

L'INTELLIGENCE ARTIFICIELLE
EN QUÊTE D'ACCEPTABILITÉ
ET DE CONFIANCE

Contribution au groupe de travail 3.2
« Anticiper les impacts économiques
et sociaux de l'intelligence artificielle »

Auteurs :

Laurence Devillers, Sorbonne/CNRS, CERNA

Danièle Bourcier, CNRS, CERNA

Stéphane Buffat, médecin, IRBA

Alexei Grinbaum, CEA, CERNA

Arnaud Lavit d'Hautefort, cabinet Ytae

Christophe Lazaro, université de Namur (Belgique), CERNA

Mars 2017

Sommaire

I. Champ d'application	4
II. Impacts éthiques et sociaux	6
III. Recommandations	9

Les opinions exprimées dans ce document engagent leurs auteurs et n'ont pas vocation à refléter la position du gouvernement.

De manière générale, il faut souligner le caractère ambivalent de toute technologie qui offre de nouvelles opportunités en même temps qu'elle crée des risques. Il ne faut pas négliger les aspects éthiques du développement de ces technologies. Les systèmes d'intelligence artificielle deviennent aujourd'hui capables d'accomplir certaines activités qui auparavant étaient l'apanage exclusif des humains. Le développement de ces systèmes conduit au renforcement progressif de leurs facultés d'autonomie propre et de cognition – la capacité à apprendre par l'expérience et à prendre des décisions de manière indépendante –, lesquelles sont susceptibles de faire de ces systèmes des agents à part entière pouvant interagir avec leurs opérateurs et leur environnement et les influencer de manière significative. Ce développement technologique est de nature à transformer considérablement la sphère de l'emploi et du travail, au point d'engendrer une nouvelle phase de révolution informatique dans l'industrie et dans les services. Celle-ci aura notamment pour conséquence la multiplication des systèmes apprenants dans l'environnement de travail, l'intensification des collaborations entre humains et machines ainsi que la délégation croissante de tâches et d'activités autrefois réalisées par les humains à des machines. Ces collaborations entre humains et machines vont revêtir toutes sortes de formes et seront différentes des collaborations classiques.

Cette conjoncture exige qu'on étudie les effets potentiels de l'IA en termes d'éthique et d'acceptabilité sociale. À ce titre, il est nécessaire de soulever toute une série de questions et de s'interroger sur la formation à l'IA, la transformation du travail grâce à l'IA, l'acceptabilité des mutations et la régulation de l'IA. Voici une liste de questions que la société doit se poser avant de décider du pourcentage d'IA susceptible d'accroître la rentabilité. En effet chaque personne doit rester le plus possible utile à la société par son travail ou ses compétences :

- **Comment l'IA va-t-elle modifier la conception de l'emploi et du travail ?**
- **Que faut-il préserver des valeurs du travail humain ?**
- **Les humains vont-ils accepter de travailler avec des IA ou sous les ordres de systèmes d'IA ?**
- **Serons-nous toujours ceux qui supervisent/surveillent ?**
- **Comment évaluer la responsabilité d'un agent humain ou numérique dans un monde où l'humain n'aura désormais plus le monopole de la décision et de l'action ?**
- **Comment créer les conditions d'une coordination des activités et d'une collaboration saines entre l'homme et les systèmes d'IA ?**

Cette annexe au rapport *Anticiper les impacts économiques et sociaux de l'intelligence artificielle* entend proposer un certain nombre de pistes de réflexion destinées à construire à court et moyen terme une relation de confiance entre les travailleurs, les organisations et les systèmes d'IA. Elle a également pour objectif de formuler diverses recommandations visant à garantir le respect des valeurs humaines aux stades de la conception, de l'acceptation, de l'appropriation et de l'utilisation de ces systèmes.

I. Champ d'application

a) Les métiers bouleversés par l'IA

L'IA va bouleverser de nombreux métiers, comme l'ont fait les machines-outils lors des révolutions industrielles précédentes. D'innombrables tâches, actions et fonctions pourront être réalisées et supervisées par une IA. Des agents autonomes participent déjà à différents types de négociations, de transactions (exemple : transactions boursières), d'interactions avec les clients, les usagers, les entreprises et organisations. À court terme vont être affectés de manière encore plus large les métiers du calcul et ceux qui suivent une procédure formellement établie. Certains métiers pourront être entièrement remplacés par des machines autonomes (exemple : conducteurs) tout en créant des nouvelles tâches de supervision. D'autres resteront entièrement le propre de l'homme car l'IA sera considérée comme inacceptable, inefficace ou non bénéfique pour la société. Dans d'autres cas encore, des formes inédites de coopération homme-machine verront le jour dans des environnements de travail où les IA seront des partenaires, des assistants plutôt que des remplaçants. De nouveaux métiers vont également apparaître comme ceux de concepteur, d'entraîneur ou de gestionnaire (cogniticien) de données et d'informations, surveillant ou encore réparateur d'IA, etc. **La valeur du travail ne sera plus le faire mais le contrôle, la vérification, l'approbation et la validation.**

Exemple : *on peut imaginer des phénomènes de retour à des métiers/tâches avec une réalisation partiellement manuelle et des résultats apparents. Vincent de Gaulejac : « Cette tendance peut s'expliquer par deux éléments. C'est un double changement, à la fois vers le manuel : de l'abstrait vers le concret. Et il y a cette volonté de posséder à nouveau ce que l'on fait. Les cadres ont de plus en plus l'impression d'être pris dans un univers où leur travail ne fait plus sens. »* Paru sur le site internet Atlantico.

La rémunération reviendra aux concepteurs d'un procédé ou aux entraîneurs, aux vérificateurs et non pas à celui qui l'exécute car c'est la machine qui exécutera. La signature et la responsabilité finales pourraient en revanche rester humaines, c'est un choix. **Les systèmes d'IA ne prendront pas la place des humains dans toutes leurs tâches et fonctions. Il sera également nécessaire de créer de nouveaux métiers d'entraides sociales où l'humain sera au centre.**

Cette annexe met l'accent sur les métiers de service et vise, en particulier, à appréhender l'impact des IA sur certains secteurs de l'économie qui semblent plus « protégés » que les autres, mais qui ne sont pas pour autant à l'abri de changements significatifs.

Ces secteurs concernent les métiers revêtant une dimension :

- (1) **affective** (exemple : soins infirmiers, assistance sociale, psychothérapie) ;
- (2) **corporelle** (exemple : médecine, physionomie, analyse kinésique) ;
- (3) **mimétique** (exemple : activités sportives, éducation) ;
- (4) **téléologique** (exemple : management, justice, prêtrise) ;
- (5) **créative** (exemple : activités artistiques et de recherche).

Ces différentes dimensions peuvent bien sûr se retrouver au sein d'une même activité ou profession.

b) Les phases de transformation du travail

Il existe deux échelles temporelles de transformation du travail :

1. Une lente évolution des métiers dans le contexte et l'environnement existants : les choix des utilisateurs portent sur « l'ancien » ou le « nouveau », mais l'ancien reste une norme parfaitement acceptée et disponible.

2. Une intégration totale des fonctions par le nouvel environnement de la vie humaine (ville intelligente, maison intelligente, objets connectés). Les contextes « anciens », sans disparaître complètement, deviennent archaïques ou exceptionnels (exemple : la maison haussmannienne non adaptée aux exigences technologiques modernes, le cheval remplacé par l'automobile). L'environnement « nouveau » s'impose en vertu des choix par défaut faits par la majorité, des décisions du secteur marchand ou des gouvernements. Ces options seront de toute façon stimulées par l'offre du marché. Dans la deuxième phase, les métiers subissent un changement de rupture, tandis que, dans la première, les compétences évoluent graduellement. Le passage de la première étape à la seconde n'est pas automatique, mais dépend de la réaction de la société, des changements générationnels et du niveau d'éducation.

L'acceptabilité des technologies apparaît dans ce processus comme un va-et-vient permanent entre les innovateurs et les divers types d'utilisateurs des nouvelles technologies.

- **Notion d'acceptabilité** : acceptabilité sociale (avant la mise sur le marché, tient compte de la pluralité des usages et perceptions) : essayer à travers des méthodes qualitatives (par exemple des scénarios) de cerner les questions d'usage que se pose le public : focus group, feedback à l'ingénieur.

Remarque : *quel choix va réellement s'offrir à l'individu : un choix social, d'éducation et de formation ?*

- **Notion d'appropriation** : sur le plan individuel, on peut définir l'appropriation comme un processus reposant sur l'acquisition progressive d'une certaine maîtrise technique et cognitive des objets techniques ainsi que sur l'intégration significative de ces objets dans la pratique quotidienne.

Remarque : *il y a deux niveaux d'appropriation celui de l'interface de médiation humain-IA et celui de la compréhension fine des mécanismes sous-jacents. Cela peut être intéressant en cas de scénario « d'accident » pour décider de la responsabilité.*

II. Impacts éthiques et sociaux

Nous discuterons les concepts et les idées dont la signification est susceptible d'évoluer sous l'effet de l'IA (responsabilité, mérite et valeur, créativité), puis les effets positifs (moins d'oubli, moins de perte de temps, prise en charge de tâches très répétitives ou des routines procédurales sans intérêt, etc.) ; enfin, les effets susceptibles d'être négatifs (manque d'exercice de la mémoire, etc.) si la société ne met pas de garde-fous et ne régule pas l'IA.

a) *Changements de concepts*

1. Responsabilité

L'entraîneur est celui qui réalise l'apprentissage du système d'IA en lui fournissant des données. L'entraîneur peut être le concepteur de la machine mais aussi celui qui a construit le corpus de données ou les utilisateurs conscients ou non conscients du fait que la machine apprend au cours de son utilisation. **La responsabilité sera partagée entre le concepteur, l'entraîneur et l'utilisateur selon des modalités qui restent à établir. Cela nécessite de modifier le droit pour prendre en compte ce type de partage.**

2. Mérite et valeur

La valeur du travail ne sera plus le *faire* mais le *contrôle*, la *vérification*, l'*approbation* et la *signature*. La rémunération reviendra au **responsable** d'un procédé et pas à celui qui l'exécute (car la machine exécutera).

3. Créativité

Un système d'IA convenablement programmé peut créer de l'innovation par expérimentation ou par sérendipité, mais seul l'humain donne un sens à la réalité nouvelle. Il est important de distinguer les créations des machines et les créations humaines.

b) *Risques*

Les aspects positifs de l'IA sont nombreux à condition de respecter un certain nombre de valeurs : une mémoire élargie, moins de perte de concentration ou d'oubli d'information pertinente, un conseil expert plus efficace, moins subjectif ou arbitraire, mieux justifié, etc., **à condition que la machine soit neutre, loyale et non discriminante.** Nous ne développons pas ici tous ces points déjà suffisamment soulignés et discutés par les acteurs de l'IA, mais nous analysons les risques éthiques. À chaque fois, il faut réfléchir aux aspects positifs et plus négatifs des situations et trouver les conditions pour transformer ces risques en atouts.

1. Désengagement : l'utilisation de l'IA et de machines autonomes peut conduire à un *désengagement* de l'humain. Remplacer le facteur humain par de l'IA pourrait conduire à une déshumanisation des pratiques et un appauvrissement des interactions sociales, lesquels constituent très souvent la raison d'être de certains métiers. C'est pour cela qu'il s'agit non pas de remplacer l'humain pour toutes les tâches sociales (comme par exemple dans la santé) mais de trouver les garde-fous pour accompagner au mieux la collaboration humain-machine. N'oublions pas que nombre de « décisions » sont actuellement prises

automatiquement sans que le destinataire en soit informé. L'IA peut restituer au moins les traces et les raisonnements activés.

Exemple : le secteur de la santé et, en particulier celui des soins apportés aux patients, aux personnes en situation de handicap et aux personnes âgées va, dans les prochaines années, être profondément transformé par le développement de technologies offrant une grande diversité d'applications allant de l'assistance au contrôle, jusqu'à l'accompagnement quotidien. La forte dimension affective au cœur de la relation de soin va évoluer et, dès lors, va modifier à la fois le travail des aides-soignants et thérapeutes ainsi que la vie des patients. Sans contester l'utilité indéniable des IA et des agents autonomes (robot portant, robot soignant ou robot coach) pour faciliter et soulager le travail des prestataires de soins, il est important de souligner que le contact humain est l'un des aspects essentiels de la relation de soin.

2. Déresponsabilisation : dans la continuité du point précédent, l'utilisation de l'IA et de machines autonomes peut entraîner une *déresponsabilisation* du travailleur, susceptible de ne plus devenir qu'un simple relais de la décision de la machine. Dans les cas de plus en plus nombreux de décisions prises sur la base d'un traitement automatisé de données, l'être humain pourrait avoir tendance à s'en remettre exclusivement à la proposition de la machine, en évitant d'engager sa responsabilité. Dévier de la solution préconisée par la machine entraînerait une prise de risque trop grande, susceptible de lui faire encourir d'éventuelles sanctions en cas de problème.

Exemple : le banquier qui s'en remettrait uniquement à la décision suggérée par la machine sur la base d'un profil pour accorder ou refuser un prêt à un client. Il y a également un problème de représentation des échelles, dans le trading par exemple, et la question de la numérosité à travers l'IA semble intéressante à creuser.

3. Transformation du mérite du travail : l'utilisation de l'IA et de machines autonomes peut bouleverser radicalement l'évaluation des performances des travailleurs et l'*idée de mérite*. En transformant les modalités de réalisation des tâches, les technologies feront évoluer les évaluations du mérite individuel et donc de l'équité des bénéfices induits par l'effort et des récompenses dues en vertu des talents acquis par l'individu.

Exemple : comment évaluer les performances d'un joueur d'échecs, d'un chirurgien ou d'un trader assisté d'une IA ? Qui est à la source de la performance ? Est-ce la technologie, le talent propre de l'humain ou l'association des deux ? Leur récompense (que ce soit en termes de salaire, de réputation) est-elle juste ? Est-elle méritée ?

4. « Prolétarianisation » des savoirs et des savoir-faire : l'utilisation de l'IA et de machines autonomes peut contribuer à une « *prolétarianisation* » des savoirs et des savoir-faire. À force de ne plus être entièrement partie prenante dans l'élaboration et l'expression d'une décision dans l'environnement de travail, les humains risquent de se voir déposséder de leur expertise en terme de *know-how*, et de perdre un ensemble de capacités et de compétences, qui non seulement peuvent être utiles à la société, mais qui contribuent aussi à alimenter le respect de soi. Ce risque pourra survenir si l'humain n'est pas associé à l'entraînement et à l'évaluation des modèles et au choix des données pour améliorer leur

performance. **Se pose la question d'un conservatoire des métiers, avant l'émergence massive des intelligences artificielles.**

Exemple : le douanier dans un aéroport auquel on ne demandera plus d'exercer ses compétences et son sens de la physionomie acquis au fil des années à travers une expérience sur le terrain et qui devra s'en remettre aux prédictions de la machine en matière de profils de dangerosité.

5. Autonomie et créativité des travailleurs : l'utilisation de l'IA et de machines autonomes peut altérer *l'autonomie et la créativité des travailleurs*. Derrière les vertus facilitatrices de certains dispositifs peut se déployer de manière sous-jacente une normativité, prenant la forme d'un « paternalisme technologique » aux multiples visages : ainsi, les IA alertent, recommandent, rappellent à l'ordre, bloquent, interdisent, voire influencent.

Exemple : pour des raisons légitimes de sécurité ou d'organisation, certains dispositifs pourraient empêcher les êtres humains d'adopter des comportements considérés comme « sub-optimaux », de tenter certaines expériences, voire de commettre des erreurs qui souvent sont à la source de nouveautés et de découvertes.

6. Vie privée et intimité au travail : l'utilisation de l'IA et de machines autonomes peut nuire à la *vie privée et à l'intimité* des travailleurs en imposant un contrôle et une surveillance disproportionnés de leurs activités. Sans contester les bénéfices que peuvent procurer les IA dans la réalisation de multiples tâches, leur utilisation risque d'être mise en œuvre de manière abusive afin d'évaluer la productivité et le rendement des travailleurs et de les soumettre à un contrôle strict.

Exemple : pour les avocats dans un cabinet d'avocats ou les journalistes dans un groupe de presse et pour de nombreux autres métiers.

7. Épanouissement dans le travail : l'utilisation d'IA et de machines autonomes peut nuire à *l'épanouissement* des travailleurs sur le lieu de travail en perturbant le rythme de travail, les routines et en battant en brèche les conditions qui permettent de se réaliser dans le travail.

Exemple : le degré d'appréciation de la dangerosité d'un conducteur par un policier n'aura plus de valeur face aux contrôles automatiques. L'interruption du processus de pensée par une sollicitation non hiérarchisée de stimulation venant de plusieurs agents peut contribuer à désorganiser, voir à annihiler la capacité à prendre des décisions (*task switching*).

8. Disparités sociales affaiblies ou aggravées : l'utilisation de l'IA et de machines autonomes peut conduire à *affaiblir certaines disparités sociales* (exemple : valeur du « capital culturel ») *ou au contraire les aggraver* (les personnes non initiées perçoivent encore moins l'utilité de l'acquisition de ces connaissances à l'apparence « superflue »).

Exemple : la machine mémorise pour le travailleur, du coup celui-ci fait de moins en moins d'efforts pour mémoriser les informations nécessaires pour résoudre un problème, comprendre et surveiller ce que fait la machine.

9. Pratique des managers et des représentants du personnel : dans une société qui se numérise, les pratiques du digital tendent à redéfinir les modèles d'interaction entre les individus. En entreprise, le numérique contribue à *redéfinir la pratique des managers et des représentants du personnel*, qui jouent un rôle majeur d'interface entre la direction et les salariés.

Exemple : DIGILABSOCIAL - Digital Management & Relations sociales

Bénédicte Tilloy, secrétaire générale et DRH de SNCF Réseau, et Jean-Paul Bouchet, secrétaire général de l'Union confédérale des cadres de la CFDT, ont pris l'initiative de lancer avec Sciences Po et le cabinet Ytae un laboratoire d'idées permettant de mieux dessiner les contours de ce que seront demain les fonctions de manager et de représentant du personnel dans une entreprise digitalisée, et de proposer des pratiques à expérimenter. Le laboratoire associe, à parts égales, une cinquantaine de managers et de représentants du personnel issus d'entreprises appartenant à des secteurs d'activité différents.

10. La « normalisation » d'une nouvelle technologique (au sens de Kuhn, où elle cesse d'être vue comme révolutionnaire) passe toujours par la normalisation (= standardisation). Or avec l'IA (comme avec les nanotechnologies), nous avons une technologie transversale qui ne se limite pas à un secteur industriel donné. D'où le dilemme suivant : soit on instaure des normes par secteur, soit on crée des « normes d'IA » pour tous les secteurs.

Exemple : la France a expérimenté la dernière approche avec les nanotechnologies et cela a très mal marché. En revanche, l'Europe fait de la normalisation par secteur avec les nanotechnologies et cela se présente plutôt comme un succès.

11. Transformation des capacités humaines : les facilités qu'apporte la machine, par exemple sur la mémorisation, vont **transformer certaines des capacités humaines**. Devons-nous anticiper ce besoin dans l'enseignement ? (en soulignant l'utilité d'exercices sollicitant la mémoire court terme et long terme (exemple : calcul mental)).

Remarque : « *Aujourd'hui, j'oublie sans conséquence car la machine (téléphone, ordinateur) me rappelle ce que j'ai oublié. De ce fait, je n'exerce pas ma mémoire, mais mon attention aux indices de rappel... la machine aide à se souvenir mais rend dépendant* » (<https://lejournel.cnrs.fr/articles/le-numerique-nous-fait-il-perdre-la-memoire>).

III. Recommandations

Afin de répondre aux impacts éthiques et sociaux des systèmes issus de l'IA et de faciliter leur adéquation et leur appropriation par tous les acteurs du monde du travail, il est important de réfléchir dès maintenant aux différentes étapes et conditions pour instaurer un climat de confiance.

La confiance se décline à un triple niveau :

(1) **la conception :** certaines valeurs doivent être prises en compte dès la phase de conception technologique et intégrées dans celle-ci (« value by design ») ;

(2) ***l'appropriation*** : il est nécessaire de faciliter l'apprentissage par l'humain et son acquisition progressive de la maîtrise technique et cognitive des dispositifs techniques ainsi que l'intégration significative de ces dispositifs dans la pratique quotidienne ;

(3) ***l'utilisation*** : il faut informer les utilisateurs de l'origine automatique ou non de la décision, anticiper les conséquences possibles de l'adoption de ces systèmes de l'IA par les êtres humains et cerner en quoi ces systèmes vont redéfinir les modalités concrètes du travail.

Afin de garantir une transition plus aisée de la recherche à la commercialisation, nous proposons au gouvernement de promouvoir des programmes de recherche qui intègrent des mécanismes de vérification des résultats à court et moyen terme afin de comprendre les opportunités et les risques réels qui sont associés au déploiement et à la dissémination de ces technologies. La formation à l'IA et aux mutations est essentielle. La conduite du changement doit être menée en bonne intelligence avec tous les acteurs : dirigeants de société et syndicats.

Des ressources suffisantes doivent être consacrées à la résolution des défis éthiques et sociaux soulevés par le développement et les applications de ces technologies.

De manière générale, l'approche développée doit être graduelle, pragmatique et prudente. Dans cette perspective, nous formulons diverses recommandations.

1. Code d'éthique : l'éducation des ingénieurs à la réflexion éthique, par toutes les méthodes pédagogiques y compris le développement d'un code d'éthique, est nécessaire pour encadrer le développement et la production des systèmes d'IA. À moyen et long terme, cela complètera de façon réfléchie et informée le cadre légal existant à l'échelle nationale et européenne. Cette réflexion éthique des ingénieurs doit être fondée sur les principes de bienfaisance ou non-malfaisance, d'autonomie et de justice, ainsi que sur les principes et valeurs consacrés dans la Charte européenne des droits fondamentaux. **Pourquoi pas une déontologie des ingénieurs en IA ?**

Objectif 1 : « **la formation** » afin d'augmenter la compréhension de l'éthique par les ingénieurs.

Moyen : intégrer l'éthique dans le maquettage des formations.

Coûts : ECTS.

Évaluations : applicabilité de la formation à l'utilisation des ressources de l'éthique.

Objectif 2 : « **les ressources de l'éthique** » afin d'accompagner la conception de l'IA, et la réévaluation nécessaire.

Moyen : rédaction de chartes éthiques, méthodologie d'évaluation, mise en œuvre d'outils pour valider le suivi des chartes dans différents secteurs, par exemple un robot d'aide à la personne, aide au diagnostic, benchmarks sur les capacités des IA.

Coûts : ETP + plateformes/réseaux.

Évaluations : degré d'utilisation, appropriation (degré), mise à jour.

2. Éducation : avec l'arrivée de l'IA, une révolution sociétale est en marche qui nécessite de la part des dirigeants de définir un nouveau projet de société clair et apparent : faire évoluer les formations initiales et l'enseignement dans les écoles et lycées ainsi que les formations aux métiers qui vont être bouleversés.

3. Surveillance et formations : il faut surveiller de près le marché du travail, spécialement la création et la perte d'emplois et de savoir-faire dans différents champs professionnels. Il est fondamental d'identifier le plus tôt possible divers tendances et de chiffrer le taux de remplacement par l'IA dans différents secteurs de l'automatisation. **Les comités d'entreprise et les organismes de représentation des travailleurs (syndicats) devront être associés à ces études et formés à ces nouveaux enjeux.**

4. Recherches et études de terrain : il est nécessaire de **tester les dispositifs de l'IA dans des situations réelles (« real-life scenarios ») sur le long terme** (ou au moins à plusieurs moments d'utilisation dans le temps) afin d'identifier et d'évaluer les risques qu'ils peuvent engendrer au-delà de la phase d'études à court terme.

5. Implication et information : afin de promouvoir l'acceptabilité sociale des technologies et de faciliter leur appropriation, Il faut **informer les personnes sur les développements technologiques en cours dans la société et les alerter des conséquences potentielles sur leur travail**. Les usagers doivent être adéquatement représentés dans la définition des politiques publiques en matière d'IA et pris en compte dans les processus d'innovation. Les métiers et opportunités de cette révolution ne seront accessibles dans un premier temps qu'aux personnes averties et sensibilisées à ces outils. L'information doit aussi se faire auprès des élèves et étudiants afin de les attirer en amont vers ces formations.

6. Engagement : dans certains secteurs d'activité, il est important de **préserver le contact humain et la dimension affective et corporelle de l'activité**. À ce titre, il faut prévoir d'aménager des « espace-temps » où les interactions sociales sont fortes et où l'implication subjective dans l'activité demeure importante.

7. Prise de décision : il est nécessaire d'amener le travailleur à **coopérer en synergie avec la machine, de le faire participer au processus décisionnel** et de lui conférer un rôle susceptible d'engager sa responsabilité en sollicitant l'engagement personnel et la mobilisation subjective dans le processus de travail. Il faut que le travailleur puisse reprendre le contrôle de la machine (reprise en main) si nécessaire. À cette fin, il est utile de créer de nouveaux métiers tels que « compreneur » ou « interprète » des processus d'apprentissage non supervisé afin de **créer de nouveaux outils que le travailleur pourra utiliser pour connaître l'état de la machine, comprendre son comportement et la surveiller. Il doit aussi pouvoir apporter des suggestions sur des améliorations ou des corrections éventuelles.**

8. Autonomie : préserver l'autonomie et la créativité des individus, notamment en envisageant des possibilités d'improvisation et de spontanéité est une autre priorité. **Dans certains cas, le travailleur doit pouvoir désactiver les dispositifs issus de l'IA** afin d'exercer son activité en toute indépendance et liberté.

9. Intimité et vie privée : il est nécessaire de **garantir la vie privée et une sphère d'intimité dans l'environnement de travail** afin que chacun ait, quand il le souhaite, des moments de répit, propices à la concentration et à la réflexion. Le travailleur et l'utilisateur doivent être informés des données privées collectées ou traitées durant l'utilisation de la machine.

10. Co-travail avec l'IA : dans le co-travail avec une IA, la machine doit respecter les valeurs de l'humain à travers le travail. À cette fin, les travailleurs et les usagers seront formés pour leur permettre de se familiariser avec les exigences techniques de leurs secteurs et de **développer de nouvelles compétences**. Les standards d'excellence ou de productivité devront probablement être revus pour éviter une compétition perdue d'avance sur des domaines où l'humain n'est pas le plus performant.

11. Normalisation : les autorités doivent lancer une concertation en vue du développement des normes techniques et industrielles pour le déploiement de l'IA. Cela doit comporter des normes spécifiques par secteur industriel ainsi que des standards généraux portant sur les aspects transversaux de l'IA.

12. Expérimentation : création de laboratoires ou plateformes d'expérimentation des nouveaux métiers, sur l'exemple qui est cité (DIGILABSOCIAL).