santé mentale et l'état émotionnel des travailleurs sont-elles “de plus en plus accessibles grâce à l'utilisation de logiciels de reconnaissance vocale installés dans les microphones, au contrôle de l'activité cérébrale des travailleurs”. Certaines entreprises allouent des primes aux salariés acceptant de se munir de bracelets connectés permettant de surveiller leur sommeil, voire d'[évaluer leur moral](#).

Comme le note le [Centre commun de recherche de la Commission européenne](#), les algorithmes fonctionnent ainsi comme des « instruments de surveillance qui remplacent l'encadrement direct et créent des asymétries de puissance ». La captation de données relationnelles, de santé, émotionnelles, etc. pourrait conduire à mettre en œuvre des modalités très pernicieuses de surveillance et d'influence.

La réduction de ces asymétries et surtout la prévention des pratiques de surveillance abusives constituent une priorité pour les travailleurs.

(2) Le deuxième risque, déjà évoqué, porte sur de possibles **discriminations accrues**, fussent-elles [involontaires](#) voire même non désirées mais liées à l'opacité des critères effectifs de décision des algorithmes. [Amazon a ainsi dû renoncer à un programme d'algorithme de recrutement](#), ayant constaté qu'il défavorisait les candidatures féminines pour les emplois tech.

A cet égard, un risque de discrimination encore peu présent dans le débat public concerne la « [collusion algorithmique](#) », une forme de coopération implicite entre plusieurs (grands) employeurs avec comme objectif un abaissement des taux de rémunération. Ainsi, l'utilisation généralisée des algorithmes dans la rémunération des employés augmente ou crée les conditions d'abaissement des conditions de travail.

(3) Le management algorithmique va conduire un nombre croissant de travailleurs à voir leur travail être influencé voire dirigé par des processus que les travailleurs ne maîtrisent pas. La perception d'un transfert de la capacité de décision du travailleur à la machine nourrit des craintes de **perte d'autonomie**, y compris chez les travailleurs

qualifiés et les managers. Comme le souligne [Franck Pasquale](#), les cadres et les recruteurs ne savent généralement pas comment sont conçues les décisions proposées par les algorithmes ni pourquoi elles le sont, en particulier quand les entreprises utilisent des logiciels et des programmes du commerce.

Cette perte d'autonomie peut même aller jusqu'à un **sentiment d'aliénation**, lié à la perception de devenir un appendice de la machine. Sur les chaînes logistiques du e-commerce, les préparateurs de commande peuvent passer leurs journées avec un casque sur la tête leur dictant en permanence quoi faire. Dans la distribution, l'algorithme Percolata permet de réaffecter en direct et sans intervention managériale les vendeurs en fonction du flux de clients et de la performance de chaque vendeur. On assiste ainsi à une **désintermédiation du manager**.

Il importe donc que l'IA augmente les capacités humaines sans s'y substituer. Comme l'a d'ailleurs écrit le [Comité économique et social européen](#), **l'humain doit rester aux commandes**, car *“il n'est pas éthiquement acceptable qu'un être humain soit contraint par l'IA ou soit considéré comme un exécutant de la machine qui lui dicterait les tâches à accomplir”*.

Face à ces changements, les travailleurs ne sont pas inactifs, et de fait diverses stratégies d'évitement ou de résistance ont pu être observées : contournement du recours aux machines, organigramme informel, procédures alternatives, mise en place d'une socialité alternative, etc. Observant par exemple la manière dont les juges et procureurs se servaient d'algorithmes prédictifs introduits dans différents tribunaux américains, [Angèle Christin explique](#) que *“ce qui m'a frappé, c'est qu'ils ne s'en servent "surtout pas". Il y a une certaine résistance des magistrats à utiliser ces algorithmes prédictifs.”*

(4) Enfin, comme toute évolution technologique liée au monde du travail, le management algorithmique va faire évoluer le périmètre et les compétences requises d'un certain nombre d'emplois. A l'évidence, ces transformations, qui sont **d'ores et déjà à l'origine de disparités salariales**, font peser pour un certain nombre de travailleurs un **risque d'obsolescence de compétences**, en interne, et de **moindre employabilité**, en externe, qui en retour appellent des mesures spécifiques de la part des entreprises, voire des pouvoirs publics, pour les former aux différentes modalités de travail avec l'IA.

Selon le ministère du travail américain ([US Bureau of Labor](#)), les travailleurs les plus touchés devraient être les ouvriers (ouvriers sur les chaînes de montage, machinistes, caristes, ouvriers du domaine des transports...).

Mais le champ d'application de l'IA touchant l'ensemble de l'entreprise, tous les métiers sont concernés par l'obsolescence des compétences, du trader dont les ordres d'achat et de vente sont réalisés par un algorithme, au chargé de clientèle qui voit ses offres bancaires présentées par un chatbot, en passant par [l'avocat dont les recommandations s'appuient sur un algorithme](#) analysant à sa place la jurisprudence.

Selon les DRH français interrogés en 2020 par le Groupe Cegos, [47% des emplois présentent ainsi un risque d'obsolescence des compétences](#) dans les trois années à venir.

1C. Enjeux pour les entreprises et les travailleurs

Plus largement, **le management algorithmique présente 5 types d'enjeux pour les travailleurs et prestataires** : des enjeux d'**explicabilité**, d'**objectivité**, de **responsabilité**, d'**adaptation des droits des travailleurs** et enfin plus largement de **partage de la valeur**.

1. En premier lieu, comment garantir **l'explicabilité des algorithmes**, dont les règles de fonctionnement sont aujourd'hui opaques pour les travailleurs et même pour les régulateurs ?

Cette question est loin d'être théorique. Elle engage par exemple directement la **capacité de certification des régulateurs**, pour vérifier que les algorithmes respectent le cadre normatif en vigueur. Les conséquences tragiques de la validation par le régulateur américain du logiciel MCAS d'aide au pilotage des Boeing 737 Max, sans en réalité le comprendre, faute d'avoir l'expertise requise -dont seul Boeing disposait- montre l'ampleur et l'actualité de ces enjeux.

2. Ensuite viennent les **questions sur l'objectivité des algorithmes et les risques de biais et de discriminations**, qui sont d'ailleurs celles qui viennent le plus souvent à l'esprit en matière d'IA au travail. A titre illustratif, le tribunal de Bologne vient de [juger discriminatoire l'algorithme de Deliveroo](#), qui utilisait un indice de "fiabilité" pour classer les livreurs et ainsi prioriser l'attribution des courses.

L'algorithme de Deliveroo ne distinguait pas selon les causes d'annulation des quarts de travail et discriminait donc potentiellement les travailleurs malades ou en grève. Cette décision rappelle qu'il ne faut pas attendre d'un algorithme qu'il ne discrimine pas - c'est son objet même. L'exigence d'objectivité des algorithmes doit donc être replacée dans son contexte. Cependant l'algorithme doit discriminer de façon *juste*, soit en connaissance de tous les paramètres et de toutes les causes possibles d'un phénomène.

La décision du tribunal de Bologne rappelle donc non seulement qu'un algorithme peut être discriminatoire mais également que son set d'apprentissage doit être exhaustif et lui-même non biaisé.

Le risque n'est pas spécifique à l'employé: un algorithme de recrutement mal spécifié peut faire passer l'entreprise à des candidats potentiels. Ainsi, une [étude récente](#) montre dans quel mesure, l'entreprise se fiant uniquement à un recrutement algorithmique peut se retrouver de plus en plus désavantagée sur un marché de travail hyper-compétitif et en manque de talent.

Ceci pose en retour la question des standards éthiques que l'écosystème entreprise - fournisseurs - travailleurs doit mettre en place : comment garantir l'objectivité des ingénieurs - fournisseurs qui ont créé l'algorithme ? Comment éviter que les données alimentant le "machine-learning" soient source de biais ? Quelles directives éthiques ont été appliquées au sein du fournisseur ?

3. Les questions de **responsabilité** sont tout aussi importantes.

Le traitement algorithmique dissimule dans une « boîte noire » un ensemble de décisions et d'instructions prises en amont. En effet, comme le souligne [Antonio Casilli](#), les algorithmes ne sont pas des processus de calculs neutres, rationnels, objectifs, mais des instructions issues d'une programmation faite avec des objectifs précis, se basant sur un certain nombre de principes. Or l'algorithme floute l'identité du donneur d'ordre, et ce d'autant plus en substituant la suggestion ("*nudge*") à l'injonction.

Ces formes d'influence posent de véritables défis au droit : ils ne relèvent pas de l'exercice classique de l'autorité, contractualisée. Reprenant ce que Michel Foucault a nommé gouvernementalité (néologisme associant gouvernement et rationalité), Antoinette Rouvroy utilise le terme de gouvernementalité algorithmique pour désigner cette normativité nouvelle, qui entoure chaque individu d'une bulle individualisée d'influences et d'incitations (on contrôle, mais en douceur, en orientant les comportements par des "coups de pouce"). Notamment en cas de litige, il devient alors plus commode de se cacher derrière une décision algorithmique plutôt que de développer un argumentaire alternatif.

Au total, qui est responsable du fonctionnement des algorithmes et de leurs résultats ? Le codeur ? Le manager ? Le travailleur qui agit en fonction de ce que lui dit l'algorithme (mais qui garde la possibilité de ne pas le suivre) ? Quelle que soit son autonomie, l'algorithme ne peut être responsable à la place de l'homme. Que reste-t-il de la responsabilité humaine, et comment l'attribuer ?

Plus spécifiquement, **quatre questions nous semblent essentielles** :

- Comment permettre aux entreprises de garder le contrôle des algorithmes qu’elles utilisent, notamment pour éviter les biais involontaires introduits par les algorithmes (comme cela a de fait été le cas dans [l’exemple d’Amazon déjà cité](#) où l’algorithme d’aide au recrutement défavorisait les femmes).
- Comment gérer et accompagner la redéfinition des processus de prise de décision au sein de l’entreprise, notamment face au risque de désintermédiation et de déresponsabilisation des managers ?
- Plus largement, comment garantir, dans un contexte d’amélioration continue de la performance et sans doute de la place des algorithmes, que l’humain reste aux commandes et que chaque travailleur préserve une marge d’autonomie par rapport aux machines ?
- Quelles sont les responsabilités des employeurs à l’égard des travailleurs en matière de management algorithmique (par exemple pour les accompagner face aux nouveaux usages, maintenir leur employabilité) ?

Plus largement, **les algorithmes ne vont-ils pas contribuer à redéfinir les contours même des entreprises ?** D’ores et déjà, les plateformes numériques, à l’image de Uber, s’appuient sur les algorithmes pour constituer des “[modèles d’hybridation marché-entreprise](#)”. Comme le notent O. Chagny et Florian Forestier, en utilisant les algorithmes pour remplacer les mécanismes de coordination usuels au sein des firmes, et en réduisant ainsi la raison d’être des firmes comme [lieu de réduction des coûts de transaction](#), les plateformes substituent une relation commerciale à une relation de travail, un contrat de prestation de service à un contrat de travail.

4. Le quatrième type d’enjeux concerne **l’adaptation des droits des travailleurs (et le maintien de leur employabilité)**.

Le droit du travail s’est construit historiquement en partie pour rééquilibrer les droits des travailleurs face aux asymétries de pouvoir avec leurs employeurs. Or il ne prévoit aujourd’hui à peu près rien sur la régulation des algorithmes, qui pourtant introduit de nouvelles asymétries avec les travailleurs, notamment d’information et de capacité de décision.

Par exemple, les outils numériques de surveillance et les algorithmes offrant une transparence quasi-totale aux employeurs sur les travailleurs et leurs pratiques, comment encadrer la collecte et l’usage de ces données pour éviter une surveillance abusive (sachant qu’il ne s’agit pas seulement de données personnelles) ?

Comment assurer pour les travailleurs face au contrôle de leur action par les algorithmes un niveau de garanties équivalent à celui dont ils disposent aujourd’hui en matière de contrôle managérial par les employeurs (dont ils sont informés des

modalités et encadré notamment par les droits fondamentaux et libertés individuelles des salariés) ?

Quelles mesures concrètes prendre pour éviter les autres risques évoqués pour les travailleurs : discriminations, risques psycho-sociaux, risques de perte d'autonomie par les travailleurs, obsolescence des compétences et perte d'employabilité ?

Comment plus largement rééquilibrer les rapports entre le travailleur et le donneur d'ordre face au management algorithmique ? Comment garantir l'accès effectif à ces droits (le RGPD notamment en offrant déjà, mais très complexes à mettre en œuvre, ce qui les rend *de facto* en partie inopérants) ?

5. *Last but not least*, le **partage de la valeur** constitue un enjeu essentiel. Le fait que les machines soient désormais capables d'apprentissage pourrait donner lieu à une captation de l'expertise humaine non seulement accélérée mais déséquilibrée dans la rétribution de la valeur créée. Les IA qui modélisent une expertise (médicale, juridique, etc.) le font à partir de données produites par du travail réel. Quelle part de la valeur économique créée est redistribuée aux travailleurs concernés ? Par exemple, les plateformes VTC qui tentent d'évoluer vers des modèles avec véhicules autonomes reverseront-elles une part de leurs bénéfices aux chauffeurs qui auront alimenté les IA à l'origine de ces véhicules autonomes ?

Ces questions devraient conduire à mieux évaluer ces créations de valeur, et à réfléchir à des mécanismes renouvelés de redistribution (envers des individus ou des professions ubérisées par exemple).

II. Principes directeurs pour promouvoir un management algorithmique responsable

Alors que le droit du travail est encore très centré sur l'organisation des relations de travail pré-numérique, prendre la mesure des enjeux posés par le management algorithmique sur le travail et agir dès maintenant pour l'encadrer est absolument critique.

On l'a vu, le management algorithmique constitue, pour les travailleurs, une source à la fois d'opportunités et de menaces. Il a le potentiel de rendre l'activité des travailleurs plus productive, moins soumise à l'arbitraire, et plus intéressante. Il ouvre à l'inverse aussi pour eux des risques de surveillance abusive, de discriminations, de perte d'autonomie voire de risques psycho-sociaux accrus, et de perte d'employabilité.

Or, comme le souligne [Jeremias Adams-Prassl](#), il ne faut pas tomber dans le piège du déterminisme technologique : aucune de ces tendances n'est inexorable et l'histoire a montré combien la régulation par les pouvoirs publics et les pratiques des entreprises peuvent orienter et encadrer les évolutions technologiques.

Comme l'écrit d'ailleurs le Parlement européen en 2019 dans sa résolution sur une [politique industrielle européenne globale sur l'intelligence artificielle et la robotique](#) : *“les évolutions dans le domaine de l'IA peuvent et devraient être conçues de façon à préserver la dignité, l'autonomie et l'auto-détermination des individus”*.

L'enjeu n'est autre que de **reprendre le contrôle de l'IA au travail : promouvoir un management algorithmique responsable** pour exploiter son potentiel tout en garantissant des conditions de travail qui préservent la dignité, la santé, l'égalité de traitement et l'autonomie des travailleurs.

Or, si l'on veut exploiter son potentiel tout en garantissant des conditions de travail qui préservent la dignité, la santé, l'égalité de traitement et l'autonomie des travailleurs, il est essentiel de définir en quoi consisterait un management algorithmique responsable. Nous proposons à cette fin plusieurs **principes directeurs**, inspirés notamment par ceux du [BIT](#), de [l'OCDE](#) et des [TUC](#).

Nous avons choisi de proposer des principes directeurs pour les trois grands types d'acteurs concernés par le management algorithmique : les entreprises, les travailleurs et les pouvoirs publics, en commençant par les entreprises car ce sont elles qui les mettent en oeuvre et avec, pour elle, une réflexion à la fois sur leurs droits et sur leurs devoirs en la matière.

2A. 10 principes directeurs relatifs aux entreprises

L'usage de l'IA par les entreprises pour organiser le travail humain relève, d'abord et avant tout, de la liberté d'entreprendre : c'est un moyen pour elles d'améliorer leur efficacité et de mieux remplir les missions qui sont les leurs.

Avant de parler de leurs responsabilités en la matière, et des contraintes que les pouvoirs publics pourraient leur imposer, il importe donc de rappeler leurs **droits**, et le principe général que ces contraintes qui pourraient peser sur leur liberté d'entreprendre doivent être strictement nécessaires et proportionnées au but qu'elles poursuivent (comme la protection des droits et libertés fondamentales des personnes).

Il nous semble ainsi que les entreprises ont droit à ce que le cadre réglementaire relatif au management algorithmique :

1. Respecte la **propriété intellectuelle**, par exemple en matière de "transparence" des algorithmes, ceux-ci étant souvent au coeur de leur propriété intellectuelle,
2. Garantisse un **jeu concurrentiel équitable** entre les acteurs,
3. Préserve la capacité d'**innovation**, en n'imposant que des limites strictement nécessaires, par exemple fondées sur les droits et libertés fondamentales des personnes
4. garantisse leur **sécurité juridique**, malgré la nécessaire évolutivité de ce cadre compte tenu de la rapidité des évolutions technologiques sous-tendant le management algorithmique, notamment en étant assez clair pour éviter les incertitudes techniques et juridiques et en étant facilement opposable aux fournisseurs.

Les enjeux très particuliers liés au management algorithmique font peser sur les entreprises des **responsabilités** spécifiques en la matière :

5. **garantir les droits de leurs travailleurs** (*cf. infra. Principes directeurs pour les travailleurs*), dans les procédures internes et en amont dans la relation avec les fournisseurs,
6. **informer les travailleurs** sur leurs pratiques en matière d'usages de l'IA pour organiser le travail, en particulier sur la nature des données personnelles collectées, l'usage qui en est fait, les critères et paramètres de fonctionnement des algorithmes, les modalités de contrôle humain sur le fonctionnement des algorithmes,
7. **former leurs travailleurs** aux pratiques de management algorithmique,
8. **investir dans l'employabilité des travailleurs** si celle-ci est remise en cause par l'IA,
9. **faire de l'IA et de ses usages au travail un thème de dialogue social**,
10. **se prêter à des audits externes** pour vérifier le respect des points ci-dessus

2B. 10 principes directeurs relatifs aux travailleurs

Pour les travailleurs, déjà “protégés” par le droit du travail, il est logique que les principes directeurs relatifs au management algorithmique soient liés à la prévention des risques spécifiques posés par celui-ci (surveillance abusive, discriminations, perte d'autonomie, perte d'employabilité).

Nous proposons ainsi **10 principes directeurs pour les travailleurs**, relatifs à leur droit à la dignité, leur droit à l'égalité de traitement, leur droit à la santé et leur droit à l'autonomie.

Préserver la **dignité** des travailleurs

1. **Droit à la vie privée** : les algorithmes ne doivent pas donner lieu à des abus dans la surveillance des travailleurs ou la collecte des données les concernant. Leur mise en œuvre n'étant pas apparente pour le travailleur dans ses tâches quotidiennes, les algorithmes permettent plus aisément, à dessein ou non, des atteintes aux libertés individuelles des travailleurs et contribuent à effacer la frontière déjà mouvante entre vie professionnelle et vie privée.
2. **Droit d'information et de consentement** : les travailleurs doivent être informés de la nature des données et informations collectées sur eux et des usages qui en sont faits et les données collectées doivent l'être avec leur consentement éclairé,
3. **Droit aux interactions humaines** : les travailleurs ne doivent pas devoir travailler uniquement avec des algorithmes mais doivent pouvoir avoir des interactions humaines dans leur travail.

Préserver l'**égalité de traitement** des travailleurs

4. **Droit à la non-discrimination** : les algorithmes au travail ne doivent pas introduire de discriminations parmi les travailleurs (dans l'entreprise mais aussi à l'embauche)

Préserver la **santé et le bien-être psychologique** des travailleurs

5. **Droit à la santé** : les algorithmes ne doivent pas être source de risques de santé notamment psycho-sociaux

Préserver l'**autonomie** des travailleurs

6. **Droit à la compréhension (explicabilité des algorithmes)** : les travailleurs doivent non seulement être informés qu'ils travaillent avec un algorithme, mais comprendre comment cette interaction fonctionne (principe d'intelligibilité des algorithmes),

7. **Droit à la décision humaine** : les travailleurs ont droit à ce que les décisions qui les concernent soient prises par des individus,
8. les travailleurs doivent pouvoir contester ou faire appel de toute décision prise sur le fondement d'un algorithme,
9. **Droit à la formation à l'usage de l'IA** : les travailleurs ont le droit d'être formés à l'usage des algorithmes,
10. **Droit de négociation** : les travailleurs ont le droit, via leurs représentants, à être informés et à négocier les conditions d'usage des algorithmes au travail dans l'entreprise

2C. 4 principes directeurs relatifs aux pouvoirs publics

A l'évidence, le management algorithmique soulève des enjeux qui appellent une intervention spécifique de la part des pouvoirs publics. Le cadre réglementaire actuel n'est pas suffisant, il doit être adapté à ces nouveaux enjeux.

Les pouvoirs publics ont ainsi une quadruple responsabilité en matière de management algorithmique :

1. **Profondeur et autonomie de l'expertise** : les pouvoirs publics doivent se donner les moyens de produire une **réglementation pertinente** sur l'IA au travail, équilibrée entre le respect de la liberté d'entreprendre et la protection des droits des travailleurs. Cette considération pourrait paraître évidente mais la nouveauté radicale, la technicité et la rapidité d'évolution du sujet imposent aux pouvoirs publics un investissement spécifique pour disposer d'une compréhension fine des réalités et des spécificités du management algorithmique et d'une expertise véritablement indépendante des acteurs.
2. **Accès effectif aux droits** : compte tenu de la très grande complexité du sujet, voire de son opacité et du risque que les abus éventuels puissent être largement invisibles pour les acteurs eux-mêmes (par exemple en matière de surveillance abusive ou de discrimination), promouvoir une réglementation pertinente ne suffira pas, les pouvoirs publics doivent veiller à ce que les **règles édictées soient bien respectées et les droits garantis soient effectivement accessibles**, en particulier pour les travailleurs individuels,
3. **Soutien à l'employabilité** - dès lors que l'usage croissant de l'IA au travail va avoir des effets considérables sur les conditions de travail des travailleurs, et notamment leurs compétences requises, avec un risque de "décrochage" ou d'obsolescence de compétences à une échelle non seulement individuelle mais collective, il appartient aux pouvoirs publics de **promouvoir des dispositifs de sensibilisation et de formation** pour, en complément des actions menées par les entreprises, doter les individus des compétences dont ils ont

besoin dans le domaine de l'IA et soutenir ainsi leur employabilité dans la durée,

4. **Adaptabilité du cadre réglementaire** - la nature et les usages de l'IA au travail se transforment tout le temps, et très vite, dès lors son cadre de régulation ne doit pas être défini ni figé une fois pour toutes mais doit être très évolutif, au risque sinon de perdre toute pertinence, avec des clauses de revoyure plus fréquentes pour adapter la loi et le cadre réglementaire, et une ouverture plus grande à des expérimentations.

III. 20 propositions concrètes pour promouvoir un management algorithmique responsable

Comment concrètement faire en sorte que ces principes directeurs soient appliqués ?

Nous proposons de donner une véritable portée juridique à ces principes en les ajoutant parmi les considérants au projet de Règlement européen sur l'IA, en leur donnant ainsi un statut de "préambule" à ce Règlement. Nous proposons également de les inclure aux [principes directeurs déjà définis par l'OCDE en matière d'IA](#) et adoptés par ses États membres.

Au-delà, pour leur donner une assise juridique encore plus solide nous proposons de les décliner autour de **quatre axes de recommandations concrètes**, toutes pleinement compatibles avec la promotion de l'innovation, le respect de la propriété intellectuelle et des règles de concurrence et le respect de la sécurité juridique que les acteurs économiques sont légitimement en droit d'attendre.

- I. **PRÉVENIR - Adopter des règles de conception sécurisant une IA au travail responsable**
- II. **PROTÉGER - Adapter les droits individuels des travailleurs aux transformations du travail portées par les algorithmes**
- III. **GARANTIR - Garantir le respect effectif des principes de l'IA responsable et des droits des travailleurs**
- IV. **ACTUALISER - Permettre l'adaptation du cadre réglementaire dans le temps aux évolutions de l'IA**

*

AXE 1 : PRÉVENIR - Adopter des règles de conception sécurisant une IA au travail responsable

Proposition #1 - Introduire un objectif d'explicabilité et d'éthique "by design" des algorithmes

Nous proposons de **consacrer un objectif d'explicabilité du code ou de l'algorithme dans le droit européen, via le projet de Règlement sur l'IA.**

Plusieurs arguments motivent cette proposition :

1. **L'explicabilité de l'algorithme est indispensable pour que les travailleurs et prestataires puissent s'engager contractuellement en connaissance de cause.** L'usage de l'algorithme, en tant qu'instrument contractuel indirect, est en effet créateur de droits et d'obligations. Or les règles de fonctionnement des algorithmes sont aujourd'hui opaques pour les travailleurs, voire pour les employeurs. Il n'est pas rare que seul l'auteur du code, prestataire externe à l'entreprise, soit capable de l'expliquer en intégralité. Sans transparence ni lisibilité, la décision issue du management algorithmique ne peut être motivée pour son destinataire : elle est dès lors... arbitraire.
2. **L'explicabilité est un prérequis indispensable pour circonscrire les risques de biais, de discrimination, ou d'autres atteintes aux droits des travailleurs et prestataires,** faire du contenu de l'algorithme un objet de négociation et permettre l'audit de l'algorithme par un tiers et notamment un juge en cas de litige ;
3. Il y a déjà un précédent juridique majeur, le [Conseil Constitutionnel](#) ayant reconnu valeur constitutionnelle à l'objectif d'accessibilité et d'intelligibilité de la loi. Cet objectif d'intelligibilité constitue désormais une nouvelle composante de la sécurité juridique. Or, si nous considérons, au sens figuré, que "[Code is law](#)", il est légitime, similairement, de consacrer un objectif d'intelligibilité de l'algorithme
4. Inversement, cet objectif d'explicabilité ne conduira ni à nuire à la promotion de l'innovation, ni aux droits de propriété intellectuelle attachés aux algorithmes.

Un algorithme se doit d'être interprétable, et un résultat explicable. Seule cette intelligibilité, érigée en principe dès la conception de l'algorithme, permettra aux parties - travailleurs, employeurs ou donneurs d'ordre- d'accéder aux motifs d'une décision pour la comprendre et la corriger le cas échéant, ce qui préservera la responsabilité de chacun.

Certes, le Règlement européen [Platform to Business](#) (P2B) introduit déjà la notion de transparence dans les relations entre travailleurs et plateformes, mais cette transparence est limitée aux plateformes et n'est envisagée qu'en aval de l'algorithme, dans les effets de sa mise en œuvre. Or il faudrait intervenir en amont, par **l'introduction d'un impératif d'explicabilité de l'algorithme à mettre en œuvre dès sa conception dans le Règlement européen sur l'IA.**

Cet objectif d'explicabilité doit se traduire par **des critères d'accessibilité, d'auditabilité, voire d'association de différentes parties prenantes à la conception des algorithmes** pour encadrer à bon escient leur élaboration et devenir ainsi "*ethical by design*".

Le Règlement européen sur l'IA pourrait ordonner *a minima* **l'explicabilité du code objet** -et non du code source- de l'algorithme, de même que la **communication sur demande** de tout travailleur ou prestataire concerné par sa mise en œuvre, d'un **document intelligible et opposable à son émetteur**, exposant de manière claire et complète les conditions et les effets de la mise en œuvre de l'algorithme sur le management dont il est l'objet.

Cet objectif d'explicabilité doit être étendu aux résultats des tests et audits des algorithmes.

L'article 13 du projet de Règlement européen dispose :

“La conception et le développement des systèmes d'IA à haut risque sont tels que le fonctionnement de ces systèmes est suffisamment transparent pour permettre aux utilisateurs d'interpréter les résultats du système et de l'utiliser de manière appropriée. Un type et un niveau adéquats de transparence permettent de veiller au respect des obligations pertinentes incombant à l'utilisateur et au fournisseur énoncées au chapitre 3 du présent titre.”

La transparence n'y est donc définie que par certaines finalités (permettre aux utilisateurs d'interpréter les résultats du système) et son champ d'application n'est pas circonscrit. Par ailleurs, seul l'employeur ou le donneur d'ordre peut bénéficier de cette transparence en l'état actuel du texte.

Il est donc loisible de préciser les dispositions de cet article 13 afin notamment de :

- introduire la notion d'explicabilité de l'algorithme, plus précise que la notion de transparence qui est envisagée dans le projet de règlement à différents degrés (ainsi le simple fait d'informer les personnes de la mise en œuvre d'un algorithme est, selon certaines dispositions, une transparence suffisante),
- entériner le fait que cette explicabilité doit bénéficier non seulement à l'utilisateur de l'algorithme mais également à la personne qui fait l'objet d'une décision dictée ou assistée par l'algorithme,
- préciser le champ d'application de cet objectif d'explicabilité, qui doit intervenir dès la conception de l'algorithme et en aval, pour assurer la lisibilité des résultats d'audits et de tests.

Proposition #2 - Assurer la traçabilité des décisions algorithmiques

La mise en œuvre de l'algorithme dans le contexte managérial impacte le devenir d'une relation contractuelle.

En effet, l'algorithme va tantôt intervenir comme aide à la décision managériale, tantôt prendre la décision seul (*de facto*). Son intervention doit pouvoir être connue du travailleur, comme doit l'être toute décision managériale impactant l'exécution du contrat (en matière d'organisation du travail par exemple).

Il est essentiel d'assurer la traçabilité des décisions algorithmiques affectant l'exécution du contrat et ce, pour deux motifs :

Tout d'abord, un travailleur ou prestataire dont le contrat est affecté par la mise en oeuvre d'un algorithme doit non seulement en être informé mais **il doit également savoir** en détail, à première demande, **quelle décision a été prise par ou au moyen de l'algorithme le concernant** : nature, sens, date, heure, impact sur l'exécution de son contrat.

Ensuite, **le travailleur doit être en mesure de questionner cette décision algorithmique** dans sa motivation, son bien-fondé, et de connaître avec précision la place de la décision algorithmique dans la décision finalement adoptée (aide à la décision pour le supérieur ou donneur d'ordre, ou décision autonome de l'algorithme).

Seule cette traçabilité permettra de retracer la réalité d'un biais, d'une discrimination et éventuellement d'en retrouver la source (biais purement humain, code de l'algorithme lui-même ou plus fréquemment, faiblesse dans la sélection des données d'apprentissage du modèle algorithmique).

La traçabilité apparaît donc, dans ce cadre contractuel, comme l'une des **protections du travailleur ou du prestataire contre l'arbitraire** et un impératif pour responsabiliser l'employeur ou le donneur d'ordre qui met en oeuvre l'algorithme.

Cette traçabilité, déjà envisagée *a minima* par le projet de Règlement européen en son article 20 ("Journaux générés automatiquement"), doit permettre non seulement la création systématique du journal ou log, relatif aux opérations de l'algorithme mais également - et de manière effective, l'intelligibilité et l'accessibilité de ces journaux pour les travailleurs ou prestataires ayant fait l'objet de décisions algorithmiques ou aidées par l'algorithme.

Enfin, la décision algorithmique trouvant sa source tant dans l'objectif poursuivi par le donneur d'ordre, que dans le travail algorithmique d'analyse et parfois d'apprentissage et de comparaison, cette traçabilité ne doit pas se limiter à la décision elle-même. Une telle limite ne permettrait pas en effet de conserver la trace de la motivation de la décision - indépendamment de sa lisibilité.

L'article 20 du projet de règlement européen pourrait donc être complété pour préciser les activités de l'algorithmes soumises à la génération automatique d'un journal et pour

affirmer la finalité du journal ainsi produit : assurer la traçabilité des décisions et du travail algorithmique pour préserver les droits des travailleurs dont l'activité en est impactée.

La piste d'une utilisation des blockchains pour opérationnaliser cette traçabilité mérite d'être étudiée, même si elle pose bien sûr plusieurs questions importantes, dont celle de son articulation avec le droit à l'oubli garanti par le RGPD.

Proposition #3 - Rendre obligatoires certaines séquences au sein des algorithmes pour garantir les droits des travailleurs et inversement interdire des séquences aux effets délétères

Le management algorithmique affecte des relations avant tout contractuelles (entre des donneurs d'ordres et des prestataires ou entre des employeurs et des salariés). Il devrait donc pouvoir faire l'objet du même encadrement législatif et du même contrôle judiciaire que le contrat dont il affecte l'exécution.

C'est d'autant plus important qu'une [étude du cabinet de conseil BCG](#) montre que moins de la moitié des entreprises qui ont déployé l'IA dans leur organisation l'ont fait avec une IA responsable.

- Tout comme il existe des mentions obligatoires dans certains types de contrats de travail, il est possible de prévoir des **blocs de code à caractère impératif**. Ils auraient vocation à entériner certaines garanties pour les travailleurs : la consécration législative de droits, une information plus complète, le caractère effectif et systématique de l'accès à la donnée, etc.
- A l'inverse le pouvoir judiciaire peut écarter l'application de certaines clauses et les juger réputées non écrites, telles que les "clauses léonines". De la même manière, **certains blocs de code aux effets délétères devraient pouvoir être soumis au juge pour voir leur application écartée de la relation de travail** (s'ils ne sont pas déjà exclus en amont par le double effet de l'encadrement normatif et de la négociation collective).

Proposition #4 - Encadrer la conception des algorithmes, notamment par un contrôle en amont des bases de données d'apprentissage pour garantir leur diversité et minimiser les biais des algorithmes

Les biais des algorithmes auto-apprenants peuvent provenir directement de la base de données utilisée lors de la phase d'apprentissage des algorithmes.

En effet, le travail de préparation de données est très souvent négligé au profit de celui de la génération des modèles. Or cette phase de construction des bases de données d'apprentissage est totalement déterminante pour garantir la qualité du modèle généré et donc minimiser les biais algorithmiques.

Deux points d'attention sont donc essentiels:

- Le premier s'inscrit dans la phase de **mise en qualité des données** afin de les rendre exploitables par les algorithmes. Un contrôle de qualité doit donc être nécessaire avec un accès aux Métadonnées afin de connaître les caractéristiques des données d'apprentissage (types, provenance, statistiques ...) pour mesurer les niveaux d'applicabilité des modèles entraînés.
- Le deuxième concerne l'**enrichissement des bases de données**. On peut parler (i) d'enrichissement simple avec l'ajout de données publiques externes et (ii) d'enrichissement complexe qui lui provient d'une génération plus ou moins automatisée de nouvelles informations par croisement des données initiales. Cette phase d'enrichissement de données a pour but d'augmenter la masse d'information disponible pour l'entraînement des algorithmes et ainsi pouvoir minimiser les biais d'apprentissage.

Au même titre qu'un algorithme doit être explicité sur demande, par une équipe d'experts de la data (type data scientist ou data engineer), les Métadonnées des bases de données d'apprentissage ainsi que les caractéristiques et conditions d'enrichissement de ces mêmes bases doivent également faire l'objet d'une explicitation sur demande par une équipe d'experts de la donnée.

Proposition #5 - Prévoir une mutualisation des données d'apprentissage auprès d'institutions tierces, relevant de l'Etat, ou prenant la forme de fiduciaires de données ('data trusts') pour minimiser les biais

Pour limiter les biais liés aux bases de données, il faut enrichir celles-ci, avec une quantité suffisante de données, bien décrites et étalonnées, suffisamment diversifiées. [L'Institut Montaigne](#) dans son rapport "Algorithmes : contrôle des biais S.V.P" mentionne à ce sujet la question emblématique de la reconnaissance faciale, pour laquelle les développeurs ne disposent pas d'une base de données suffisamment représentative de la population française pour tester l'absence de biais de leur algorithme.

Cet élargissement de la collecte des données ne se fera cependant qu'à condition d'instaurer un cadre permettant la confiance.

L'Institut Montaigne propose pour sa part (c'est la proposition 5 du rapport "Algorithmes : contrôle des biais S.V.P") la mise en place d'une ou plusieurs bases de données publiques comportant des variables protégées au titre des 25 critères de discrimination. Celles-ci seraient utilisées pour tester la méthodologie, avant ou après entraînement sur les données propres. Selon l'Institut Montaigne, le Laboratoire national de métrologie et d'essais (LNE) pourrait assumer cette responsabilité. Ce laboratoire intervient en effet *"déjà dans des domaines d'application de l'IA tels que les véhicules autonomes, les dispositifs médicaux intelligents (notamment d'aide au diagnostic), les robots industriels collaboratifs ou les drones. Il est responsable d'un dé destiné à évaluer des robots agricoles, et a déjà noué des partenariats pour développer des méthodes d'évaluation des algorithmes de transcription automatique de la parole."*

Une solution alternative que nous privilégions du fait de sa plus grande capacité à assurer un cadre de confiance serait de [développer des "data trusts" ou "fiduciaires de données"](#) dédiées aux outils de management algorithmiques. Ces "data trust" sont des intermédiaires indépendants ayant l'obligation fiduciaire d'agir au nom des personnes concernées dont elles collectent et gèrent les données individuelles. Ils présentent un double intérêt. 1) Simplifier et rendre plus transparentes la collecte et l'utilisation des données personnelles des utilisateurs, 2) permettre aux acteurs de bénéficier de pools de données diversifiées, tout en garantissant qu'elles soient utilisées de façon conforme aux volontés des parties-prenantes, et 3) compléter, en cas de besoin, via des approches de [données synthétiques](#), d'éventuelles sous-représentations d'une partie de l'échantillon qui permettront de préserver la protection de données personnelles.

L'Open Data Institute (ODI), institut de recherche travaillant entre autres sur la question du data trust, [propose six piliers](#) autour desquels organiser un tel organisme : un but précis, un processus décisionnel défini en amont de sa création, un modèle de partage des bénéfices, une véritable structure juridique, des droits aux utilisateurs sur les données qu'ils partagent, et un financement durable. [Un data trust peut prendre différentes formes](#), de l'entreprise à l'association à but non-lucratif, en passant par l'autorité administrative, dès lors que sa structure est définie en amont.

Les data trust sont actuellement mobilisés dans le cadre de politiques sectorielles (dans le domaine de la santé, l'exemple le plus connu est le Health Data Hub) ou territoriales, en particulier pour les projets de "smart cities" ou de [territoires d'intelligence](#)", selon la formule de Luc Bélot, par exemple pour coordonner les nouvelles mobilités multimodales intelligentes.

Nous proposons de créer de telles structures pour les données utilisées pour les solutions de management algorithmique. Sous réserve qu'elles atteignent une taille critique, celles-ci permettraient de :

- réduire les risques de biais et rendre les algorithmes plus sensibles aux faits, grâce à la [multiplication et à la diversification les modes d'apprentissage des algorithmes](#) ;
- proposer un modèle pouvant jouer un rôle de filtre pour des contenus générés par les modèles linguistiques utilisés pour entraîner à large échelle les algorithmes à moindre coût pour les opérateurs qui vont les exploiter, par exemple pour les PME, tout en offrant, grâce à l'intervention de l'acteur tiers, des garanties quant à leur qualité et à la prévention des biais qu'ils pourraient incorporer.

Proposition #6 - Placer certains types de données sous le statut de “données d'intérêt général”

La possibilité d'accéder à une grande quantité de données (sourcing, tri des CVs, entretiens et évaluation) constitue également une barrière à l'entrée importante pour la conception de solution, confortant la position de quelques entreprises et mettant celles-ci en position de force, les laissant imposer des standards sans que les entreprises clientes et les autres parties-prenantes disposent d'un pouvoir de décision ou d'une capacité de contrôle suffisantes.

Cette situation n'est évidemment pas limitée aux questions de management et concerne de nombreux domaines, comme l'énergie, l'eau, ou encore les mobilités. Dans ces domaines toutefois, de nombreux jalons ont déjà été posés afin de construire un cadre permettant de sécuriser l'accès des acteurs publics à ces données. Ainsi, la loi pour une République numérique de 2016, a rendu obligatoire la mise à disposition des données « dont la publication présente un intérêt économique, social, sanitaire ou environnemental ». Elle oblige à ce titre les délégataires des services publics qui récoltent les données de la gestion de l'eau, de l'énergie ou des déchets à ouvrir l'accès à ces données.

Un élargissement des possibilités de collecte de « données privées d'intérêt général » est par ailleurs discuté, donnant aux acteurs publics les moyens de l'appliquer, par exemple en conditionnant la délivrance d'autorisations d'activité à l'obligation de transmission de certaines données. Pour l'instant, l'attention se concentre sur les données des plateformes comme waze, Uber ou AirBnB, indispensables pour la transition numérique des territoires. En 2017, le rapport Belot a proposé un statut de «données d'intérêt territorial» lequel permettrait d'engager au niveau local «un

recensement des données d'intérêt territorial disponibles», produites par les entreprises ou associations.

Nous proposons d'inscrire certaines données concernant les pratiques managériales au sein de cette réflexion.

D'abord, en identifiant précisément les types de données concernées.

Puis, en proposant des modalités facilitant leur mise à disposition et leur partage.

- Soit en obligeant les acteurs du secteur privé à pratiquer l'*open data* concernant ces données (à l'image de ce qui a été prévu sur la mise à disposition de données locales d'énergie par la loi de transition énergétique).
- Soit en permettant à une instance tierce d'utiliser les données tout en laissant leur traitement aux acteurs privés¹.
- Soit enfin, de façon plus ambitieuse, en prolongeant jusqu'au bout la logique des "données d'intérêt général" en inventoriant l'ensemble des données relevant de cette problématique et en donnant à celles-ci le statut de "communs". Le statut de la donnée comme «commun» est défini à partir d'un faisceau de droits (bundle of rights) décomposant («démembrant» au sens juridique du terme) la propriété des données en un ensemble de droits distincts.

Proposition #7 - Mettre en place un encadrement spécifique du Cycle de Vie des Modèles pour minimiser les biais et le risque de dérive en qualité de modèles au cours du temps

De la même manière que les méthodes de gouvernance des données visent à s'assurer du maintien de la qualité des données lors de leur ingestion et de leur redistribution, le cycle de vie des modèles s'assure que la qualité des modèles reste stable au cours du temps. En effet, un modèle en production est "branché" sur des données en perpétuelle évolution qui sont émises par exemple quotidiennement.

Ainsi le modèle prédictif est effectivement généré à partir de données passées (contrôlées, identifiées, validées ...). Or il se nourrit de données récentes pour propager une information vers le futur. Ainsi l'hypothèse selon laquelle les données sont

¹ « Engager un rééquilibrage privé/public par les données », *Cahiers Innovation & Prospective* no 05, Commission nationale de l'informatique et des libertés, 2017.

indépendantes et identiquement distribuées lors de la phase d'apprentissage devient erronée dans le monde réel. Par exemple, pensons à un système de détection de fraude où l'adaptation des fraudeurs rend caduque le modèle construit. On parle de dérive de modèle alors qu'en réalité ce sont les données elles même qui ont changé. Il devient alors essentiel de bien construire et contrôler le cycle de vie des modèles.

Le Cycle de vie des modèles englobe donc les évènements qui interviennent au cours de l'existence d'un modèle, c'est-à-dire dès lors qu'il est mis en service.

La première phase réside dans la conduite du changement définie lors de l'insertion du modèle dans le processus existant en définissant par exemple des droits d'accès, d'utilisation (etc ...). Deuxièmement, le cycle de vie des données assure que les données servies au modèle obéissent bien aux standards définis lors de sa phase d'apprentissage.

Ce n'est que lors de la troisième étape - celle du monitoring des modèles - que la qualité du modèle est surveillée et contrôlée et ce en termes de bon fonctionnement, performance, retour d'expérience à l'usage ...

Dans un projet impliquant une IA on peut distinguer trois niveaux d'intervention sur le cycle de vies des modèles:

- Organisation métier - qui garantit l'impact métier d'un modèle en veillant à l'applicabilité du modèle au cours du temps.
- Scientifique - pour évaluer la robustesse des modèles à l'efficacité d'un système de détection de dérive et aux différentes solutions de redressement de la performance: ré-entraînement des modèles.
- Technologique - gestion de l'architecture du système en production, assurer l'interface entre cycle de vie des données et modèle et garantir le déploiement rapide de solutions préconisées par le niveau scientifique.

Nous recommandons donc de mettre en place une équipe d'experts responsables du Cycle de Vie des modèles pour s'assurer de la non-dérive des modèle dans le temps :

- responsable qualité des données - Assurer la pérennité du référentiel de données
- responsable robustesse du modèle - Contrôler non seulement la performance du modèle mais également sa robustesse en suivant son incertitude : Monitoring des modèles
- responsable production - Garantir l'orchestration et la mise en oeuvre de la chaîne de traitement sur la plateforme de production
- responsable surveillance & alerte - Surveiller les indicateurs de performance et gérer les alertes en cas de problème.

Proposition #8 - Encadrer la mise en production des algorithmes avec différentes parties prenantes

Dans le cas des algorithmes, le droit ne peut pas se contenter de cadrer *a posteriori* les possibilités techniques et chercher à rééquilibrer après-coup une asymétrie de pouvoir inscrite dans le design même des algorithmes : il doit construire un cadre pour orienter leur développement, limiter d'emblée la mise en place de telles asymétries et la façon dont elles se confortent.

Or, comme le souligne le CNnum dans son [rapport sur le travail à l'ère des plateformes](#), un des outils les plus controversés du management algorithmique est le design abusif. Celui-ci relève de choix opérés par les architectes et concepteurs de l'application qui mobilisent les biais cognitifs de l'utilisateur de manière à influencer son comportement.

La notion de design trompeur désigne des pratiques visant à représenter des éléments de telle manière qu'ils pourraient induire l'individu en erreur. Les designs dangereux correspondent à des méthodes qui peuvent rendre le travailleur directement ou indirectement vulnérable. Ces designs peuvent inciter le travailleur à commettre des infractions. Par exemple, des chronomètres étaient déclenchés automatiquement sur certaines applications de livraison à vélo lors de l'acceptation d'une course. Cela pouvait inciter les coursiers à commettre des infractions au Code de la route en jouant sur des biais cognitifs.

Il convient donc de former les designers aux enjeux d'éthique et plus précisément de droit du travail, par exemple par la création de formations croisées au sein des cursus de droit et de design.

Les designs trompeurs pourraient par ailleurs être condamnés sur le fondement des pratiques commerciales déloyales. Ainsi, le 15 mars 2021, la loi californienne sur la protection de la vie privée des consommateurs ([California Consumer Privacy Act, CCPA](#)) a été actualisée pour interdire aux entreprises d'utiliser certains 'dark patterns' (techniques utilisées par les plateformes pour dissuader un internaute de se retirer d'un service auquel il a souscrit ou l'amener à partager involontairement ses données privées).

Le terme de « *dark pattern* », apparu pour la première fois en 1994 dans le magazine *Wired*, désigne des interfaces (considérées comme non éthiques) destinées à obtenir l'action recherchée de la part de l'utilisateur. Comme l'expose la CNIL dans son rapport sur "[La forme des choix](#)"², la question de l'indication par le design est cependant vaste : au-delà des dark patterns à proprement parler, de nombreux designs sont conçus pour

² La forme des choix, p. 26. Woodrow Harzog distingue design abusif, design trompeur et design dangereux.

endormir la vigilance de l'utilisateur, susciter des comportements automatiques, et, comme le souligne Dominique Boullier, *“réduire à l'extrême les hésitations et les arbitrages conscients, pour créer une forme de naturalité (...) qui semblera très économique sur le plan cognitif.”* (p.13)

Cette dimension appellerait un cadre de régulation plus vaste.

AXE 2 : PROTÉGER - Adapter les droits des travailleurs aux transformations du travail portées par les algorithmes

Proposition #9 - Introduire un véritable droit à la décision humaine

Le RGPD a créé des droits pour les travailleurs en matière de management algorithmique³ :

- le droit pour les travailleurs de ne pas faire l'objet de décisions fondées exclusivement sur un traitement automatisé ([article 22](#)),
- l'encadrement du profilage, c'est-à-dire du traitement automatique de données pour évaluer des personnes ou prédire leur comportement ([article 4](#)).

Le [Digital Service Act](#), pour sa part, devrait d'ici 2022 régir la responsabilité des intermédiaires -et plus spécifiquement les plateformes-, soumis à des obligations de transparence (notamment algorithmique) et de reddition de compte accrues. Le Règlement européen [Platform to Business](#) (P2B), entré en vigueur le 12 juillet 2020, pose déjà des bases de ces exigences de transparence, d'équité et de responsabilité.

Mais **ces droits consacrés par le RGPD et le futur Digital Service Act sont avant tout formels**. Il suffit à l'employeur ou au donneur d'ordre d'apporter une réponse toute aussi formelle à la requête du titulaire des droits, pour les rendre **ineffectifs** : l'employeur peut confirmer avoir contrôlé la décision de l'algorithme, la plateforme peut affirmer que son code est intelligible pour un data scientist compétent ou que sa complexité est incompressible.

La CNIL a [recommandé](#) que le droit de ne pas faire l'objet de décisions fondées exclusivement sur un traitement automatisé (article 22 RGPD) soit considéré plus largement comme un **droit à la supervision** de la décision algorithmique par l'employeur ou le donneur d'ordre, sur des séries plus ou moins nombreuses de décisions. Ceci permettrait d'introduire *“des formes de délibération humaine et contradictoire encadrant et accompagnant l'utilisation des algorithmes en examinant et interrogeant le paramétrage mais aussi tous les effets - directs et indirects - du système”*.

³ Le [Digital Service Act](#) devrait aussi d'ici 2022 régir la responsabilité des intermédiaires -et plus spécifiquement les plateformes-, soumis à des [obligations de transparence](#) (notamment algorithmique) et [de reddition de compte accrues](#). Le Règlement européen [Platform to Business](#) (P2B) posait déjà depuis juillet 2020 les bases de ces exigences de transparence, d'équité et de responsabilité.

Le Règlement européen sur l'IA devrait introduire le **droit positif à une décision humaine** (avec le cas échéant mise en œuvre d'un procédé algorithmique) vs. actuellement le droit "en creux" de ne pas faire l'objet d'une décision algorithmique:

- Ce contrôle humain de la décision algorithmique doit être défini par les pouvoirs publics, selon des critères permettant de responsabiliser l'employeur ou le donneur d'ordre, tout en rendant ce droit effectif pour le travailleur.
- Ce droit à la décision humaine devrait être **mis en avant par le donneur d'ordre ou la plateforme**, tant lors de la conclusion du contrat qu'au cours de son exécution, avec la mise en place de **procédures internes facilitant sa mise en œuvre**.

L'article 14 du projet de Règlement entend imposer une conception et un mode de fonctionnement de l'algorithme permettant à son utilisateur de conserver le contrôle sur le travail et les décisions de l'algorithme. Toutefois il n'est pas créateur de droits pour le travailleur ou prestataire impacté par la décision dictée ou assistée par l'algorithme.

Comme exposé, ce droit peut prendre la forme d'une évolution de l'article 22 du RGPD. Cependant, il peut également trouver sa place à l'article 14 du projet de règlement européen, en ce que l'algorithme utilise également pour aboutir à ses décisions, des données qui ne sont pas personnelles et échappent par nature au champ d'application du RGPD.

Il est donc loisible de compléter cet article 14 pour introduire, à la suite des moyens mis à la disposition de l'utilisateur de l'algorithme pour conserver le contrôle sur l'IA, un véritable droit pour le travailleur de solliciter l'utilisateur et lui imposer l'usage effectif des moyens que ce dernier tient de l'article 14.

Ce droit à la décision humaine pourrait en outre être inclus dans la "Charte des droits numériques" que va prochainement proposer la Commission européenne.

Proposition #10 - Introduire un droit de communication d'un document intelligible et opposable sur les algorithmes au travail

La communication d'un document intelligible sur l'algorithme et opposable à son émetteur, devrait être un droit pour chaque travailleur et pour les institutions représentatives du personnel.

Ce document devra **exposer de manière claire et complète les effets de la mise en œuvre de l'algorithme** sur le management dont les travailleurs sont l'objet.

L'émergence de la formalisation d'une obligation d'information - contractuelle voire précontractuelle, à la charge de l'employeur ou du donneur d'ordre est comparable au devoir d'information précontractuelle déjà en vigueur en droit de la consommation ou en droit commun des contrats.

L'algorithme doit être explicité sur demande, par une équipe d'experts de la data (type data scientist ou data engineer, dont les services pourront être sollicités par les institutions représentatives du personnel, comme le sont déjà ceux d'autres experts).

Proposition #11 - Assurer un meilleur partage de valeur créée liée à l'usage de l'IA au travail entre l'entreprise et les travailleurs producteurs de données

Une perception trompeuse de la création de valeur liée au management algorithmique pourrait conduire à penser que c'est l'IA seule qui crée la valeur et donc qu'il est légitime que celle-ci revienne, intégralement, à l'entreprise.

En réalité, quelle que soit la pertinence de ses apports, l'IA n'intervient qu'au bout de la chaîne de la création de valeur, qui ne pourrait en tout état de cause exister sans les données elles-mêmes sur lesquelles elle s'appuie ni donc sur les travailleurs qui en sont à l'origine.

Même si cette production de données ne résulte pas d'un travail spécifique additionnel du travailleur par rapport à son activité normale au sein de l'entreprise, il joue donc un rôle essentiel dans la création de valeur liée au management algorithmique, qui justifie qu'il en perçoive une partie.

Cela est d'autant plus vrai lorsque les données captées par la machine relèvent de l'expertise personnelle du travailleur et que la machine apprend de cette expertise et la modélise (ex. dans le domaine médical ou juridique), sans que ce "transfert de savoir et de compétences" soit prévu en tant que tel dans le contrat de travail ni donc dans la rémunération du travailleur.

Par exemple, les plateformes VTC qui tentent d'évoluer vers des modèles avec véhicules autonomes reverseront-ils une part de leurs bénéfices aux chauffeurs qui auront alimenté les IA à l'origine de ces véhicules autonomes ?

Certes, on pourrait arguer du fait que les travailleurs sont déjà susceptibles de bénéficier de cette valeur par le biais de conditions de travail potentiellement plus faciles, plus intéressantes et plus productives. Il n'en reste pas moins qu'*in fine* c'est

bien l'entreprise qui aujourd'hui capte le surcroît de création de valeur financière, alors même qu'elle est en partie due aux travailleurs eux-mêmes.

Au surplus, les données ainsi créées ont une valeur économique qui va au-delà de leur finalité première. Quand ces données ont une seconde vie, notamment par leur agrégation et leur commercialisation à des fins distinctes de celles visées par l'activité professionnelle initiale, il est légitime que la valeur perçue par l'entreprise soit directement partagée avec ceux qui ont contribué à la créer.

Dans le cas de la captation de données purement contingentes dans le cadre de l'activité du travailleur, comme dans celui où intervient une part d'expertise du travailleur, l'entreprise parvient à créer de la valeur indirectement à partir de l'activité du travailleur sans que cette création constitue une tâche contractuelle. N'étant pas liée au contrat de travail, elle ne fait donc pas partie de l'équilibre économique de ce contrat, au détriment des intérêts du travailleur.

Or il est pleinement légitime que le travailleur perçoive une partie de la valeur que son activité a contribué à créer, même si ce n'était pas l'objet direct de son travail, dès lors que l'entreprise elle-même en tire profit.

C'est pourquoi nous proposons d'introduire dans le droit du travail une disposition prévoyant des mécanismes de partage de valeur au profit des travailleurs en cas de captation et d'exploitation de données produites par le travailleur qui donnent lieu à création de valeur pour l'entreprise sans que celle-ci constitue une tâche contractuelle.

Naturellement, la manière la plus simple de régler cette question serait d'inclure dans les clauses du contrat de travail des modalités de partage de valeur qui portent sur l'intégralité de la valeur créée par l'entreprise à partir de l'activité du travailleur.

Certes il n'appartient pas au code du travail de régler cette question dans les détails, mais au moins d'en affirmer le principe, afin que les partenaires sociaux puissent, par le dialogue social, trouver des solutions adaptées en fonction des différentes situations.

Ce dialogue social risque toutefois de se heurter à la technicité du sujet et à la difficulté d'évaluer précisément la valeur créée et d'en tracer les modalités de création, avec le risque qu'il soit peu opérant.

Mais compte tenu de l'ampleur des enjeux, et de la dynamique exponentielle de la valorisation des données produites de manière incidente par les travailleurs, il nous semble essentiel de poser un premier jalon et de lancer une dynamique de dialogue social.

AXE 3 : GARANTIR - Garantir le respect effectif des principes de l'IA responsable et des droits travailleurs

Proposition #12 - Passer de l'autoévaluation de la conformité réglementaire des algorithmes par des fournisseurs à une évaluation par des organismes tiers

Le projet de Règlement de la Commission relatif à l'IA ne déroge pas au “nouveau cadre législatif” de l'UE, visant à élaborer la base juridique européenne de la conformité de tous produits non alimentaires. Il reprend donc les notions de contrôle *ex ante* et *ex post*, ici appliqués à l'algorithme, tout comme il consacre par avance l'intervention des organismes d'harmonisation des normes au niveau européen que sont le Comité Européen de Normalisation (CEN) et le Comité Européen de Normalisation en Électronique et en Électrotechnique (CENELEC).

En effet, au-delà des principes que ce projet de Règlement entend fixer, le CEN et le CENELEC sont amenés à élaborer les normes harmonisées que les fournisseurs d'algorithmes pourront suivre, à défaut de tenter d'interpréter eux-mêmes le Règlement à venir.

Or, le respect par un produit - ici par un algorithme, de ces “normes CE”, soit les normes édictées par ces comités, vaut présomption de conformité.

S'inscrivant toujours dans son “nouveau cadre législatif”, l'UE distingue les systèmes d'IA destinés à être utilisés en tant que composants de sécurité de produits par ailleurs réglementés, tels que les jouets ou les dispositifs médicaux, et les systèmes d'IA autonomes à haut risque tels qu'énumérés à l'annexe III du projet de Règlement.

Les premiers devront être soumis aux procédures de contrôle *ex ante* et *ex post* régissant déjà les produits dont ils sont les composants, tandis que les seconds devront faire l'objet d'un nouveau système de mise en conformité, qui repose de fait presque exclusivement sur un contrôle interne *ex ante* par le fournisseur du système d'intelligence artificielle.

Or, comme l'estime [Valerio De Stefano](#), cette auto-évaluation par le fournisseur est insuffisante pour fournir des protections équivalentes à celles mises en place pour d'autres systèmes à haut risque “*qui doivent être soumis à des procédures plus strictes d'évaluation de la conformité, avec la participation d'un organisme notifié*”. D'autant, que cette disposition semble n'avoir fait l'objet d'aucune forme de dialogue social au niveau de l'UE.

Il est par ailleurs notable que plusieurs lois nationales au sein de l'UE vont plus loin que le règlement et "interdisent ou limitent considérablement le recours à des outils technologiques pour contrôler les travailleurs (Aloisi et Gramano, 2019)".

Les acteurs privés comme les décideurs publics nationaux semblent donc prêts à relever le niveau d'exigence et de transparence applicable à la conformité des systèmes d'intelligence artificielle, en actant la réalité des risques comme la nécessité de tiers certificateurs spécifiques.

A fortiori, nous proposons que cette évaluation de la conformité des algorithmes à haut risque soit établie, *ex post* et/ou *ex ante*, sous la responsabilité d'organismes tiers tels que les organismes notifiés - acteurs déjà connus de la conformité sous le régime du "nouveau cadre législatif" de l'UE, et non par le seul biais de l'auto-évaluation du fournisseur qui pourrait avoir l'effet de rendre ineffective la conformité recherchée, conjointement avec la présomption de conformité attachée au respect des normes édictées par le CEN et le CENELEC.

Proposition #13 - Mettre en place des procédures d'estimation des risques des algorithmes, sur le modèle du stress test bancaire

Les entreprises et institutions sont de plus en plus nombreuses à se confronter au problème du "risque algorithmique" et à recourir à de nouveaux experts, spécialistes de l'audit algorithmique, capables d'auditer et d'interpréter les résultats d'un modèle et d'évaluer sa conformité réglementaire.

Mais que recouvre cette expression de "risque algorithmique" ? Il faut ici adopter une vision large et ne pas se limiter aux seuls risques liés à la fiabilité des systèmes. Comme cela a déjà cours dans certains secteurs industriels (par exemple l'aérospatiale, les soins de santé, et surtout les études d'impact environnementales), il faut inclure dans les risques la question du préjudice social que les algorithmes sont susceptibles de susciter.

Les entreprises sont de facto déjà sensibles à cette dimension du fait des risques réputationnels auxquels peuvent les exposer certains effets indésirables des algorithmes : on a par exemple beaucoup parlé, depuis 2018, du problème de [l'outil de recrutement développé par Amazon, qui éliminait les femmes](#). Pour autant, il n'existe pas encore de cadre systématique permettant d'aborder la [question des risques algorithmiques](#), en particulier en matière de management.

L'évaluation en matière d'algorithmes implique de trouver un compromis entre trois exigences : le niveau de précision, de reproductibilité d'une décision issue d'un SIA par apprentissage, l'explicabilité d'une décision en fonction du niveau (concepteur, utilisateur, usager) et du domaine concerné, les [risques selon le domaine envisagé](#). Pour

cela, il faut aussi distinguer clairement les risques liés aux données des risques liés aux algorithmes. Si l'utilisation d'IA symbolique déterministe (analyse experte des règles de décisions codées dans l'algorithme) permet en effet une évaluation indépendante des algorithmes, l'apprentissage automatique et plus encore le deep learning engendrent de facto un effet boîte noire intimement lié aux données utilisées pour exercer la solution IA.

Il s'agit ensuite de mettre en place des protocoles de contrôle et d'audit.

En matière de discrimination, on peut se référer à certains protocoles déjà existant, par exemple le guide établi par Philippe Besse visant à mieux [détecter, évaluer les risques des impacts discriminatoires des algorithmes d'IA](#). Celui-ci dresse un protocole passant par un ensemble de questions permettant d'accompagner une organisation, concernant la conception des solutions, des processus pour tester et contrôler les biais éventuels au cours de la phase de mise au point, de déploiement et d'utilisation du système, l'analyse ou des indicateurs utilisés.

L'évaluation peut ensuite prendre plusieurs formes.

- Interne, ce qui implique un travail de formation des salariés, la mise à disposition d'outils (c'est-à-dire principalement de bibliothèques d'audit algorithmique, c'est-à-dire d'algorithmes capables d'analyser d'autres algorithmes). Le contrôle interne peut porter sur les biais les plus fréquents et les mieux identifiés : **le biais des données** (un algorithme de recrutement entraîné sur une base de données dans laquelle les hommes sont sur-représentés exclura les femmes), **le biais de stéréotype** (une femme a tendance à cliquer sur des offres d'emplois qu'elle pense plus faciles à obtenir en tant que femme, **le biais de variable omise**, **le biais de sélection** (une banque utilisera des données internes pour déterminer un score de crédit, en se focalisant sur les personnes ayant obtenu ou pas un prêt, mais ignorant celles qui n'ont jamais eu besoin d'emprunter, etc).
- Un contrôle externe, qui peut être exercé par des structures tierces, comme le [recommande l'Institut Montaigne](#). La création d'un [pôle d'expertise sur les algorithmes](#) de vingt personnes à la Direction générale des entreprises (ministère de l'Economie) capable d'auditer les algorithmes, initiée en 2019, pose des jalons dans cette direction, même s'il serait souhaitable d'en élargir le périmètre et d'y associer des experts et intervenants d'horizons divers (collectifs de travailleurs, sociologues, juristes, etc.) pour identifier les points sensibles en matière de management. Mais il faudrait également prévoir l'intervention d'autres acteurs, semblables aux sociétés d'audit et envisager des procédures de certifications algorithmiques, comme c'est déjà le cas pour les questions comptables. Une autre possibilité est celle de la

responsabilisation, laissant les organisations libres de la forme de contrôle qu'elles choisissent de mettre en œuvre, mais sanctionnant fortement a posteriori la réalisation du risque.

- Une évaluation sectorielle afin de prévenir toute collusion implicite ou explicite entre algorithmes. Ceci peut se faire via des procédures de benchmarking ou de comparaison inter-sectorielle, par exemple, afin de pouvoir détecter une possible convergence des conditions et de rémunération de travail vers le plus petit dénominateur commun. Ceci ne peut se faire sans collaboration entre différents régulateurs, notamment les autorités à la concurrence qui devrait être sensibiliser à l'impact qu'une telle collusion puisse avoir sur le marché du travail.

Ces instances de contrôles, internes comme externes, devront par ailleurs pouvoir bénéficier d'outils adaptés et efficaces, ce qui implique de soutenir les recherches actuellement en cours en matière d'évaluation algorithmique. On peut citer en exemple le [projet "Matilda" de stress-tests d'algorithmes](#) de l'Université de Melbourne qui fournit des outils publics pour une évaluation plus fine et multifactorielle de la performance algorithmique, de façon à adapter la procédure de test aux critères et finalités d'une solution. En effet, souligne le Professeur Kate Smith-Miles de University of Melbourne, responsable du projet, les algorithmes sont actuellement testés selon des procédures très standards, sur des échantillons pré-établis, bien balisés par la littérature scientifique, mal adaptées aux contextes d'utilisation, ce qui limite en particulier la possibilité d'anticiper et d'évaluer les effets humains de leur implémentation.

Proposition #14 - Intégrer la mesure des risques liés au management algorithmique dans les risques au sein des entreprises, notamment psychosociaux

Les études montrent que le déploiement de l'IA dans l'entreprise accentue les écarts de salaires entre ceux qui maîtrisent ou ne maîtrisent pas l'usage des algorithmes.

Au-delà, l'usage de l'IA impacte aussi la manière pour les travailleurs d'appréhender bien-être, stress et santé au travail.

Par la profondeur de ses impacts professionnels et la diversité des tâches concernées pour le travailleur, l'IA démultiplie les risques psychosociaux à trois niveaux :

1. une **hyper-standardisation des tâches pour les travailleurs**, l'IA laissant à l'humain les tâches non encore automatisées ;
2. une **hyper-complexification des tâches pour les travailleurs**, l'IA supportant les tâches les plus simples. Une [étude japonaise](#) menée en 2017

auprès de 10.000 travailleurs a montré que si l'usage de l'IA contribuait à la satisfaction au travail en concentrant une partie des travailleurs sur des tâches moins répétitives, l'IA introduit dans ces cas-là un niveau de stress supplémentaire lié à la complexité des tâches restant à réaliser par l'humain.

3. enfin, **les relations entre l'humain et les robots collaboratifs ou cobot** induisent [trois risques](#)⁴ en terme de Sécurité et Santé au Travail (SST) :
 - i) Les risques de collision robot-humain, dûs à un comportement imprévisible du robot suite à l'apprentissage automatique, ii) les risques de sécurité, lorsque la connexion Internet des robots affecte l'intégrité de la programmation logicielle, entraînant l'apparition de vulnérabilités, iii) Les risques environnementaux, en cas de dégradation des capteurs ou de comportement humain imprévu dans des environnements non structurés

Notons que ces “cobots” sont une réalité déjà bien implantée : l'organisation néerlandaise pour la recherche scientifique appliquée [écrivait dès 2018](#) qu’“Amazon dispose ainsi de 100 000 cobots augmentés IA, qui ont permis de réduire à moins de deux jours le temps de formation des employés. Airbus et Nissan utilisent quant à eux des cobots pour accélérer la production et améliorer l'efficacité.”

La prévention des risques psychosociaux est d'autant plus importante qu'elle ne peut se fonder exclusivement sur des modèles génériques et requiert au contraire, pour un minimum d'efficacité, une démarche proactive d'audit et d'analyse spécifique à l'entreprise.

Cela est d'autant plus vrai à l'occasion de la mise en œuvre d'un algorithme dans le cadre de la relation de travail.

Nous recommandons d'intégrer la mesure des risques liés au management algorithmique dans le cadre de la prévention des risques psychosociaux en répondant aux prérequis suivants :

- l'implication des dirigeants de l'entreprise,
- la formation massive de l'ensemble des collaborateurs aux usages, opportunités et risques liés à l'IA. Une [étude menée par IBM](#) indique que, au sein des dix premières économies mondiales, quelque 120 millions de travailleurs devront être re-formés et leurs compétences recyclées pour prendre en charge l'IA et l'automatisation intelligente,
- le développement d'une culture de la prévention des risques psychosociaux liés à la mise en oeuvre d'un algorithme dans les relations de travail,

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

- un travail de réflexion en profondeur pour savoir comment conjuguer au mieux les compétences des humains et des robots dans le contexte spécifique de l'entreprise concernée⁵,
- l'implication des représentants du personnel et leur formation en conséquence,
- la possibilité pour les représentants du personnel d'être assistés de tiers dans l'identification et l'évaluation de ces facteurs de risque spécifiques,
- le cas échéant, l'accompagnement et la formation d'acteurs des services de santé au travail sur les problématiques spécifiques à la mise en oeuvre des algorithmes au sein de l'entreprise,
- l'information des salariés sur les démarches préventives et leur accès aux outils de cette prévention,
- la possibilité pour les salariés et représentants du personnel d'initier des alertes en matière de risques psychosociaux liés à la mise en oeuvre des algorithmes, sans conséquence sur leur emploi ou leur mandat,
- le développement d'une section spécifique à la mise en oeuvre des algorithmes au sein de l'entreprise, dans le Document Unique d'Evaluation des Risques,
- le développement de plans d'action après identification des risques et la budgétisation de leur mise en oeuvre, suivie à intervalles réguliers mais réalistes, d'un nouvel audit pour tenir compte du caractère évolutif des algorithmes, de leurs usages et donc de leur impact en terme RPS.

Proposition #15 - Élargir les responsabilités du DPO au champ du management algorithmique

Le RGPD impose déjà à certaines entreprises et institutions la création d'un poste de DPO, pour Data Protection Officer (Délégué à la Protection des Données).

Il serait loisible d'élargir les responsabilités de celui-ci, ou d'imposer aux organisations qui recourent massivement à des processus de management algorithmique la création d'un **poste d'Algorithmic Management Officer** (délégué à la Protection d'un Management Algorithmique Responsable).

Pour les entreprises, nommer un responsable, l'intégrer dans un organigramme... est en effet souvent le premier moyen pour initier une politique, des processus, des normes et des cadres qui permettront de la déployer. Sa mission serait d'une part d'informer et conseiller son organisation et d'autre part, de contrôler les développements des applications et interfaces pour un usage responsable des outils de management algorithmique. Cet acteur pourrait être salarié de l'entreprise ou prestataire extérieur et

⁵ Par exemple, l'institut fédéral allemand chargé des questions de sécurité et santé au travail (BAuA) organise des ateliers annuels sur le thème de la «collaboration homme-robot». <https://osha.europa.eu/en/publications/future-work-robotics/view>

relèverait le cas échéant du plus haut niveau de l'organisation. Il animerait un réseau de relais au sein des filiales de l'organisation et/ou une équipe d'experts en interne (design, juridique, informatique, ressources humaines) et serait, à l'image du DPO, l'interlocuteur naturel des autorités comme des représentants du personnel pour toute problématique liée au management algorithmique.

Proposition #16 - Intégrer les droits individuels et collectifs relatifs aux algorithmes dans les codes de conduite des entreprises

Ces codes de conduite constituent des annexes au règlement intérieur. Ils sont donc soumis aux procédures d'information et de consultation des institutions représentatives du personnel.

Nous proposons d'inclure dans la partie réglementaire du code du travail, l'obligation pour les entreprises de faire mention des droits individuels et collectifs relatifs à la mise en œuvre d'algorithmes par l'employeur, dans les codes de conduite et chartes informatiques.

Nous proposons également d'intégrer dans les obligations de déclaration RSE des entreprises un volet "IA responsable" décrivant les pratiques de transparence en termes de management algorithmique et les mesures prises par l'entreprise pour la maintenir dans la durée.

Parmi ces bonnes pratiques, la mobilisation des acheteurs est essentielle pour garantir un code de transparence avec les fournisseurs et les salariés. La formation et l'accompagnement de ces acheteurs permettent le déploiement d'IA responsables et évitent les effets "boîtes noires" d'algorithmes achetés et utilisés au sein de l'entreprise sans connaître leur réel mode de fonctionnement .

Proposition #17 - Former partenaires sociaux et salariés aux enjeux et aux pratiques de l'IA au travail

L'usage de l'IA dans les ressources humaines s'est développé à vitesse considérable ces dernières années. En effet, [la moitié des applications](#) IA développées entre 2010 et 2020 vise l'automatisation des processus de management d'entreprise. D'ores et déjà, près de [40% des sociétés internationales - soit un quadruplement en 4 ans- utiliseraient des algorithmes prédictifs dans leurs activités RH](#) : recrutement, gestion de carrière, GPEC notamment.

Or l'intégration de l'IA transforme profondément les processus RH : processus de recrutement et d'intégration des salariés, gestion de leur évolution de carrière, évaluation et gestion de performance, gestion active et prévisionnelle des emplois et compétences, plans de formation adéquats par rapport aux besoins des compétences futures de l'entreprise. La machine devient à tout le moins une aide à la décision humaine par sa capacité à brasser quantités de données sur les travailleurs : compétences, postes occupés, performances...

Comme le met en évidence le [rapport du CNNum sur le travail à l'ère des plateformes](#), la transformation numérique pose des défis inédits au dialogue social. C'est encore plus vrai en matière de management algorithmique.

Il est donc indispensable de former les **représentants des personnels**, à l'ensemble de ces enjeux afin de leur donner les moyens de mener à bien leur mission de négociation sur tous les sujets RH impactés par l'IA (recrutement, gestion de carrière, GPEC, formation...).

Quelles types de données sont collectées par l'entreprise via ces logiciels et algorithmes RH, pour quels usages, quelles données alimentent les algorithmes, de quelle façon on s'assure de l'absence de biais, quelle est la part de l'humain et celle des algorithmes dans les décisions, etc. ?

L'employeur et les organisations syndicales doivent avoir cet échange dès la mise en place des logiciels et algorithmes RH dans l'entreprise afin d'explicitier l'objectif, les fonctionnalités et les résultats attendus du logiciel dans les activités de gestion RH.

Une meilleure connaissance et une formation adaptée à ces outils doit permettre une discussion éclairée et équilibrée entre les partenaires sociaux sur des sujets humains clés.

Cette formation relative aux usages des algorithmes dans l'organisation du travail humain et la gestion des "ressources humaines" pourrait s'inscrire dans le cadre du droit à congé de formation économique, sociale et syndicale et du droit à la formation sur les questions de santé, sécurité et conditions de travail prévues respectivement aux articles [L.2145-1](#) et [L.2315-18](#) du Code du Travail.

De même, il semble souhaitable que **tous les salariés** travaillant avec des algorithmes soient formés non seulement au fonctionnement spécifique de ces algorithmes mais plus largement aux enjeux de l'IA au travail et du management algorithmique.

Ils devraient notamment être éclairés sur la nature des données personnelles collectées et la façon dont elles sont collectées et traitées, et sur les usages qui en sont faits par les algorithmes.

L'introduction de l'IA dans l'organisation du travail humain modifie profondément les activités et les métiers des travailleurs concernés : hyper-standardisation de certaines tâches pour les travailleurs, l'IA laissant à l'humain les tâches non encore automatisées, à l'inverse grande complexification des tâches pour les travailleurs, l'IA supportant les tâches les plus simples, gestion des interactions entre l'humain et les robots collaboratifs ou cobot.

Pour les travailleurs qui seront amenés à développer des activités de supervision de la machine, d'analyse et de résolution de problèmes et plus généralement d'activités à plus forte valeur ajoutée, avoir une connaissance approfondie des systèmes IA et de leurs fonctionnalités est indispensable afin de ne pas perdre la maîtrise du poste et la capacité à pouvoir réagir en cas de défaillance.

Proposition #18 - Constituer une capacité d'expertise publique approfondie et former les administrations concernées pour permettre aux pouvoirs publics de produire une réglementation pertinente et indépendante sur l'IA au travail

Il est essentiel que les pouvoirs publics aient les moyens, ou plus exactement se donnent les moyens, de produire une **réglementation pertinente** sur l'IA au travail et de veiller à la fois à son respect et à son actualisation en fonction de l'évolution des techniques et des usages de l'IA appliquée au travail humain.

Cette considération pourrait paraître évidente mais la nouveauté radicale, la très grande technicité et la rapidité d'évolution du sujet imposent aux pouvoirs publics un investissement spécifique pour disposer d'une compréhension fine des réalités et des enjeux.

Produire une réglementation équilibrée entre le respect de la liberté d'entreprendre et la protection des droits des travailleurs, et plus largement entre les intérêts des entreprises et ceux des travailleurs, sera déjà, en soi, une tâche particulièrement complexe.

Or l'enjeu va encore bien au-delà. En effet, s'agissant de sujets aussi sensibles que les données et les libertés individuelles et la possibilité de manipulation d'êtres humains

par des machines, c'est le **fondement du pacte de confiance** entre les travailleurs et les employeurs qui est en jeu.

C'est pourquoi nous recommandons la **constitution d'une capacité d'expertise publique approfondie sur l'IA appliquée au travail humain**, qui soit indépendante de celle des opérateurs. C'est en effet la condition nécessaire à l'élaboration d'une politique et d'une réglementation véritablement adaptées aux spécificités du sujet, qui ne soit ni inspirées par les solutions "proposées" par les opérateurs économiques, ni plaquées sur des solutions adoptées pour des problèmes différents.

Cette capacité d'expertise pourrait être incluse au sein du [PEReN : Pôle d'expertise de la Régulation numérique](#).

Qui aujourd'hui, à Bercy ou au Ministère du Travail par exemple, peut se targuer de bien connaître les questions de management algorithmique, de nudge, de designs trompeurs, d'explicabilité et d'auditabilité des algorithmes... ? La constitution de cette capacité d'expertise ne saurait suffire à elle seule : il importe de **former les administrations concernées** par la réglementation à ces enjeux.

Proposition #19 - Mettre en place des dispositifs incitatifs, de type labellisation, pour mobiliser l'ensemble de l'entreprise

La mise en visibilité sur la place publique des pratiques vertueuses, pourrait avoir pour conséquence des phénomènes de punition et d'incitation par le marché. Cela implique de mettre en place des outils permettant de répertorier, catégoriser, inventorier et analyser les bonnes et mauvaises pratiques (...) sur le modèle par exemple de [GoodUI](#).

Le [Conseil national du numérique](#) préconise ainsi de mettre en place un digiscore qui s'appliquerait aux plateformes d'emploi type Uber ou Deliveroo. Celui-ci, pensé sur le modèle du "nutriscore", qui permette d'éduquer les citoyens consommateurs en leur donnant une indication claire, avant qu'ils aient recours à un opérateur, de la manière dont celui-ci respecte certains critères, en particulier sociaux, mais également technique, en particulier la façon dont les algorithmes sont utilisés comme outils de management.

Dans la même perspective, mais s'inspirant pour sa part de l'application [Yuka](#), [l'Institut Montaigne préconise](#) de rendre publique une liste du type de données utilisées pour entraîner l'algorithme et des objectifs que l'algorithme doit atteindre.

Enfin, la labellisation pourrait également s'avérer un outil approprié pour valoriser les usages responsables des algorithmes en matière de management.

Une première option serait une labellisation des outils techniques, imposant un ensemble de critères de transparence et d'explicitation. Cette labellisation pourrait être confiée à l'AFNOR, sous couvert des pouvoirs publics, et impliquer des autorités comme le Défenseur des droits (pour les questions de discrimination). Elle pourrait se concentrer "sur l'auditabilité des algorithmes, sur la qualité des données ou encore sur la présence d'un processus d'évaluation des risques de biais au sein de l'entreprise", comme le suggère aussi l'Institut Montaigne, lequel souligne aussi que mettre en place ce type de label permettrait également d'inciter à développer d'avantages de compétences et de méthodes d'audit des algorithmes. Le même Institut préconise par ailleurs de faire porter les processus de labellisation sur "l'équipe développant l'algorithme (ses processus internes, ses pratiques, sa composition) plutôt que sur les algorithmes eux-mêmes". [Fair data use](#) est un exemple de label de ce type.

Une seconde option serait d'inscrire plus fortement la question du management algorithmique dans certains labels déjà existant concernant les fonctions RH. Le **Label Diversité**, créé en 2008 et propriété de l'État, semble par exemple très bien se prêter à ce type de question. Visant à prévenir les discriminations et à promouvoir la **diversité** dans les secteurs public et privé, il se concentre sur les processus de ressources humaines d'une structure (recrutement, gestion des carrières, formation, communication), et plus généralement sur l'ensemble des activités de celle-ci (relation avec les usagers et clients, avec les sous-traitants, et autres partie-prenantes), et englobe un ensemble de champs à caractère évolutifs, auquel la question des algorithmes utilisés à des fins managériales pourrait être explicitement incluse.

AXE 4 : ACTUALISER - Permettre l'adaptation du cadre réglementaire dans le temps aux évolutions de l'IA

Proposition #20 - Inclure dans le périmètre des activités considérées comme “à haut risque” par le projet de Règlement européen sur l'IA les décisions liées à la rémunération

L'approche de la Commission européenne en matière d'intelligence artificielle se fonde d'abord sur une identification et un encadrement des risques selon les domaines d'applications – publics ou privés – concernés :

- les **risques inacceptables** dans le champ de la sécurité, des moyens de subsistance et des droits des personnes ;
- les **risques élevés** dans les domaines touchant les infrastructures critiques (énergie, transports...), l'éducation, la formation professionnelle, l'emploi, les ressources humaines, les services privés et publics essentiels, le maintien de l'ordre, la justice, les processus démocratiques...

L'annexe III du projet de Règlement européen sur l'IA de la Commission européenne (Annexe III) prévoit déjà d'inclure le recrutement, la prise de décisions de promotion et de licenciement, l'attribution des tâches et l'évaluation des performances dans les activités à haut risque. Toutefois, les décisions liées à la rémunération y sont formellement absentes et pourtant cruciales. Le texte ne doit laisser aucune incertitude quant au fait qu'elles sont bien couvertes par le règlement, parmi les activités à haut risque.

Proposition #21 - Faire du management algorithmique un thème du dialogue social à part entière

Pour [Valerio de Stefano](#), la négociation des algorithmes pourrait voire devrait dans les années à venir devenir un objectif crucial pour l'ensemble des organisations syndicales.

Il souligne en effet que la négociation collective demeure le meilleur outil pour assurer le respect des intérêts des travailleurs en matière de management algorithmique et éviter les impacts négatifs sur leur santé physique, mentale et leur sécurité. *“Les représentants des travailleurs et les syndicats doivent être associés aux processus de prise de décision qui aboutissent à la définition et à l'utilisation d'algorithmes. Il est indéniablement plus efficace d'aborder ces questions en amont que de devoir limiter les dégâts par la suite”*.

Comme le soulignent Odile Chagny et Franck Bonot, co-animateurs du [réseau Sharers & Workers](#), c'est en principe déjà le cas, les représentants du personnel ayant un droit de consultation prévu dans toutes les situations d'introduction de nouvelles technologies. Pour autant, ils sont très rarement consultés sur ces thèmes.

Or le dialogue social est crucial pour mettre en place dans l'entreprise une approche *“l'humain aux commandes”* et pour réguler les sujets aussi sensibles que le périmètre, l'usage et la propriété des données collectées sur les travailleurs.

Les accords collectifs pourraient fixer les limites de la surveillance exercée sur les travailleurs au moyen de l'IA. Ils pourraient également définir des critères pour accroître la transparence des processus décisionnels fondés sur l'IA (Dagnino et Armanoli 2019), afin de permettre une meilleure compréhension de la manière dont ces dispositifs tirent leurs conclusions, ou plus largement traiter de l'impact des algorithmes sur les pratiques de recrutement et de gestion de la main d'œuvre et sur l'organisation du travail humain, ainsi que le contenu des emplois et les compétences requises.

[Valerio de Stefano rappelle](#) à cet égard que l'[article 88 du RGPD](#), portant sur le traitement de données dans le cadre des relations de travail, reconnaît l'importance des conventions collectives pour la protection des droits en matière de traitement des données et pour la réglementation des processus de prise de décision algorithmiques.

Cet article définit les conventions collectives comme des instruments importants s'agissant de garantir un traitement juste et légal des données dans le cadre des relations de travail. Il mentionne explicitement le traitement des données à des fins de recrutement et de management, ce qui signifie que des conventions collectives pourraient offrir des moyens de protection appropriés dès lors que des outils fondés sur l'IA et des pratiques de management algorithmique sont utilisés sur le lieu de travail

(Dagnino et Armaroli 2019; Hendrickx 2018). Elles pourraient, par exemple, exiger des informations sur la manière dont les employeurs utilisent les données à caractère personnel des travailleurs ainsi que sur la façon dont ces données sont traitées par les systèmes d'IA (TUC 2021). Elles pourraient aussi interdire les utilisations les plus intrusives de ces technologies, telles que la neurosurveillance (De Stefano 2020).

Par ailleurs, le management algorithmique recourt massivement à des formes d'influence nouvelles, qui ne relèvent pas de l'exercice classique de l'autorité, contractualisée et basée sur l'instruction, mais de formes plus ou moins subreptices d'influence. Là encore, ce n'est pas tant l'usage de ces formes qui dilue et invisibilise l'autorité en poussant chacun.e à en intérioriser la contrainte qui est nouvelle, que la manière dont elles sont systématiquement mises en œuvre par certaines entreprises.

Pour faire face à ces pratiques qui affectent directement la capacité de l'individu à exercer son autonomie, il convient de faire de ces questions les objets de discussions collectives relevant du dialogue social à part entière.

Le management algorithmique pourrait ainsi contribuer à une refonte conséquente du droit du travail en permettant de ramener le « contenu » du travail au cœur des négociations⁶. Né de la révolution industrielle, le droit du travail cadre en effet les conditions de travail (horaires, rémunération) tout en laissant dans une certaine mesure le contenu et le mode d'accomplissement de la tâche à la discrétion de l'employeur.e.

Or, la pertinence de cette partition est remise en cause par le constat de plus en plus partagé d'une « perte » de sens du travail, lié à la perte d'autonomie des travailleur.euse.s qui se sentent de plus en plus dépossédé.e.s de la maîtrise de leur expertise. Or, en structurant de manière difficilement dissociable les conditions de travail et le contenu du travail, le management algorithmique conduit *de facto* à refaire du contenu du travail un sujet de délibération collective.

L'élaboration d'un droit social de l'algorithme est aussi une opportunité historique de fournir un véritable cadre juridique à ce mouvement de réappropriation.

⁶ Chagny O. et 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. Capitalisme, algorithmes et démocratie"

Conclusion et prochaines étapes

Historiquement, comme le rappelle Alain Supiot, les syndicats se sont battus “[contre l’exploitation au travail, pas contre l’oppression dans le travail](#)”, ce qui a conduit à évacuer “le contenu et le sens du travail de la notion de justice sociale”.

A l’heure du constat de plus en plus partagé d’une « perte » de sens du travail, liée notamment à un sentiment de perte d’autonomie de la part des travailleurs, c’est peut-être là, *in fine*, **l’enjeu fondamental posé par le management algorithmique : remettre le sens et le contenu du travail au centre du débat.**

Ce n’est que par la prise de conscience de tous les acteurs, entreprises, travailleurs, syndicats, pouvoirs publics, instances internationales, que nous trouverons les réponses adaptées.

Au total, s’il est clair que nous ne sommes encore qu’à l’orée du déploiement du management algorithmique et de ses effets sur le travail, l’ampleur des enjeux et leur actualité imposent d’agir dès maintenant.

Consacrer le principe d’intelligibilité de l’algorithme, introduire un encadrement législatif de certains éléments de contenu des algorithmes pour renforcer leur objectivité, garantir l’accès effectif des travailleurs aux droits nouveaux garantis par les Règlements européens, introduire de nouveaux droits individuels et collectifs : il y a clairement matière pour la France à porter une initiative forte dans le cadre de sa présidence de l’UE, pour promouvoir une IA responsable conciliant le respect de la liberté d’entreprendre et la protection des droits des travailleurs et de montrer l’exemple en faisant évoluer dès maintenant notre cadre législatif et réglementaire national.

Annexe 1 - Récapitulatif des ajustements proposés au Projet de Règlement européen sur l'IA

Le projet de Règlement sur l'IA proposé par la Commission européenne représente une avancée majeure. Il constitue en effet la première tentative mondiale de réglementation horizontale des systèmes d'IA et de prévention des usages abusifs qui pourraient en être faits. En particulier, son approche par catégorisation selon l'évaluation des risques est très pertinente et permet d'affirmer de manière claire et pédagogique les interdits et les points de vigilance indispensables.

Toutefois ce texte présente des marges d'amélioration importantes, liées notamment à l'écart entre les finalités poursuivies et le caractère insuffisant ou ineffectif des mesures avancées pour les atteindre. Il prend aussi insuffisamment en compte les enjeux spécifiques de l'IA au travail.

La raison première de cette insuffisance tient à l'insertion de ce texte dans le nouveau cadre législatif européen qui régit les produits non alimentaires au sein de l'UE, des jouets aux radiateurs électriques, avec les mécanismes usuels de conformité, d'intervention d'organismes notifiés, de responsabilité et de protection des consommateurs. Or ce cadre réglementaire est connu pour être perfectible. En l'espèce, il ne répond pas efficacement aux risques spécifiques posés par l'IA.

Dans le cadre de nos recommandations relatives au contrôle de l'IA au travail, nous proposons ainsi plusieurs ajustements au texte de la Commission européenne, avec 2 axes prioritaires, :

1. adapter le projet de règlement pour lui permettre de dépasser la déclaration d'intention et de faire respecter les droits des travailleurs,
2. assurer, compte tenu de la complexité singulière des sujets liés à l'IA, la compétence de tous les acteurs de cette réglementation, au niveau national et européen (autorités, CEN, CENELEC, organismes notifiés, fiduciaires de données, etc.).

1. Intégrer les principes directeurs du présent Livre blanc aux considérants du projet de Règlement de la Commission

2. Interdire expressément les algorithmes qui influencent les comportements par l'usage des biais cognitifs

Insérer à l'**article 5**, qui énumère les pratiques interdites en matière d'IA, les systèmes d'IA qui mobilisent les biais cognitifs de l'utilisateur de manière à influencer son comportement (design trompeur pouvant par ailleurs constituer une pratique commerciale trompeuse).

3. Introduire un principe d'explicabilité des algorithmes

Préciser les dispositions de l'**article 13** afin notamment de :

- introduire la notion d'explicabilité de l'algorithme, plus précise que la notion de transparence qui est envisagée dans le projet de règlement à différents degrés,

- entériner le fait que cette explicabilité doit bénéficier non seulement à l'utilisateur de l'algorithme mais également à la personne qui fait l'objet d'une décision dictée ou assistée par l'algorithme,
- préciser le champ d'application de cet objectif d'explicabilité, qui doit intervenir dès la conception de l'algorithme et en aval, pour assurer la lisibilité des résultats d'audits et de tests.

4. Relever le niveau de la transparence sur l'usage des algorithmes en imposant la communication aux personnes concernées d'un document complet et opposable sur le système d'IA mis en oeuvre

Compléter l'**article 13** sur la transparence, en y ajoutant la formalisation d'une obligation d'information - contractuelle voire précontractuelle, à la charge de l'employeur ou du donneur d'ordre, allant plus loin que la simple information de la mise en oeuvre d'un traitement algorithmique et imposant à l'employeur ou au donneur d'ordre de fournir à la personne faisant l'objet du traitement algorithmique un document complet opposable sur les données utilisées, la nature, la finalité, et les conséquences pour elle de ce traitement algorithmique.

5. Consacrer un droit à la décision humaine, plus effectif et plus large que le droit de ne pas faire l'objet d'une décision purement algorithmique

Compléter l'**article 14** pour introduire un véritable droit pour le travailleur à la décision humaine, imposant à l'entreprise qui met en oeuvre l'algorithme un droit pour le travailleur à la décision humaine, plus large et plus effectif que le droit que les travailleurs tiennent déjà de l'article 22 du RGPD.

6. Renforcer l'obligation de conformité en impliquant les organismes notifiés de manière systématique pour les systèmes d'IA à haut risque

Compléter l'**article 17** en imposant pour les systèmes d'IA à haut risque, une évaluation par des organismes notifiés *ex ante* ou *ex post*, parallèlement à l'obligation déjà prévue d'une auto-évaluation par les fournisseurs de systèmes d'IA.

7. Ajouter à l'obligation de conformité la phase de construction des bases de données d'apprentissage des algorithmes

A l'**article 17**, relatif au système de gestion de la qualité des systèmes d'intelligence artificielle, insérer la phase de construction des bases de données d'apprentissage (mise en qualité et enrichissement) dans le périmètre de l'obligation de conformité prévue par le Règlement pour l'IA à haut risque (conformité *ex ante* et *ex post* avec auto-évaluation par le fournisseur et évaluation par un organisme notifié).

8. Renforcer les mécanismes de traçabilité des algorithmes pour inclure l'ensemble des décisions prises ou recommandées

Préciser les termes de l'**article 20** du projet de Règlement en définissant les activités de l'algorithmes soumises à la génération automatique d'un journal et affirmer la finalité du journal ainsi produit : assurer la traçabilité des décisions et du travail algorithmique pour préserver les droits des travailleurs dont l'activité en est impactée.

Cette précision devrait s'articuler avec celle relative à l'explicabilité, en ce que les journaux ainsi générés devraient également être soumis à l'impératif d'explicabilité dans un souci de lisibilité de l'information pour ses destinataires, au rang desquels doivent figurer les travailleurs qui en font la demande.

9. Introduire des fiduciaires de données pour mutualiser et améliorer la qualité des données d'apprentissage des algorithmes

Introduire au **chapitre 4 du titre III (articles 30 à 39)**, relatif aux organismes notifiés, des “data trusts” ou fiduciaires de données. Ces organismes peuvent être créés au niveau national et éventuellement à l'échelle sectorielle mais leur consécration dans le Règlement constituerait une avancée notable.

10. Ajouter aux exigences applicables aux systèmes d'IA à haut risque une évaluation des algorithmes sur le modèle du stress test bancaire

Ajouter au **chapitre 5 du Titre III** (articles 40 à 51) des procédures d'estimation des algorithmes sur le modèle des stress tests bancaires.

11. Compléter la liste des activités à haut risque liées à l'IA au travail (activités affectant la rémunération des travailleurs)

Élargir le périmètre des activités considérées comme “à haut risque”. L'**Annexe III** intègre déjà les systèmes d'IA mis en œuvre en matière de recrutement pour assister aux décisions de recrutement elles-mêmes, et en matière de gestion de la main d'œuvre. Cependant les décisions liées à la rémunération ne sont pas expressément prévues et devraient selon nous faire l'objet d'une mention explicite.

Annexe 2 - Références bibliographiques

- 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. et Greene D. (2019) Platforms at work: automated hiring platforms and other new intermediaries in the organization of work, in Vallas S.P. et Kovalainen A. (dir.) Work and labor in the digital age, Bingley, Emerald Publishing, 61-91.
- Ajunwa I., Crawford K. et 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. et De Stefano V. (2021) [Frankly, my rider, I don't give a damn](#), Rivista il Mulino, 7 janvier 2021.
- Aloisi A. et 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 mars 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., Cléménçon S., Waelbroeck P. (2019), "[Algorithmes : biais, discrimination et équité](#)", TelecomParisTech et Fondation Abéona
- Bodie M.T., Cherry M.A., McCormick M.L. et 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), "[Algorithmes : contrôle des biais S.V.P.](#)", Institut Montaigne
- Cappelli, Peter and Tambe, Prasanna and Yakubovich, Valery, [Artificial Intelligence in Human Resources Management: Challenges and a Path Forward](#) (April 8, 2019)
- Chagny O. et 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. Capitalisme, algorithmes et démocratie"
- Comité économique et social (2017) Avis du Comité économique et social européen sur « L'intelligence artificielle : les retombées de l'intelligence artificielle pour le marché unique (numérique), la production, la consommation, l'emploi et la société » (avis d'initiative), Journal officiel de l'Union européenne, 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. et 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), "[Management algorithmique et négociation collective](#)", Note de perspectives, 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 juin 2018.
- Fischbach K., Gloor P.A., Lassenius C., Olguin Olguin D., Pentland A.S., Putzke J. et 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), "[Face aux plateformes, il faut un véritable droit social de l'algorithmie](#)", Politis
- Giniaux-Kats J., Campanini E., Dufour M., Forestier F. (2021), "[Le management algorithmique, nouveau paradigme d'organisation du travail ?](#)", AOC
- Gonfalonieri A. (2020) [What brain-computer interfaces could mean for the future of work](#), *Harvard Business Review*
- Guszczka, 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. et Upchurch M. (2018) Digitalisation of work and resistance, in Moore P.V., Upchurch M. et Whittaker X. (dir.) Humans and machines at work, Cham, Palgrave Macmillan, 2018, 17-44.
- OCDE (2019), [Recommendation of the Council on Artificial Intelligence](#), OECD/LEGAL/0449 (May 12, 2019),
- OCDE (2021), [The impact of AI on the labour market: what do we know so far?](#), Janvier 2021
- O'Donovan C. (2018) [An invisible rating system at your favorite chain restaurant is costing your server](#), BuzzFeed.News, 21 juin 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 décembre 2019.
- Satariano S. (2020) [How my boss monitors me while I work from home](#), The New York Times, 7 mai 2020.
- Schubert C. et 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 mars 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

en partenariat avec

**GENEVA
MACRO
LABS**