

## LA SST ET L'AVENIR DU TRAVAIL: AVANTAGES ET RISQUES DE L'UTILISATION D'OUTILS D'IA DANS UN ENVIRONNEMENT DE TRAVAIL

### Introduction

Le concept d'intelligence artificielle (IA) apparaît en 1956, lors d'une série d'ateliers académiques organisés au Dartmouth College, une université privée située dans le New Hampshire, aux États-Unis. Lors de cette conférence, un groupe de scientifiques se propose d'apprendre à des machines à utiliser le langage, à formuler des concepts, à s'améliorer (en tant que machines) et à résoudre des problèmes initialement «réservés aux humains» (McCarthy et al., 1955). John McCarthy et ses collègues nourrissaient l'ambition de parvenir à ce résultat en quelques semaines seulement. Si la conférence n'a pas été un grand succès en soi, elle a toutefois marqué les débuts des travaux de recherche et de développement sur l'intelligence artificielle.

Bien que ces premiers pas puissent faire sourire aujourd'hui par leur optimisme naïf, l'intérêt pour l'intelligence artificielle n'a pour autant pas disparu. Au contraire, les débats et les expérimentations relatifs à l'IA ont suscité l'engouement et connu des hauts et des bas, passant de l'espoir que les machines puissent être entraînées pour se comporter exactement comme les êtres humains et atteindre un niveau d'intelligence égal, comme l'imaginaient les participants aux ateliers de Dartmouth, aux affres de la désillusion. Les premiers robots expérimentaux, tels que «WOBOT» et «Shakey», n'ont pas permis d'atteindre l'objectif d'intelligence artificielle universelle qu'ils promettaient. La recherche en la matière connaît deux périodes sombres, de 1974 à 1980 et de 1987 à 1993, baptisées les «hivers de l'intelligence artificielle» et marquées par l'échec de plusieurs expérimentations et la chute des financements. L'année 2019 marque cependant un regain d'intérêt pour la discipline.

Les pays développés accordent aujourd'hui des budgets importants, se chiffrant en milliards, à la recherche et au développement de l'intelligence artificielle, avec en tête les États-Unis, suivis de près par la Chine et Israël (Delponte, 2018). Selon les prévisions, l'IA devrait permettre au produit intérieur brut (PIB) chinois de faire un bond de 26 % d'ici 2030. Pour l'Amérique du Nord, la hausse prévue serait de 14,5 % (PwC, 2018a), certains prédisant également que l'intelligence artificielle permettrait de créer autant d'emplois qu'elle en éliminerait (PwC, 2018b). Les prévisions des sociétés de conseil et des groupes de réflexion rejoignent une série de rapports de haut niveau émis par des organismes gouvernementaux, régionaux et internationaux qui prédisent l'impact majeur de l'intelligence artificielle sur les économies et les sociétés, notamment aux États-Unis (Bureau des sciences et technologies de la Maison Blanche (OSTP), 2018), au Royaume-Uni [département des Affaires, de l'Énergie et des Stratégies industrielles (BEIS), département du Numérique, de la Culture, des Médias et du Sport, 2018)], au sein de l'Union européenne (Commission européenne, 2018) et avec l'Organisation internationale du travail (OIT).

Dans la plupart des cas, ces rapports de haut niveau indiquent que l'intelligence artificielle permettra d'améliorer la productivité. Les discussions sur la productivité concernent les implications directes de l'IA pour les travailleurs et leurs conditions de travail, bien entendu, mais peu se penchent sur la question de savoir si l'introduction de l'IA sur les lieux de travail sera à l'avantage ou au contraire en défaveur des travailleurs eux-mêmes en ce qui concerne la santé et la sécurité au travail (SST). Afin de jeter les bases de ce rapport d'expertise, qui traite précisément de cette question jusqu'ici peu abordée, nous commencerons par examiner la signification du concept d'intelligence artificielle, afin de partir d'une base claire pour débattre de son incidence sur les travailleurs. Dans le chapitre 2, nous présenterons les différentes utilisations de l'intelligence artificielle dans divers outils et applications de travail assisté, ainsi que dans les processus de prise de décision professionnels, de même que les risques et les avantages en découlant en matière de SST. Nous commencerons par les ressources humaines (RH), avec les «people analytics» et les entretiens vidéos, et poursuivrons par l'intégration d'outils robotiques à intelligence augmentée par l'IA, avec notamment les robots collaboratifs (cobots) et les chatbots (dialogueurs) utilisés dans les usines, les entrepôts et les centres d'appel. Nous identifierons

ensuite les usages possibles pour les technologies «prêt-à-porter» et les tablettes assistées sur les chaînes d'assemblage en production, puis nous étudierons les processus algorithmiques à l'œuvre dans la *gig economy*. Dans le chapitre 3, nous présenterons les réponses des parties prenantes internationales aux nouveaux risques et avantages associés à l'IA sur le lieu de travail. Dans sa conclusion, au chapitre 4, le rapport propose diverses recommandations afin de mieux gérer et atténuer les risques les plus critiques pouvant découler de l'utilisation de l'intelligence artificielle dans le monde du travail.

## Qu'est-ce que l'intelligence artificielle (IA)?

La définition exacte des concepts que recouvre le terme d'intelligence artificielle fait aujourd'hui débat. Il existe d'ailleurs un véritable battage autour de l'intelligence artificielle, qui surpasse la réalité même. Toutefois, alors que les gouvernements investissent massivement des capitaux dans la recherche et le développement, publiant des rapports de haut niveau sur l'impact majeur qu'aura l'IA sur le PIB et la productivité, mieux vaut prendre le sujet de l'IA au sérieux. Le débat autour de l'authenticité de l'IA est toutefois pertinent. Prenons donc le temps de définir clairement ce qu'est l'intelligence artificielle, afin d'éviter toute confusion dans la suite de ce rapport. McCarthy et ses collègues, dont mention a été faite dans l'introduction, ont défini la «question de l'intelligence artificielle» comme étant «le moyen de fabriquer une machine se comportant de manière telle qu'elle serait considérée comme intelligente s'il s'agissait d'un être humain» (McCarthy et al., 1955). Les auteurs de la conférence de Dartmouth étant les inventeurs du concept d'intelligence artificielle, il n'est pas inutile de rappeler leur définition. Les machines peuvent-elles se comporter comme des êtres humains? Bien que répondre à cette question philosophique ne soit pas l'objet de cet article, il est intéressant de noter que des questions plus larges relatives aux êtres humains et à notre relation aux machines ont été au cœur des premières orientations de la recherche en matière d'IA (par ex., Simon, 1969; Dreyfus, 1972; Weizenbaum, 1976), et sous-tendent encore les expérimentations et les applications de l'IA aujourd'hui. Une interrogation centrale, évidente, mais rarement énoncée, est au cœur de ces questionnements: *pourquoi* souhaitons-nous que les machines se comportent comme nous, voire mieux que nous? D'un point de vue social, que nous manque-t-il pour qu'une telle «amélioration» soit nécessaire? Dans le cadre de ce rapport, parmi les nombreuses définitions existantes pour le concept d'intelligence artificielle, nous utiliserons celle établie par McCarthy comme base pour identifier les problématiques émergentes d'un point de vue épistémologique.

La définition de la Commission européenne, telle qu'établie dans sa Communication de 2018, est également adoptée pour le présent rapport. Dans celle-ci, le terme d'intelligence artificielle se réfère à tous «les systèmes qui font preuve d'un comportement intelligent en analysant leur environnement et en prenant des mesures – avec un certain degré d'autonomie – pour atteindre des objectifs spécifiques» (Commission européenne, 2018). Un autre rapport de 2018, intitulé *European artificial intelligence leadership, the path for an integrated vision* (Leadership européen en matière d'intelligence artificielle – Vers une vision intégrée), affine cette définition, le terme d'intelligence artificielle devenant un «terme générique regroupant l'ensemble des techniques associées à l'analyse des données et à la reconnaissance des formes» (Delponte, 2018, p. 11). Ce rapport, commandité par la Commission de l'industrie, de la recherche et de l'énergie du Parlement européen, distingue l'intelligence artificielle des autres technologies numériques, dans le sens où «l'IA doit apprendre de son environnement afin de prendre des décisions autonomes» (Delponte, 2018, p. 11). Ces définitions établissent une base claire pour les discussions sur les enjeux de cette technologie, alors que les machines et les systèmes IA sont de plus en plus intégrés à l'environnement de travail, où ils affichent des compétences décisionnelles et prédictives bien plus rapides et précises que les humains, tout en offrant un comportement semblable à celui des êtres humains et des capacités d'assistance aux travailleurs.

Les experts classent l'intelligence artificielle en deux catégories: l'intelligence artificielle forte et l'intelligence artificielle faible. L'IA *faible* désigne une machine qui s'appuie sur un logiciel pour guider ses questions et ses réponses. Ce type d'IA n'atteint pas un niveau de conscience ou une «sentience» semblable aux êtres humains, mais elle peut résoudre des problèmes dans un champ d'application donné. Le terme d'IA *faible* désigne donc les systèmes experts, ainsi que les systèmes de reconnaissance de texte et d'image. L'intelligence artificielle forte, également appelée «intelligence artificielle universelle» (Hutter, 2012), se réfère quant à elle à une machine capable

de faire preuve d'un comportement équivalent, voire supérieur, à celui des êtres humains en termes de compétences et d'aptitudes. Il est de tous temps celui qui a le plus fasciné les chercheurs tels qu'Alan Turing. Avant même la conférence de 1956 réunissant McCarthy et ses collègues, Alan Turing s'est posé en 1950 la question de savoir si les machines étaient capables de penser. Une intelligence artificielle universelle existe lorsqu'un agent universel unique est capable d'apprendre en vue de se comporter de façon optimale dans n'importe quel environnement, lorsqu'un robot fait preuve de compétences universelles telles que la marche, la vue et la parole. Aujourd'hui, alors que les capacités mémorielles des ordinateurs ne cessent de croître et que les programmes logiciels deviennent de plus en plus sophistiqués, la création d'une telle intelligence artificielle universelle semble plus plausible que jamais. Ces avancées viendraient compléter les capacités d'automatisation existantes, avec des robots possédant des compétences de travail équivalentes à celles des êtres humains, tout en étant exempts de leurs caractéristiques non productives telles que la fatigue ou la maladie. Le grand public semble plus à l'aise avec le concept d'IA faible, qui permet d'améliorer les capacités des machines afin de leur permettre de jouer un rôle d'assistant pour les humains, plutôt que de les remplacer en tant qu'ouvriers ou d'assumer des fonctions de gestion humaine.

Dans le prochain chapitre, nous étudierons les différents usages de l'intelligence artificielle sur le lieu de travail, ainsi que les risques et les avantages potentiels et avérés de l'IA en matière de SST, sur la base de recherches documentaires et d'une série d'entretiens avec des experts menés par l'auteure.

## 2 L'intelligence artificielle dans l'environnement de travail

Bien que les possibilités de l'IA en matière de progression et d'amélioration de la productivité soient importantes, des questions majeures liées à la santé et la sécurité au travail (SST) se posent lors de l'intégration de systèmes d'IA dans un environnement de travail. Stress, discrimination, précarité accrue, troubles musculosquelettiques, intensification du travail et pertes d'emploi potentielles comptent parmi les risques psychosociaux existants, avec la violence physique, dans les environnements de travail digitalisés (Moore, 2018a). Ces risques sont exacerbés lorsque l'IA sert à «augmenter» des outils technologiques existants ou est introduite pour la première fois sur le lieu de travail aux fins de gestion et de conception. Il est évident que l'intelligence artificielle accroît les risques SST dans les environnements de travail digitalisés, car elle permet une surveillance et un suivi accrus pouvant mener à un phénomène de micro-management, l'une des principales sources de stress et d'anxiété pour les travailleurs (Moore, 2018a). L'IA souligne la nécessité d'accorder davantage de crédibilité, voire d'autorité, à ce qu'Agarwal et ses collègues (2018) appellent les «machines prédictives», c'est-à-dire les outils robotiques et les processus algorithmiques au travail. Il est toutefois à noter qu'aucune technologie seule, en elle-même, n'engendre de risques ou de bénéfices en matière de SST. Il s'agit bien de la *mise en œuvre* des technologies qui crée des conditions positives ou négatives.

### 2.1 L'intelligence artificielle dans les ressources humaines

Dans le domaine de la gestion des ressources humaines, l'une des intégrations IA les plus populaires est appelée «people analytics» (outils analytiques RH), qui consiste principalement à utiliser le Big Data et les outils numériques pour «mesurer, suivre et comprendre les performances des employés, les différents aspects de la planification des effectifs, la gestion des talents et la gestion opérationnelle» (Collins, Fineman et Tsuchida, 2017). L'informatisation, la collecte des données et les outils de contrôle permettent aux organisations d'appliquer «des fonctions analytiques en temps réel, à l'endroit précis où cela est nécessaire dans le processus métier, offrant une compréhension plus approfondie des problématiques et des informations exploitables par l'entreprise» (ibid.). Les algorithmes machine de prédiction appliqués à ces processus se situent souvent dans des «boîtes noires» (Pasquale, 2015), dont peu de personnes comprennent vraiment le fonctionnement; malgré cela, il est octroyé aux programmes informatiques une autorité

suffisante pour leur permettre d'effectuer des «prédictions[s] par exception»<sup>(1)</sup> (Agarwal, Gans et Goldfarb, 2018).

Tous les outils de *people analytics* ne sont pas nécessairement associés à un processus d'IA. Cependant, une réponse de programme intelligente aux équations algorithmiques permet un apprentissage machine, qui génère des prédictions et pose des questions connexes sans intervention humaine, excepté lors de la phase de saisie des données; il s'agit là de processus d'intelligence artificielle au sens établi par l'UE, tel qu'indiqué précédemment. Le Big Data, un phénomène lucratif en croissance forte depuis plusieurs années et représentant un investissement attractif, implique de collecter des données sur tout, tout le temps. Le Big Data porte aujourd'hui ses fruits dans le domaine des RH, alors que les importants ensembles de données désormais disponibles sont utilisés pour former les algorithmes à la réalisation d'analyses et de prédictions concernant le comportement des travailleurs, via l'apprentissage machine, facilitant ainsi la prise de décision managériale. Sur la base des schémas identifiés, l'IA crée un algorithme capable de produire des solutions et des réponses aux demandes relatives aux schémas apparaissant dans les données, et ce bien plus rapidement que ne saurait le faire un humain. Les réponses issues de l'apprentissage machine (apprentissage automatique) sont souvent différentes de celles qu'un humain donnerait, voire serait capable de donner. Les données relatives aux travailleurs peuvent être recueillies auprès de différentes sources, directement sur le lieu de travail ou en dehors, notamment le nombre de clics de clavier, les informations provenant des réseaux sociaux, le nombre d'appels téléphoniques et leur contenu, les sites Web visités, la présence physique, les lieux visités (positionnement GPS) en dehors du lieu de travail, les déplacements au sein des bureaux, le contenu des courriers électroniques, voire même le ton employé pour parler et les mouvements corporels, dans une approche sociométrique (Moore, 2018a, 2018b).

Également appelée «analytique humaine», «analytique des talent» ou «analytique des ressources humaines», à l'ère des RH stratégiques, cette application des outils IA se définit globalement par l'utilisation des données individuelles des personnes pour faciliter le processus décisionnel des managers et des professionnels RH dans le cadre du recrutement, par exemple afin de savoir qui engager, de l'évaluation des performances et de l'octroi des promotions, ou encore pour identifier les employés susceptibles de quitter leur emploi ou sélectionner les futurs dirigeants. Les outils de *people analytics* sont également utilisés pour rechercher des schémas récurrents au sein des données relatives aux employés, notamment afin d'identifier les tendances en matière de présence, de moral des employés et de problèmes de santé au niveau de l'organisation.

Environ 40 % des fonctions RH au sein des compagnies internationales s'appuient aujourd'hui sur l'intelligence artificielle. Ces compagnies sont principalement basées aux États-Unis, toutefois certaines entreprises européennes et asiatiques se lancent également dans la course. Une étude réalisée par PwC démontre qu'un nombre croissant d'entreprises internationales commence à reconnaître la valeur ajoutée de l'IA dans la prise en charge du personnel (PwC, 2018a). Un rapport indique ainsi que 32 % des services du personnel dans les entreprises, technologiques et autres, repensent actuellement leur organisation avec l'aide de l'intelligence artificielle, dans un souci d'optimisation, «à des fins d'adaptabilité et d'apprentissage, de manière à mieux intégrer les informations recueillies auprès des employés et via les technologies à disposition» (Kar, 2018). Une récente 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. Ce rapport indique que les deux tiers des PDG estiment que l'intelligence artificielle sera génératrice de valeur pour les RH (IBM, 2018). Un rapport du cabinet Deloitte indique que 71 % des compagnies internationales estiment que les outils de *people analytics* constituent une priorité élevée pour leur organisation (Collins, Fineman et Tsuchida, 2017), car ils leur permettent non seulement d'obtenir des informations métier précieuses, mais également de prendre en charge les «problèmes humains» (ibid.).

<sup>(1)</sup> Le terme «prédiction par exception» se réfère aux processus permettant aux ordinateurs de gérer de gros ensembles de données, de délivrer des prédictions fiables fondées sur des données régulières et de routine, d'identifier les aberrations, voire même d'envoyer des notifications indiquant à l'utilisateur d'effectuer des contrôles ou qu'une assistance ou une intervention humaine est requise.

Ces «problèmes humains», également appelés «risques humains» (Houghton et Green, 2018), concernent sept aspects de l'entreprise selon un rapport du *Chartered Institute for Personnel Development (CIPD)* (Houghton et Green, 2018):

1. Gestion des talents
2. Santé et sécurité
3. Éthique du personnel
4. Diversité et égalité
5. Relations avec les employés
6. Continuité de l'activité
7. Risque de réputation

Toutefois, peut-être que l'«humain» n'est pas le seul problème. Si l'on se base sur la définition originale de l'intelligence artificielle, selon laquelle des machines pourraient avoir la capacité de se comporter comme des êtres humains, il n'est pas étonnant que les IA fournissent des réponses partiales, étant donné que les humains eux-mêmes ont un comportement partiel et discriminatoire. En d'autres termes, l'apprentissage automatique fonctionne uniquement avec les données qui lui sont fournies, et si ces données révèlent un historique de pratiques discriminatoires en matière d'embauche et de licenciement, les résultats du processus algorithmique reflèteront cette conduite discriminatoire. Si les informations recueillies ne sont pas contrebalancées par des informations qualitatives sur les expériences personnelles et une consultation avec les employés, l'IA risquera de prendre des décisions injustes (voir ci-dessous pour en savoir plus à ce sujet).

Les pratiques RH augmentées IA peuvent aider les directeurs à obtenir des informations censément objectives sur les employés potentiels, avant leur embauche, dans la mesure où la direction a accès aux données qui les concernent, ce qui présente des implications importantes pour la personnalisation de la protection des travailleurs et la prévention des risques SST au niveau individuel. Idéalement, les outils de *people analytics* peuvent aider les employeurs à «mesurer, suivre et comprendre les performances des employés, les différents aspects de la planification des effectifs, la gestion des talents et la gestion opérationnelle» (Collins, Fineman et Tsuchida, 2017). En effet, la prise de décision algorithmique offerte par les *people analytics* peut être utilisée pour offrir un soutien au personnel, en alignant la rémunération des performances et la participation aux bénéfices, ainsi que les coûts de main d'œuvre, sur la stratégie commerciale de l'entreprise et en apportant un appui à des travailleurs spécifiques (Aral et al., 2012, cités dans Houghton et Green, 2018, p. 5). Les employés doivent être personnellement responsabilisés par l'accès à de nouvelles formes de données, qui les aideront à identifier leurs domaines d'amélioration, afin de stimuler le développement individuel et d'obtenir un engagement plus fort.

Cependant, si les processus décisionnels algorithmiques des *people analytics* n'impliquent pas d'intervention humaine ni de considérations éthiques, cet outil RH peut exposer les employés à un stress et à des risques structurels, physiques et psychosociaux accrus. En effet, comment les employés peuvent-ils être certains que les décisions sont prises de façon équitable, exacte et honnête, s'ils n'ont pas accès aux données que leur employeur détient et utilise? Les risques de stress et d'anxiété augmentent lorsque les travailleurs estiment que les décisions sont fondées sur des chiffres et sur des données auxquels ils n'ont pas accès et sur lesquels ils n'ont aucun contrôle. Ce phénomène est particulièrement inquiétant lorsque les données de *people analytics* entraînent des restructurations, des remplacements d'emploi ou encore des changements de description de poste. Les outils de *people analytics* entraîneront une augmentation logique du stress des employés si les données sont utilisées dans la gestion des évaluations et des performances, sans une diligence raisonnable de traitement et de mise en œuvre, qui questionnent sur la possibilité d'un micro-management et donne l'impression aux travailleurs d'être espionnés. S'ils savent que leurs données sont consultées en vue d'une chasse aux talents ou au contraire pour décider d'un potentiel licenciement, les employés se sentiront contraints d'améliorer leurs performances et risquent le surmenage, ce qui présente un risque SST réel. Un autre risque existe, celui de la responsabilité de l'entreprise, car il pourra être demandé aux entreprises de justifier leurs affirmations concernant les capacités prédictives des systèmes d'IA utilisés et les services du personnel devront peut-être faire face à des plaintes pour discrimination.

Un expert en liaison avec les travailleurs<sup>(2)</sup> affirme que la collecte des données relatives aux employés à des fins décisionnelles, telle que celle employée en *people analytics*, constitue la problématique la plus urgente associée avec l'utilisation des IA sur le lieu de travail. Il est en outre fréquent que le comité d'entreprise ne soit pas informé des usages possibles de ces outils de gestion. Ou bien ces systèmes sont mis en place sans que le comité d'entreprise et les employés n'aient été consultés. Ces outils engendrent d'autres risques en matière de SST, notamment du stress pour les travailleurs et des pertes d'emploi, lorsque les technologies sont mises en œuvre dans l'urgence, sans une consultation et une formation ou une communication appropriées. Dans ce contexte, il est intéressant d'étudier un projet mis en place au siège de la compagnie IG Metall, dans le cadre duquel les programmes de formation du site ont été examinés pour l'année 2019 dans le contexte de l'*Industrie 4.0* (voir également section 3.4)<sup>(3)</sup>. Cette étude démontre que les besoins en formation doivent être revus, non seulement pour préparer les travailleurs aux risques physiques qu'ils encourent, comme dans toute formation SST de l'industrie lourde, mais également pour leur permettre de faire face aux risques mentaux et psychosociaux introduits par la transformation numérique du travail, qui implique notamment l'utilisation d'outils de *people analytics*<sup>(4)</sup>.

Un autre outil de *people analytics* consiste à filmer les entretiens d'embauche. Cette pratique a déjà été adoptée par des organisations telles que Nike, Unilever et Atlantic Public Schools. Ces entreprises utilisent des solutions qui permettent aux employeurs de filmer les entretiens avec les candidats, puis d'appliquer des outils IA qui analysent les signaux verbaux et non verbaux émis. L'une de ces solutions est proposée par le groupe HireVue, et est actuellement utilisée par plus de 600 entreprises. L'objectif est de réduire le degré de partialité du recruteur, afin qu'il ne se laisse pas influencer par un niveau d'énergie bas chez le candidat ou par ses propres affinités avec celui-ci, par exemple en termes d'âge, d'origine ethnique ou de démographie. Il est toutefois évident que le système tiendra compte des préférences d'embauche des précédents recruteurs, celles-ci se portant généralement sur un homme hétérosexuel de race blanche, selon un rapport de Business Insider, les autres critères étant sans incidence (Feloni, 2017). Si les données fournies à l'algorithme reflètent le préjugé dominant appliqué par le passé, celui-ci favorisera vraisemblablement les expressions faciales «validées» selon ce modèle dominant, au détriment des signaux liés à une orientation sexuelle, un âge et un sexe autres.

Globalement, ces outils de *people analytics* représentent donc aussi bien des avantages que des risques en matière de SST. Étant donné que ces outils utilisent des algorithmes, les machines doivent être soumises à des tests approfondis avant d'être appliquées dans des situations RH en conditions réelles. Il est également possible de concevoir les algorithmes de *people analytics* spécifiquement pour éliminer préjugés et partialité; ce n'est toutefois pas une tâche aisée. Un processus d'évaluation des risques est déjà à l'expérimentation sur les systèmes pénaux, où une IA est chargée d'informer les commissions d'application des peines et de libération conditionnelle, en vue d'éviter toute partialité. IBM a récemment médiatisé un outil visant de même à réduire les risques de discrimination. Il est à espérer que ce type d'initiative permettra de prendre en charge les risques SST croissants associés aux processus décisionnels RH assistés par IA. Cependant, ce qui fait la force de l'IA fait aussi sa faiblesse.

---

(2) L'auteure a rencontré en entretien le Dr Michael Bretschneider-Hagemes, responsable du Bureau de liaison avec les employés de la Commission allemande KAN, le 18 septembre 2018.

(3) Le terme *Industrie 4.0*, qui fait l'objet de nombreux débats, est apparu dans les cercles industriels allemands, avec pour mission de faire progresser la terminologie marketing de l'industrie manufacturière. Certains critiques affirment qu'il s'agit davantage d'une volonté narrative que d'une réalité aujourd'hui. Il est cependant communément accepté que, si l'on retrace la trajectoire des différentes révolutions industrielles, l'*Industrie 1* correspond à la première révolution industrielle marquée par l'invention du moteur à vapeur. La deuxième révolution correspond aux avancées de la science et la troisième aux inventions numériques intégrées à l'environnement de production. Aujourd'hui, l'Internet des objets, avec lequel les machines communiquent techniquement les unes avec les autres, les outils robotiques avancés et les capacités croissantes en termes de puissance de calcul et de mémoire sont considérés comme les forces motrices sous-jacentes à l'*Industrie 4.0*.

(4) Antje Utecht, qui travaille au service politique et formation du siège de la compagnie IG Metall, à Francfort (Allemagne), a fourni ces informations à l'auteure du présent rapport lors d'un entretien mené le 16 octobre 2018.

## 2.2 Les cobots utilisés en usine et dans les entrepôts

La scène est facile à imaginer: d'énormes bras robotisés orange se déplaçant en vrombissant au sein d'un gigantesque espace industriel, assemblant des pièces de voiture sur les tapis roulants des lignes de production où travaillaient auparavant des êtres humains. Dans de nombreuses usines, les robots ont directement remplacé les ouvriers sur les lignes d'assemblage, et le grand public confond parfois ces systèmes d'automatisation avec une véritable IA. Dans son sens originel, l'automatisation permet de remplacer à l'identique un bras humain par un bras robotisé. Le rapport de l'Agence européenne pour la sécurité et la santé au travail (EU-OSHA) intitulé *Prévisions concernant les risques nouveaux et émergents en matière de sécurité et de santé au travail liés à la numérisation d'ici 2025* (EU-OSHA, 2018, p. 89) explique que les robots permettent de remplacer les travailleurs humains dans les tâches et les environnements dangereux présentant des risques chimiques et ergonomiques, réduisant ainsi les risques SST pour le personnel.

Les emplois manuels et peu qualifiés ont toujours été les plus susceptibles d'être automatisés, et c'est toujours le cas aujourd'hui. Désormais, l'automatisation peut être «augmentée» par des machines au comportement autonome, parfois presque douées de pensée. Avec l'intelligence artificielle, l'automatisation permettrait donc de remplacer non seulement le corps d'un travailleur, mais également son cerveau. Comme l'explique un document de réflexion de l'EU-OSHA sur l'avenir du travail dans le contexte de la robotisation, alors qu'ils ont été initialement créés pour réaliser des tâches simples, les robots sont aujourd'hui de plus en plus sophistiqués et dotés de fonctions d'IA conçues pour leur offrir une «pensée autonome» (Kaivo-oja, 2015).

Des robots collaboratifs (cobots) sont désormais intégrés aux usines et aux entrepôts, où ils travaillent aux côtés des employés humains de façon collaborative. Ils les assistent dans une palette croissante de tâches, sans automatiser nécessairement un poste complet. 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.

Comme l'indique un récent rapport de l'organisation néerlandaise pour la recherche scientifique appliquée (Netherlands Organisation for Applied Scientific Research, TNO), il existe trois types de risque SST dans les environnements avec interactions humain-cobot (TNO, 2018, p. 18-19):

1. Les risques de collision robot-humain, dus à un comportement imprévisible du robot suite à l'apprentissage automatique
2. Les risques de sécurité, lorsque la connexion Internet des robots affecte l'intégrité de la programmation logicielle, entraînant l'apparition de vulnérabilités
3. Les risques environnementaux, en cas de dégradation des capteurs ou de comportement humain imprévu dans des environnements non structurés

Avec la démocratisation des modèles IA, de la reconnaissance vocale et de la vision par ordinateur, les emplois peu qualifiés ne sont aujourd'hui plus les seuls à risquer d'être remplacés. D'autres postes moins routiniers et répétitifs peuvent désormais être pris en charge par des cobots, ou d'autres outils et applications. Avec ces technologies, l'automatisation augmentée IA permet donc de faire prendre en charge de nombreux autres aspects du travail par des ordinateurs et des machines (Frey et Osborne, 2013). Les outils augmentés IA offrent également une protection des ouvriers en matière de SST, comme le montre l'exemple d'une entreprise de l'industrie chimique fabriquant des pièces optiques pour les machines. Les minuscules puces que cette entreprise produit doivent être contrôlées pour identifier d'éventuels défauts. Auparavant, une personne était chargée de détecter à l'œil nu ces défauts en se tenant assise, immobile, devant une suite d'images de puces répétitives, pendant plusieurs heures d'affilée. Aujourd'hui, cette tâche est entièrement prise en charge par une IA, sans intervention humaine. Les risques SST qui ont aujourd'hui pu être éliminés par ces méthodes incluent les troubles musculosquelettiques, ainsi que la fatigue et les lésions oculaires<sup>(5)</sup>.

<sup>(5)</sup> Informations obtenues lors de l'entretien avec Antje Utecht (cf. Note 4).

Les cobots peuvent réduire les risques SST en permettant aux systèmes IA d'effectuer d'autres tâches de service courantes et de routine dans les usines, qui par le passé généraient du stress, un surmenage, des troubles musculosquelettiques, voire même de l'ennui chez les ouvriers, de par leur caractère répétitif. Cependant, les robots augmentés IA utilisés dans les usines et les entrepôts peuvent être source de stress et engendrer des problèmes graves s'ils ne sont pas mis en œuvre de façon appropriée. Ainsi, une syndicaliste britannique affirme que la transformation numérique, l'automatisation et la gestion algorithmique, lorsqu'elles sont «utilisées conjointement, sont toxiques et conçues pour priver des millions de travailleurs de leurs droits fondamentaux»<sup>(6)</sup>. Les problèmes SST potentiels incluent des facteurs de risque psychosociaux s'il est demandé au personnel de travailler au même rythme que les cobots (et non l'inverse) et en cas de collisions entre des cobots et des employés<sup>(7)</sup>. Autre exemple d'interaction homme-machine liée à l'utilisation des cobots et générant de nouvelles conditions de travail et des risques SST, lorsqu'une personne est chargée de surveiller une machine et reçoit pour cela des notifications et des avis d'état sur ses appareils de communication personnels (smartphone ou ordinateur). Ce type de situation entraîne un risque de surmenage car le travailleur se sent obligé de prendre note de ces notifications en dehors de ses heures de travail, et son équilibre vie professionnelle-vie privée s'en trouve perturbé<sup>(8)</sup>.

Un expert<sup>(9)</sup> de l'intelligence artificielle dans le monde du travail a étudié les avancées dans l'utilisation de l'Internet des objets sur le lieu de travail, où des systèmes machine-à-machine connectés travaillent aux côtés des employés humains dans des usines et des entrepôts. Les problèmes de saisie des données, les inexactitudes et les défaillances des systèmes machine-à-machine engendrent des risques SST importants, ainsi que des questions de responsabilité. En effet, les capteurs, les logiciels et la connectivité peuvent être défectueux et instables, et toute vulnérabilité soulève la question de savoir qui est responsable légalement des dommages causés. En cas de collision entre un cobot et un employé, de qui est-ce la faute? Du cobot? De l'employé? De la société qui a fabriqué le cobot? Ou bien de la société qui emploie le travailleur et exploite le cobot? Les questionnements complexes ne manquent pas.

Les interactions humains-robots engendrent à la fois des risques SST et des avantages physiques, cognitifs et sociaux, cependant les cobots disposeront peut-être un jour des compétences nécessaires pour raisonner par eux-mêmes et rassurer les employés humains concernant leur sécurité. Pour ce faire, les cobots doivent bénéficier d'une perception suffisante pour distinguer les objets des êtres humains, et être capable d'anticiper les collisions, d'adapter leur comportement en conséquence et disposer d'une mémoire suffisante pour faciliter l'apprentissage automatique et une prise de décision autonome (TNO 2018, p. 16), dans la lignée des définitions du concept d'intelligence artificielle telles qu'énoncées précédemment.

## 2.3 Les dialogueurs utilisés dans les centres d'appel

Les dialogueurs, également appelés chatbots ou agents conversationnels, sont une autre forme d'outil augmenté IA capable de prendre en charge un pourcentage élevé de requêtes de service client basiques, permettant aux employés humains des centres d'appel d'avoir davantage de temps à consacrer aux questions plus complexes. Les dialogueurs travaillent aux côtés des employés, mais pas au sens physique du terme. Hébergés dans la partie dorsale (back end) des systèmes, ils servent à gérer les requêtes téléphoniques des clients au moyen d'un système de traitement automatique du langage naturel (TALN). La société Dixons Carphone utilise ainsi un agent conversationnel appelé Cami capable de répondre aux questions client de premier niveau sur son site Web Currys et via Facebook Messenger. La compagnie d'assurance Nuance a quant à elle lancé en 2017 un dialogueur baptisé Nina pour répondre aux questions de ses clients et accéder à la documentation. Morgan Stanley a mis à la disposition de 16 000 conseillers financiers des algorithmes d'apprentissage machine afin d'automatiser leurs tâches de routine.

<sup>(6)</sup> Entretien avec Maggie Dewhurst du syndicat britannique *Independent Workers Union of Great Britain (IWGB)* en 2017.

<sup>(7)</sup> Selon l'entretien avec le Dr Sam Bradbrook, spécialiste au *United Kingdom's Health and Safety Executive's Foresight Centre*, en septembre 2018.

<sup>(8)</sup> Entretien avec Antje Utecht (cf. Note 4).

<sup>(9)</sup> Entretien avec le Dr Sam Bradbrook (cf. Note 7).

Les employés de centre d'appels subissent déjà des risques SST importants de par la nature répétitive et exigeante de leur travail, soumis à des niveaux élevés de micro-surveillance et à des formes extrêmes de mesure (Woodcock, 2016). Un nombre important d'activités sont déjà enregistrées et mesurées dans les centres d'appels, et cela ne fait qu'augmenter. Les mots utilisés dans les courriers électroniques ou énoncés vocalement peuvent faire l'objet d'une exploration de données afin de déterminer l'humeur des employés, un processus appelé «analyse des sentiments». Les expressions du visage peuvent également être analysées afin de détecter des signes de fatigue ou des changements d'humeur, ces informations pouvant ensuite être utilisées pour prendre des décisions et réduire les risques de surmenage. Toutefois, les dialogueurs, bien que conçus pour être des machines d'assistance, constituent malgré tout un risque psychosocial par la peur de la perte d'emploi et du remplacement qu'ils engendrent. Les employés devraient être formés à comprendre le rôle et la fonction des robots dans leur environnement de travail et à connaître leur contribution en termes de collaboration et d'assistance.

## 2.4 Outils «prêt-à-porter» et IA dans les environnements de production (par lots)

Les dispositifs de suivi individuel prêt-à-porter sont de plus en plus courants dans les environnements de travail. Le marché des systèmes prêt-à-porter dans l'industrie et le secteur médical a connu une croissance importante ces dernières années, passant de 21 millions USD en 2013 à 9,2 milliards USD en 2020 (Nield, 2014). Entre 2014 et 2019, 13 millions de dispositifs de condition physique ont été intégrés dans les environnements de travail. Cette tendance est déjà présente dans les entrepôts et les usines où les systèmes GPS, l'identification par radiofréquence et aujourd'hui les bracelets de retour haptique, tels que ceux brevetés par Amazon en 2018, ont remplacé les blocs-notes et les stylos.

Une nouvelle dimension de l'augmentation IA des processus d'automatisation et d'Industrie 4.0 est à l'œuvre dans le domaine de la production par lots<sup>(10)</sup>. Les ouvriers, dotés de lunettes de réalité virtuelle, de type HoloLenses et Google Glasses, ou de tablettes, sont postés le long de la chaîne de production pour réaliser des tâches ponctuelles. Le modèle de ligne d'assemblage où les ouvriers effectuent une tâche spécifique de manière répétitive pendant plusieurs heures n'a pas totalement disparu, toutefois la méthode par lots est différente. Utilisée dans les stratégies de production agiles, cette méthode implique des commandes de plus petite taille fabriquées suivant des paramètres temporels spécifiques, plutôt qu'une production constante en masse n'offrant pas des clients garantis.

Dans la fabrication par lots, les ouvriers bénéficient d'une formation visuelle sur le terrain dispensée au moyen de lunettes HoloLens ou d'une tablette, qui les invite à effectuer de nouvelles tâches qu'ils découvrent sur l'instant et qu'ils ne réaliseront que sur le temps de la commande spécifique à produire. Bien que ces systèmes d'assistance *semblent* au premier abord offrir une autonomie, une responsabilisation et un développement personnel accrus, cela n'est pas nécessairement le cas (Butollo, Jürgens et Krzywdzinski, 2018).

L'utilisation de systèmes de formation sur le terrain, qu'il s'agisse de dispositifs prêt-à-porter ou autres, signifie que les connaissances ou la formation préalables nécessaires de la part de l'ouvrier sont moindres, car il se forme au cas par cas. Cette situation engendre un risque d'intensification du travail, car les casques de réalité virtuelle et les tablettes tiennent lieu de formateur en chair et en os pour les travailleurs non qualifiés. D'autre part, les ouvriers n'acquièrent plus de compétences à long terme, car il leur est demandé de réaliser des tâches ponctuelles et modulaires, sur des procédés d'assemblage spécifiques permettant de fabriquer des produits personnalisés à diverses échelles. Bien qu'elle soit favorable au rendement de l'entreprise, la méthode de production par lots génère des risques SST importants dans le sens où elle dé-qualifie les travailleurs, car le seul travail du personnel qualifié consiste à concevoir les programmes de formation sur le terrain utilisés par les ouvriers, qui n'ont plus à être eux-mêmes spécialisés.

---

<sup>(10)</sup> Entretien avec le Dr Michael Bretschneider-Hagemes, précédemment cité (cf. Note 2).

D'autres risques SST peuvent également apparaître du fait de l'absence de communication, car les ouvriers sont incapables d'appréhender suffisamment vite la complexité de cette nouvelle technologie, tout particulièrement s'ils ne sont pas non plus formés pour faire face aux dangers qui peuvent en découler. Le problème est crucial dans le secteur des PME et des start-ups, qui utilisent de façon expérimentale les nouvelles technologies, sans toujours se soucier des normes de sécurité à respecter tant qu'aucun accident ne s'est produit, et il est alors trop tard <sup>(11)</sup>. Un entretien avec les personnes impliquées dans le projet *IG Metall Better Work 2020* (Bezirksleitung Nordrhein-Westfalen/NRW Projekt Arbeit 2020) a révélé que les syndicalistes avaient engagé des discussions actives avec les entreprises sur la façon dont celles-ci choisissaient d'introduire les technologies de l'*Industrie 4.0* dans leur environnement de travail (Moore, 2018a). L'introduction de robots et de solutions de surveillance des ouvriers, de Cloud Computing, de communications machine-à-machine et autres a incité les personnes chargées du projet chez IG Metall à poser les questions suivantes aux entreprises:

- Quel sera l'impact de ces évolutions technologiques sur la charge de travail des employés?
- Le travail en sera-t-il plus facile ou au contraire plus dur?
- Le travail en sera-t-il plus ou moins stressant? La charge de travail sera-t-elle plus ou moins importante?

Les syndicalistes de IG Metall affirment que le niveau de stress des ouvriers tend à augmenter lorsque les technologies sont mises en œuvre sans une formation et un dialogue suffisants avec le personnel. Une expertise est souvent nécessaire pour réduire les risques associés à l'introduction de nouvelles technologies dans un environnement de travail.

Parlons à présent d'un domaine où l'intelligence artificielle a un impact certain, à savoir les environnements de travail de la *gig economy*.

## 2.5 Plates-formes de la *gig economy*

Dans le cadre de cette *gig economy*, le «gig work», ou «travail à la tâche», est mis à disposition des acteurs du secteur par le biais d'applications en ligne (applis), également appelées plates-formes, par des entreprises telles que Uber, Upwork ou Amazon Mechanical Turk (AMT). Le travail requis peut être effectué *en ligne* (c'est-à-dire que la mission est obtenue et exécutée sur un ordinateur à domicile, dans un café ou tout autre espace public, par exemple pour des travaux de traduction et de conception) ou bien *hors ligne* (c'est-à-dire que la mission est obtenue et exécutée dans le monde réel, par exemple pour des services de nettoyage ou de taxi). Tous les algorithmes n'utilisent pas l'intelligence artificielle, toutefois les informations produites par les services de mise en lien entre le client et le «prestataire», ainsi que l'évaluation du service de la plate-forme par le client génèrent des données qui permettent de former des profils générant à leur tour des notes globales plus élevées ou plus basses, qui amènent ensuite le client à choisir un prestataire spécifique pour ses missions.

Les coursiers et les chauffeurs de taxi connaissent depuis de nombreuses années une surveillance et un suivi au quotidien; toutefois, la montée en puissance de ces nouvelles plates-formes «hors ligne» permettant la livraison de commandes ou de repas en vélo ou des services de taxi est un phénomène relativement nouveau. Uber et Deliveroo imposent aux prestataires d'installer une application spécifique sur leur téléphone, placé ensuite sur le tableau de bord du véhicule ou le guidon du vélo, qui leur permet d'acquérir des clients par l'utilisation de technologies de cartographie par satellite et de logiciels de mise en lien basés sur des algorithmes. Les bénéfices de l'IA dans cette économie spécifique peuvent être à l'avantage du prestataire comme du client. DiDi, service chinois d'autopartage, utilise un logiciel IA de reconnaissance faciale pour identifier les prestataires lorsqu'ils se connectent à l'application. DiDi utilise cette information pour garantir l'identité du chauffeur, dans un souci de prévention des délits. Cependant, un échec récent de

---

<sup>(11)</sup> Comme indiqué par le Prof. Dr Dietmar Reinert, PDG de PEROSH, Institut allemand pour la sécurité au travail des organismes d'assurance-accident, dans un entretien avec l'auteure mené le 13 septembre 2018.

cette technologie a fait grand bruit avec un accident tragique. Un chauffeur s'est connecté avec le profil de son père, et sous une fausse identité a causé la mort de son passager.

Les prestataires de ce type de plate-forme sont responsables de leur vitesse, du nombre de livraisons par heure et de leur classement client, et ce dans un environnement très concurrentiel reconnu pour engendrer des risques SST. Dans le magazine Harper, un chauffeur témoignait, expliquant que ces nouveaux outils numériques jouaient le rôle de «fouet mental» poussant les travailleurs à la tâche, en les «intimidant pour qu'ils travaillent plus vite» (The Week, 2015). En effet, les chauffeurs et les livreurs risquent de se voir exclus de la plate-forme si leur classement client n'est pas suffisamment bon ou s'ils ne répondent pas à certains critères. Il en résulte des risques SST flagrants, tels qu'un traitement inéquitable, du stress, voire de la peur.

Les algorithmes sont utilisés pour faire le lien entre les clients et les prestataires dans cette *gig economy* (on parle également de micro-travail). La plate-forme BoonTech utilise l'outil Personality Insights intégrant l'IA Watson d'IBM pour mettre en lien les clients et les prestataires, notamment ceux qui acceptent des missions via AMT et Upwork. Des problèmes de discrimination sont apparus en lien avec les responsabilités domestiques des femmes, lorsque celles-ci exercent une activité depuis leur domicile, notamment celle de mère porteuse ou de garde, dans un contexte traditionnel. Une récente enquête réalisée par des chercheurs de l'OIT auprès de ces travailleurs à la tâche dans les économies en développement indique que les femmes sont plus nombreuses que les hommes à «préférer travailler à domicile» (Rani et Furrer, 2017, p. 14). Cette étude de Rani et Furrer révèle qu'en Afrique, 32 % des travailleuses ont des enfants en bas âge, contre 42 % en Amérique latine. La charge est donc double pour ces femmes, qui «passent environ 25,8 heures par semaine à travailler sur ces plates-formes, dont 20 heures de travail payé et 5,8 heures considérées comme du travail non rémunéré» (ibid., p. 13). L'enquête révèle également que 51 % des femmes travaillant dans la *gig economy* travaillent de nuit (22h à 5h) ou en soirée (76 % de 18h à 22h), qui sont considérées comme des «heures de travail antisociales» selon les catégories de l'OIT sur les risques de violence et de harcèlement au travail (ILO, 2016, p. 40). Rani et Furrer ajoutent que l'externalisation mondiale du travail par le biais des plates-formes a engendré le développement d'une «économie fonctionnant 24h/24 et 7j/7 [...] qui estompe les limites entre vie professionnelle et vie privée [...] [avec pour conséquence] une double charge pour les femmes, étant donné que les tâches domestiques ne sont pas réparties de manière équitable entre les sexes» (2017, p. 13). Travailler à domicile peut en soi créer un environnement de travail dangereux pour la femme, car elle peut y être soumise à la violence domestique ainsi qu'à l'absence des protections légales associées à un travail de bureau classique. En effet, «violence et harcèlement peuvent découler [...] de ces technologies qui brouillent les frontières entre lieu de travail, espace privé et espace public» (OIT, 2017, p. 97).

La numérisation des tâches non standard, telles que les missions réalisées en ligne à domicile ou les services de taxi et de livraison proposés hors ligne, constitue une méthode de gouvernance basée sur la quantification des tâches à un niveau granulaire très bas, où seul le temps de contact explicite est rémunéré. La transformation numérique peut sembler formaliser ce marché du travail, au sens entendu par l'OIT, toutefois le risque de sous-emploi et de sous-rémunération est bien réel. En termes de temps de travail, les tâches préparatoires d'amélioration de la réputation et de développement des compétences nécessaires ne sont pas du tout rémunérées dans la *gig economy*. La surveillance est la norme mais le stress est présent. D'Cruz et Noronha (2016) présentent une étude de cas sur les travailleurs en ligne de la *gig economy* en Inde, pays dans lequel il est reproché aux «services humains à la demande» ou «humans-as-a-service» comme les appellent Jeff Bezos (Prassl, 2018) de déshumaniser et de dévaloriser le travail, d'accroître la précarisation des travailleurs et même d'entraîner une informalisation de l'économie. Le travail à la tâche en ligne, comme celui proposé via les plates-formes telles qu'AMT, s'appuie sur des formes d'emploi non standard (ibid., p. 46) qui augmentent le risque de travail des enfants, de travail forcé et de discrimination. Le racisme est présent, avec pour preuve des commentaires client directs insultants et abusifs sur les plates-formes. Des comportements racistes entre prestataires sont également observés: les acteurs de la *gig economy* travaillant dans des pays plus développés reprochent à leur homologues indiens de casser les prix (ibid.). D'autre part, certaines tâches missionnées via ces plates-formes en ligne sont particulièrement désagréables, notamment le travail réalisé par les modérateurs de contenu, qui passent au crible un volume important d'images afin d'éliminer les plus dérangeantes et insultantes, avec des protections et

des aides quasi-inexistantes. Il existe dans ce contexte des risques évidents de violation des principes SST dans un environnement favorisant le stress et la violence psychosociale, la discrimination, le racisme, les brutalités, le travail des enfants et le travail forcé, face à l'absence des protections les plus fondamentales.

Dans la *gig economy*, il est imposé aux travailleurs de s'inscrire en tant que travailleurs indépendants, ce qui a pour effet de leur faire perdre les droits fondamentaux dont bénéficie tout salarié, notamment les heures garanties, les congés maladies, les congés payés et le droit d'adhérer à un syndicat. La réputation est en outre un facteur crucial pour ces «prestataires», car elle garantit l'obtention de clients. Comme mentionné précédemment, les avis et les notes des clients sont essentiels au développement d'une bonne réputation et déterminent quel volume de missions le travailleur obtient. Les algorithmes effectuent leur apprentissage sur la base des avis client et sur la quantité de missions acceptées, produisant des profils de travailleur spécifiques généralement accessibles au public. Les classements client ne tiennent aucun compte de l'état de santé des travailleurs, de leurs responsabilités familiales et domestiques, ni des circonstances hors de leur contrôle qui peuvent affecter leurs performances professionnelles, engendrant des risques SST accrus car les travailleurs se sentent obligés d'accepter davantage de missions qu'ils ne peuvent raisonnablement prendre en charge, de crainte d'être exclus du système. Les classements de satisfaction client et le nombre de missions acceptées peuvent en effet être utilisées pour «désactiver», et donc exclure, les chauffeurs de taxi qui utilisent la plate-forme, comme c'est le cas chez Uber, malgré le paradoxe et l'illusion que les algorithmes sont exempts de toute «partialité humaine» (Frey et Osborne, 2013, p. 18).

Pour résumer, l'intégration de l'IA à ces plates-formes de la *gig economy* peut offrir des avantages en termes de choix de vie et de travail, notamment avec la protection de l'identité du prestataire et la possibilité de travailler selon des horaires flexibles. Cependant, ces mêmes avantages peuvent se transformer en risques, comme l'incident du chauffeur DiDi le prouve et dans le cas de la double charge imposée aux femmes travaillant selon ce modèle. Les protections en matière de SST sont généralement faibles dans ce type d'environnement professionnel et les risques nombreux (Huws, 2015; Degryse, 2016): rémunération faible et longues heures de travail (Berg, 2016), absence endémique de formation (CIPD, 2017) et insécurité élevée (Taylor, 2017). Williams-Jimenez (2016) tire la sonnette d'alarme, affirmant que le droit du travail et les lois en matière de santé et de sécurité au travail (SST) n'ont pas suffisamment évolué pour faire face à cette transformation numérique du travail; et d'autres études arrivent à la même conclusion (Degryse, 2016). Les succès de l'IA sont aussi ses échecs.

Après avoir décrit l'introduction de l'intelligence artificielle dans les environnements de travail, ainsi que ses avantages et ses risques en matière de SST, étudions à présent quelle réponse la communauté SST internationale apporte à cette nouvelle donne, en identifiant l'évolution des politiques, et en étudiant les débats et les discussions actuels sur ces sujets.

### 3 Élaboration des politiques, réglementation et formation

L'émergence de l'intelligence artificielle et en particulier de l'écosystème et des fonctionnalités de prise de décision autonome exige une «réflexion sur la pertinence de certaines règles établies pour les questions de sécurité et de droit civil en matière de responsabilité» (Commission européenne, 2018). Il est ainsi nécessaire de revoir la réglementation horizontale et sectorielle afin d'identifier tous les risques émergents, d'offrir une protection adéquate et de garantir les bénéfices associés à l'intégration de technologies augmentées IA dans les environnements de travail. La Directive machine (2006/42/CE), la Directive RED (équipements radioélectriques) (2014/53/UE), la Directive DSGP (sécurité générale des produits) (2001/95/CE), ainsi que d'autres réglementations spécifiques relatives à la sécurité, offrent des principes directeurs en la matière. D'autres seront cependant nécessaires pour garantir la sécurité et la santé des personnels sur le lieu de travail. Un rapport de *IOSH Magazine* souligne d'ailleurs que les risques associés à l'IA «devancent les protections en place» (Wustemann, 2017) en matière de sécurité au travail.

À la lumière de cet état des lieux, le présent chapitre étudiera les perspectives des décideurs politiques et des experts face aux recommandations émergentes et de la communauté pour la réglementation de l'intelligence artificielle, en vue de réduire les risques SST, puis proposera des suggestions de formation concernant l'IA et la SST chez IG Metall.

### 3.1 Commission européenne

Le marché unique numérique est un instrument important pour le développement de l'intelligence artificielle. L'évaluation à mi-parcours de la Commission sur la mise en œuvre de ce marché unique numérique (Commission européenne, 2017) indiquait que l'IA serait à même d'apporter une solution technologique à des situations dangereuses, notamment en diminuant le nombre de morts sur les routes, en permettant une utilisation plus intelligente des ressources, en réduisant l'usage de pesticides, en améliorant la compétitivité du secteur manufacturier, en accroissant la précision des opérations chirurgicales et en offrant une assistance dans les situations de danger de type tremblement de terre ou catastrophe nucléaire. Les débats sur le sujet à l'échelle européenne soulèvent de nombreuses questions: aspects juridiques et questions de responsabilité, partage et stockage des données, risque de partialité dans les compétences d'apprentissage automatique et difficulté d'offrir un droit à une explication, notamment sur la façon dont les données relatives aux employés sont utilisées, surtout à la lumière de la réglementation renforcée par le RGPD (Règlement général sur la protection des données).

Ainsi, suite à l'évaluation de mi-parcours du marché unique numérique, qui a des répercussions sur l'IA, la SST et le monde du travail, la discussion actuelle porte sur les risques de partialité et le droit à une explication sur l'usage des données, pour lequel un consentement éclairé concernant l'utilisation des données et le droit d'accès aux données relatives à un individu est crucial. Les problématiques socio-économiques et éthiques en matière d'intelligence artificielle ont révélé leur ampleur dans les communications les plus récentes de la Commission européenne, tout particulièrement depuis la Communication d'avril 2018 sur l'intelligence artificielle en Europe et les conclusions sur le plan coordonné de développement et d'usage de l'intelligence artificielle en Europe, qui soulignent l'importance de l'éthique dans la mise en œuvre d'un avantage concurrentiel.

### 3.2 Normes internationales

Un comité de l'Organisation internationale de normalisation (ISO) a travaillé en 2018 et 2019 à la création d'une norme appliquée à l'utilisation de tableaux de bord et de métriques sur le lieu de travail. Cette norme intégrera une réglementation sur la façon dont les tableaux de bord peuvent être établis et sur la collecte et l'utilisation des données relatives aux employés. Les outils de quantification intéressent de plus en plus les employeurs; toutefois, ces données ne sont d'aucune utilité si elles ne sont pas standardisables. Bien que les représentants des éditeurs du logiciel utilisé pour standardiser les données (SAP) participent activement aux discussions avec l'ISO, il est important que d'autres acteurs soient également impliqués (en Allemagne, par exemple, IG Metall s'intéresse à la formation et à l'IA) afin de garantir une vraie cogestion et une représentation plus importante des travailleurs à l'échelle internationale. Un expert du secteur explique que les normes internationales peuvent être un moyen efficace pour garantir la concrétisation des bénéfices de ces outils, et qu'une étape importante consisterait à s'assurer que les pratiques des entreprises à l'international soient harmonisées, que les données soient standardisables et que les travailleurs soient impliqués dans les discussions et les processus de mise en œuvre<sup>(12)</sup>. D'autre part, une évaluation des risques doit être réalisée sur la base des gros volumes de données recueillis par ces outils de tableau de bord, qui peuvent être une vraie source d'avantages en matière de protection de la SST.

<sup>(12)</sup> Entretien avec Rolf Jaeger, European Industrial Relations Intercultural Communication and Negotiation, 18 septembre 2018.

### 3.3 Organisation internationale du travail

L'OIT a publié pour ses états-membres de nombreux rapports suggérant de meilleures pratiques pour l'intégration de l'intelligence artificielle aux environnements de travail.

Dans son rapport intitulé *Les plateformes de travail numérique et l'avenir du travail* (Berg et al., 2018), elle révèle l'aspect humain de l'intelligence artificielle: nombre des travaux ou des micro-tâches missionnées via les plates-formes en ligne (présentées précédemment) sont similaires au travail non qualifié qui peut, dans la plupart des cas, être automatisé. Le rapport indique que les plates-formes de micro-travail ont été inventées en partie pour faire face aux défaillances des algorithmes du Web 2.0, qui ne parvenaient pas à «classifier les nuances des images, des sons et des textes» que les entreprises souhaitaient conserver et classifier (Irani, 2015, p. 225, cité dans Berg et al., 2018, p. 7). Le travail à effectuer couvre aussi bien des tâches de «gros œuvre», par exemple la réalisation d'une enquête auprès de milliers de personnes, que la reconnaissance d'images. Amazon a d'ailleurs qualifié le type de travail effectué par les «prestataires» de sa plateforme AMT d'«intelligence artificielle artificielle» ou de «main d'œuvre humaine adaptable et à la demande réalisant des tâches que les êtres humains exécutent mieux que les ordinateurs, par exemple reconnaître des objets sur des photos» (Berg et al., 2018, p. 7).

Ce rapport recommande la réglementation des plates-formes de «crowdwork» (qui, selon lui, remplacent le travail automatisé et des IA, comme indiqué précédemment), en dressant une liste de 18 critères pour un «microtravail plus juste» qui incluent des recommandations telles que l'élimination du statut obligatoire de travailleur indépendant du fait que le travailleur est un employé salarié dans la pratique; le droit à adhérer à un syndicat et à une négociation collective; un salaire minimum; une transparence des frais (le vol salarial est un problème courant dans la *gig economy*); le droit pour les travailleurs d'accepter certaines tâches et d'en refuser d'autres, sans en être pénalisés; une protection contre les pannes informatiques; des conditions générales d'utilisation des plates-formes concises et lisibles; une protection contre l'utilisation abusive des notes et des évaluations des travailleurs; la création d'un code de déontologie; le droit pour les travailleurs de contester un non-paiement et d'autres problèmes; l'octroi d'un accès aux informations relatives aux clients, pour les travailleurs; la révision par les plates-formes des instructions relatives aux tâches à effectuer, avant la publication de l'offre; l'autorisation pour les travailleurs d'exporter leur historique de réputation; le droit pour les travailleurs de travailler en direct avec un client après l'avoir approché via la plate-forme; l'obligation pour les clients et les opérateurs de répondre aux demandes des travailleurs de façon rapide et polie; le droit des travailleurs à connaître la finalité de leur travail; et l'obligation de signaler clairement toute mission impliquant un travail psychologiquement stressant (Berg et al., 2018, p. 105-109).

Le rapport «Travailler pour bâtir un avenir meilleur» de la Commission mondiale sur l'avenir du travail indique que toute action impliquant les technologies et le monde du travail doit faire l'objet d'une réflexion centrée sur l'humain. Selon ce rapport, les cobots peuvent réellement diminuer pour les travailleurs le stress et les risques de blessure. Toutefois, cette technologie peut également réduire le nombre d'emplois disponibles pour les travailleurs humains, ce qui au final aliène les travailleurs et freine leur évolution. Les décisions qui se prennent sur le lieu de travail ne devraient jamais être fondées sur les données produites par des algorithmes, et l'intégration d'une IA devrait toujours se faire sous contrôle humain, dans un environnement où «le contrôle, la surveillance et la gestion algorithmiques, réalisées par le biais de capteurs, de dispositifs prêt-à-porter et d'autres formes de supervision, sont réglementés en vue de protéger la dignité des travailleurs» (OIT, 2019, p. 43). Le rapport se poursuit en reprenant la célèbre citation de la Déclaration de Philadelphie, selon laquelle le travail n'est pas une marchandise: «Le travail n'est pas une marchandise; ce n'est pas un robot non plus.» (ibid.).

### 3.4 Forum économique mondial et RGPD

Le *Global Future Council on Human Rights and Technology* (Conseil mondial sur l'avenir des droits de l'homme et des technologies) du Forum économique mondial (FEM) a signalé en 2018 que, même lorsque des ensembles de données valides et de qualité sont utilisés pour effectuer

l'apprentissage automatique des algorithmes, il existe des risques considérables de discrimination dans les cas de figure suivants (FEM, 2018):

1. Lorsque le modèle choisi n'est pas adapté
2. Lorsque le modèle est fondé, involontairement, sur des critères discriminatoires
3. En l'absence d'intervention et de supervision humaines
4. Lorsque les systèmes sont imprévisibles et impénétrables
5. Lorsque la discrimination n'est pas contrôlée et involontaire

Le FEM souligne en outre qu'il existe un besoin distinct pour «une auto-gouvernance plus active de la part des entreprises privées», en conformité avec la Déclaration de principes tripartite sur les entreprises multinationales et la politique sociale de l'OIT (5<sup>e</sup> édition, 2017), qui propose aux entreprises des conseils directs en matière de pratiques de travail durables, responsables et inclusives et sur la politique sociale connexe, dans le cadre de l'Objectif de développement durable (ODD) n°8.8 visant garantir la sécurité sur le lieu de travail et assurer la protection de tous les travailleurs d'ici 2030. La prévention d'une discrimination injuste et illégale est clairement à l'ordre du jour face à l'utilisation croissante des systèmes d'IA, sous l'égide des rapports du FEM (2018) et de l'OIT dont la gouvernance est vitale.

La première erreur qu'une entreprise peut faire lorsqu'elle décide d'utiliser l'intelligence artificielle, et qui peut mener à un contexte de discrimination telle que définie par le FEM, est d'arriver à une situation dans laquelle l'utilisateur applique le même algorithme à deux problèmes différents, qui ne présentent pas des contextes et des points de données identiques. Prenons par exemple une entreprise dans laquelle les candidats à l'embauche sont évalués à l'aide d'un algorithme chargé de rechercher des signes sur le type de personnalité par le biais des réseaux sociaux et de vidéos permettant de détecter les expressions du visage, ces données étant ensuite regroupées avec les ensembles de données de curriculum vitæ, en revenant parfois plusieurs années avant l'embauche. Comme le souligne le Dr Cathy O'Neil dans un entretien avec l'auteure<sup>(13)</sup>, l'algorithme doit être conçu pour être discriminatoire, ou au minimum sélectif, puisque les pratiques d'embauche reposent sur la sélectivité. Cependant, si l'algorithme recherche par exemple des personnes au caractère extraverti pour un emploi en centre d'appels, le même algorithme ne serait pas pertinent pour l'embauche d'un assistant de laboratoire, où l'aisance orale ne fait pas partie de la description de poste. Bien que l'application de l'algorithme ne puisse en soi entraîner systématiquement une discrimination illégale, il n'est pas difficile d'imaginer les possibilités en cas d'utilisation inadéquate.

La deuxième erreur, qui consiste à «créer un modèle présentant des caractéristiques discriminatoires involontaires», peut concerner l'utilisation d'une banque de données intégrant déjà des principes discriminatoires. Par exemple, au Royaume-Uni, les inégalités de salaires entre hommes et femmes ont été récemment exposées, révélant que pendant des années les femmes ont travaillé pour des salaires inférieurs à ceux des hommes, exécutant parfois le même travail qu'un homme pour une rémunération moindre. Si les données qui démontrent cette tendance étaient utilisées pour créer un algorithme assistant les décisions d'embauche, la machine «apprendrait» que les femmes doivent recevoir un salaire inférieur à celui des hommes. Cet exemple démontre que les machines sont incapables aujourd'hui de prendre des décisions éthiques indépendamment de toute intervention humaine. Un nombre croissant d'études démontre d'ailleurs que l'IA ne permet pas d'éliminer la discrimination dans les processus de décision et de prédiction, mais qu'au contraire la codification des données perpétue le problème (Noble, 2018).

La troisième erreur souligne l'importance d'une intervention humaine, qui est aujourd'hui obligatoire sur le territoire européen. En mai 2018, le RGPD est entré en vigueur, exigeant le consentement des travailleurs pour la collecte et l'utilisation des données les concernant. Bien que le RGPD concerne principalement les droits relatifs aux données des consommateurs, ses implications dans le monde du travail sont importantes, car les décisions qui s'y prennent ne peuvent se faire uniquement à l'aide de processus automatisés.

<sup>(13)</sup> Le Dr Cathy O'Neil, auteure de *Weapons of maths destruction* et PDG d'ORCAA (O'Neil Risk Consulting and Algorithmic Auditing), s'est entretenue avec l'auteure du présent rapport le 14 octobre 2018.

Dans la Section 4 du RGPD, intitulée «Droit d'opposition et prise de décision individuelle automatisée», l'Article 22, «Décision individuelle automatisée, y compris le profilage», indique que:

22(1): La personne concernée a **le droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé**, y compris le profilage, produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire.

Les fondements de ce règlement, tels qu'édictés dans les premières sections du document, stipulent que:

(71): La personne concernée devrait avoir **le droit de ne pas faire l'objet d'une décision**, qui peut comprendre une mesure, impliquant l'évaluation de certains aspects personnels la concernant, qui est prise sur le seul fondement d'un traitement automatisé et qui produit des effets juridiques la concernant ou qui, de façon similaire, l'affecte de manière significative, tels que [...] **des pratiques de recrutement en ligne sans aucune intervention humaine**. Ce type de traitement inclut le «profilage» qui consiste en toute forme de traitement automatisé de données à caractère personnel visant à évaluer les aspects personnels relatifs à une personne physique, notamment pour analyser ou prédire des aspects concernant le rendement au travail de la personne concernée, [...] sa fiabilité ou son comportement, ou sa localisation et ses déplacements, dès lors qu'il produit des effets juridiques concernant la personne en question ou qu'il l'affecte de façon similaire de manière significative.

Le non-respect de ces critères peut mener à des décisions discriminatoires inéquitables ou illégales.

Concernant la quatrième erreur, «des systèmes imprévisibles et impénétrables», la description figurant dans le rapport du FEM (2018) indique que «lorsqu'un être humain prend une décision, par exemple celle d'embaucher ou non un candidat, il est possible de lui demander d'explicitier sa décision». À l'inverse, il n'est pas possible de demander à une machine les motivations d'une décision prise sur la base de l'exploration de données. L'élimination de jugements qualifiés et l'absence d'intervention humaine ouvrent donc clairement la voie à la discrimination.

La dernière erreur en matière de mise en œuvre des IA peut se produire lorsque «la discrimination n'est pas contrôlée et involontaire». Cela peut par exemple se produire lorsqu'une entreprise ne souhaite pas embaucher une femme en raison d'une grossesse potentielle. Même si ce positionnement explicite ne tiendrait pas devant un tribunal, un système d'apprentissage automatique peut offrir à l'entreprise un moyen dissimulé de rendre cette discrimination possible par le biais d'un algorithme conçu pour filtrer et rejeter un sous-ensemble de candidates féminines sur la base de l'âge et du statut marital. Il n'est pas difficile de voir que ces pratiques ouvrent la porte non seulement à un risque de discrimination, mais vraisemblablement à la probabilité forte d'une discrimination techniquement illégale.

### 3.4 Formation aux IA et à la SST

IG Metall travaille en collaboration avec des entreprises pour l'établissement de ses programmes de formation en santé et sécurité au travail (SST), afin d'intégrer les dernières évolutions technologiques présentes sur les lieux de travail en 2019. Des discussions avec l'expert responsable de cette initiative révèlent que la formation SST ne concerne généralement qu'un ou deux agents de santé et de sécurité sur site, et n'est pas encore pleinement intégrée à l'ensemble des systèmes. Les résultats d'étude démontrent aujourd'hui qu'il est nécessaire de former tous les employés afin de leur faire acquérir des capacités d'apprentissage rapide, car les technologies évoluent très rapidement et leurs compétences doivent s'adapter en conséquence<sup>(14)</sup>. L'expert explique que la formation doit être adaptée pour conserver sa pertinence dans le contexte de

<sup>(14)</sup> Entretien avec le Dr Maike Pricelius, Secrétaire de projet, Better Work 2020, IG Metall, 12 octobre 2018.

l'Industrie 4.0 et de la transformation numérique, afin que les travailleurs soient préparés à faire face à ces risques émergents. Cependant, ce n'est pas une panacée et le processus doit faire partie d'un plan de mise en œuvre plus large. Si aucun plan d'action n'est établi pour réellement mettre en œuvre et utiliser les nouvelles connaissances et compétences proposées dans le cadre de la formation, celles-ci seront perdues. Dans un tel contexte, il est clair qu'un meilleur alignement entre la formation SST et les technologies intégrées est nécessaire. Cela étant dit, il convient également d'adapter la pédagogie de formation, car l'apprentissage est un processus continu qui doit se poursuivre tout au long de la vie active des travailleurs, tout particulièrement dans le marché du travail actuel marqué par l'instabilité et l'incertitude. Il sera également important pour les travailleurs d'acquérir les principes de base et des compétences en résolution de problèmes, en sus des compétences plus traditionnelles. Aujourd'hui, les travailleurs doivent comprendre et choisir leur propre parcours et style d'apprentissage<sup>(15)</sup>. Seul le temps confirmera le degré réel de présence de l'IA dans les environnement de travail; il est toutefois pertinent de rester en alerte sur les risques et bénéfices en matière de SST, et d'impliquer les travailleurs dans ces processus en leur dispensant une formation adéquate à chaque étape.

## 4 Conclusion

Dans les années 1920 déjà, l'écrivain E. M. Forster dépeignait un monde dystopique mêlant technologie et humanité. Dans sa nouvelle d'anticipation, *La machine s'arrête*, Forster décrit un monde où les humains doivent vivre sous la surface de la Terre, au sein d'une Machine omnipotente vénérée par le protagoniste, car celle-ci (Forster, 1928):

«[...] nous apporte nourriture, vêtements et abri; par elle nous nous parlons les uns aux autres, par elle nous nous voyons les uns les autres, en elle nous sommes ce que nous sommes». La Machine est l'amie des idées et l'ennemie de la superstition: Omnipotente, éternelle, bénie soit la Machine!

Toutefois, cette entité omnipotente et omniprésente commence bientôt à se décomposer et l'expertise humaine ne suffit plus à assurer sa maintenance, entraînant une fin sinistre pour l'ensemble de l'humanité.

Bien que cette œuvre soit un classique en matière de science-fiction, l'invisibilité et la puissance potentielle apparentes des technologies d'aujourd'hui semblent se perpétuer sans fin, car leur fonctionnement est souvent dissimulé dans ce que l'on appelle une boîte noire, où leurs rouages paraissent souvent hors de notre compréhension, bien qu'acceptés par la majorité. En dehors des ingénieurs, peu de personnes comprennent comment fonctionnent les ordinateurs et les systèmes d'IA. Cependant, les experts humains eux-mêmes sont surpris par les actions des IA, par exemple les joueurs d'échecs et de Go battus par un programme informatique.

En Chine, le gouvernement a commencé à attribuer aux citoyens un score, une note de réputation économique et personnelle, qui tient compte du paiement des loyers, de la note de crédit, de l'utilisation du téléphone, etc. Ce score vise à déterminer les conditions d'obtention d'un prêt, d'un emploi ou encore d'un visa de voyage. Les outils de *people analytics* pourraient être utilisés pour attribuer, sur ce même principe, un «score professionnel» aux travailleurs, qui serait utilisé dans les décisions d'évaluation. Cela soulèverait toutefois de nombreuses questions concernant la vie privée et la surveillance des individus. Le terme «condition algorithmique» récemment apparu dans un rapport de l'UE (Colman et al., 2018) se réfère à la logique de plus en plus normalisée des algorithmes, dans laquelle les symboles se transforment en réalité. Aujourd'hui, cette condition commence à affecter le monde du travail, par exemple lorsque la note de réputation en ligne des travailleurs est soumise à un processus algorithmique de mise en correspondance et que les profils individuels sont traités par des robots d'exploration de données. Le problème est que ces algorithmes ne tiennent pas compte des aspects qualitatifs de la vie ni du contexte. Le Dr O'Neil (cf Note 13) a fait une remarque pertinente à ce sujet lors de son récent entretien avec l'auteure.

<sup>(15)</sup> Duncan Spencer, Responsable conseil et pratiques, Institution of Occupational Safety and Health (Leicester, Royaume-Uni), a discuté de ce sujet avec l'auteure lors d'un entretien mené le 15 octobre 2018.

En regardant les livreurs de Deliveroo passer à côté d'elle sous la pluie, le Dr O'Neil observe que les plates-formes par lesquelles ces livreurs obtiennent leurs missions valorisent avant tout la rapidité et l'efficacité, incitant les cyclistes à aller vite malgré des conditions météo défavorables, mettant parfois leur vie même en jeu. Le Dr O'Neil appelle ces algorithmes «des ersatz d'univers», car ces entités qui semblent tout savoir ne connaissent en réalité du monde que ce qu'on leur en dit, et sont donc totalement aveugles à certains aspects.

Sergey Brin, co-fondateur de Google, s'est adressé à ses investisseurs dans sa lettre annuelle 2018 comme suit:

«[...] le nouveau printemps que connaît l'intelligence artificielle est l'évolution informatique la plus remarquable que je connaîtrai durant ma vie [...], cependant, des outils aussi puissants apportent avec eux de nouvelles questions et responsabilités. En quoi affecteront-ils l'emploi dans les différents secteurs d'activité? Comment comprendre leur fonctionnement réel, ce qui se passe «sous le capot»? Comment en mesurer l'équité? Comment peuvent-ils être manipulés? Sont-ils sûrs?

Ces questions éthiques autour de l'intelligence artificielle doivent cependant être discutées au-delà de la sphère de l'entreprise. Nous avons couvert dans le présent rapport les questions relatives à la santé et la sécurité au travail (SST), ainsi que les risques et les bénéfices associés à l'utilisation de ces outils. La machine omnipotente inventée par E. M. Forster dans sa nouvelle d'anticipation classique n'avait bien entendu pas fait l'objet d'une évaluation morale et éthique par des commissions d'examen avant que l'ensemble de l'humanité ne s'abrite en son sein pour vivre sous terre. Cette dystopie n'est pas vraiment représentative de ce que nous vivons aujourd'hui, bien sûr, toutefois les discussions sur le sujet de l'intelligence artificielle, qu'il s'agisse de celles entourant les communications de la Commission européenne et le Plan coordonné européen sur l'intelligence artificielle ou des groupes de révision des programmes des syndicats comme IG Metall, démontrent un intérêt marqué pour la prévention des risques les plus graves et la mise en œuvre des bénéfices de l'IA en termes de SST, alors que les systèmes d'IA sont de plus en plus intégrés aux processus décisionnels des entreprises et au travail assisté.

En conclusion, alors que la mise en œuvre des systèmes d'intelligence artificielle dans le monde du travail est un processus relativement nouveau, nous ne faisons que commencer à découvrir les risques et les bénéfices qui en découlent en matière de santé et de sécurité au travail. Le présent rapport a mis en lumière certains domaines où les bénéfices de l'IA ont été observés et pris en charge, et ses risques identifiés, ainsi que les mesures de précaution et les réglementations appliquées. Dans la prise de décision RH, l'utilisation d'outils de *people analytics* augmentés IA engendrent un risque de traitement inéquitable et de discrimination. Dans le domaine de l'automatisation et de l'Industrie 4.0, les risques incluent une formation inadaptée ou inexistante entraînant surmenage et stress (Downey, 2018), ainsi que des accidents imprévus tels que les collisions en travailleurs humains et robots. La déqualification du travail est un enjeu crucial dans le secteur de la fabrication et les autres industries, face à l'intégration de processus de production par lots et à l'utilisation de technologies prêt-à-porter dans les pratiques de formation automatisée. Des risques pour la vie privée des travailleurs en lien avec une surveillance accrue et un phénomène de micro-management ont été observés, car les dirigeants ont la possibilité d'accéder à des données intimes sur les travailleurs par le biais des technologies prêt-à-porter utilisées dans les usines comme dans les bureaux. Dans la *gig economy*, les algorithmes ne peuvent être les seuls décideurs. Les avantages des technologies d'IA doivent toujours être mis en avant.

Il est en effet important pour l'ensemble des parties prenantes de se concentrer sur les possibilités d'assistance de ces technologies dans les applications commerciales et de veiller à ce que les gouvernements et autres instances réglementaires supervisent l'utilisation des outils et des applications IA dans le monde du travail. Lorsqu'elle est mise en œuvre dans le cadre de processus appropriés, l'intelligence artificielle a un impact positif car elle aide les entreprises à réduire la partialité des décideurs humains, notamment durant les entretiens RH, dans la mesure où les algorithmes ont été conçus pour identifier les critères discriminatoires passés et où les décisions sont prises sous une supervision humaine totale, voire une action affirmative. L'intelligence artificielle peut aider à améliorer les relations avec et entre les employés lorsque les données recueillies identifient un potentiel de collaboration. Les outils RH augmentés IA peuvent améliorer le processus décisionnel grâce à la méthode de prédiction par exception, libérant du

temps pour le développement personnel et de carrière, lorsque l'IA est capable de prendre en charge des tâches répétitives et peu gratifiantes.

Pour éviter les risques SST, l'auteure recommande de mettre l'accent sur les outils IA d'assistance et de collaboration plutôt que de rechercher les compétences génériques d'une intelligence artificielle universelle. Une formation adéquate doit être mise en place à tous les niveaux, de même que des contrôles réguliers, y compris de la part des services SST et des autorités. Il est également important que les travailleurs soient consultés à chaque étape lorsque de nouvelles technologies sont intégrées à l'environnement de travail, permettant une approche centrée sur le travailleur et la mise en place d'un contrôle humain (De Stefano, 2018). Il est de la responsabilité des dirigeants d'entreprise et des gouvernements de se tenir informés des normes internationales, des réglementations gouvernementales et des activités des syndicats, qui connaissent aujourd'hui une évolution significative pour réduire les risques les plus graves associés à l'intelligence artificielle, tout en concrétisant les bénéfices et les effets positifs de celle-ci. En conclusion, ce n'est pas la technologie d'intelligence artificielle en soi qui constitue un risque pour la santé et la sécurité des travailleurs, mais bien la manière dont elle est mise en œuvre. Il est de notre responsabilité à tous de garantir une transition en douceur pour l'intégration de l'IA dans les environnements de travail.

*Auteure: Dr Phoebe V. Moore, Professeure associée en économie politique et technologie, Division management et organisation, School of Business, Université de Leicester, Royaume-Uni – Chargée de recherche, WZB Weizenbaum Institute for the Networked Society 2018-19*

*Gestion du projet: Annick Starren – Agence européenne pour la sécurité et la santé au travail (EU-OSHA)*

## Références

- Agarwal, A., Gans, J., Goldfarb, A., 2018, *Prediction machines: The simple economics of artificial intelligence*, Boston, MA: Harvard Business Review Press.
- Berg, J., 2016, *Income security in the on-demand economy: Findings and policy lessons from a survey of crowdworkers*, Conditions of Work and Employment Series No 74, Genève: Organisation internationale du travail
- Berg, J., Furrer, M., Harmon, E., Rani, U., Silberman, M. S., 2018, *Digital labour platforms and the future of work: Towards decent work in the online world*, Genève: Organisation internationale du travail
- Butollo, F., Jürgens, U., Krzywdzinski, M., 2018, 'From lean production to Industrie 4.0: More autonomy for employees?', Wissenshanftszentrum Berlin für Socialforschung (WZB), Document de réflexion SP 111 2018-303.
- CIPD (Chartered Institute for Personnel Development), 2017, *To gig or not to gig? Stories from the modern economy*. Disponible en ligne: [www.cipd.co.uk/knowledge/work/trends/gig-economy-report](http://www.cipd.co.uk/knowledge/work/trends/gig-economy-report)
- Collins, L., Fineman, D. R., Tshuchica, A., 2017, 'People analytics: Recalculating the route', Deloitte Insights. Disponible en ligne: <https://www2.deloitte.com/insights/us/en/focus/human-capital-trends/2017/people-analytics-in-hr.html>
- Colman, F., Bülmann, V., O'Donnell, A., van der Tuin, I., 2018, *Ethics of coding: A report on the algorithmic condition*, Bruxelles Commission européenne.
- D'Cruz, P., Noronha, E., 2016, 'Positives outweighing negatives: The experiences of Indian crowdsourced workers', *Work Organisation, Labour and Globalisation* 10(1), 44-63.
- De Stefano, V., 2018, 'Negotiating the algorithm: Automation, artificial intelligence and labour protection', ILO working Paper No 246/2018, Genève Organisation internationale du travail
- Degryse, C., 2016, *Digitalisation of the economy and its impact on labour markets*, Bruxelles: Institut syndical européen (ETUI).
- Delponte, L., 2018, *European artificial intelligence leadership, the path for an integrated vision*, Bruxelles: Département thématique des politiques économiques, scientifiques et de la qualité de la vie, Parlement européen.
- Downey, K., 2018, 'Automation could increase workplace stress, unions warn', *IOSH Magazine*, 23 avril 2018. Disponible en ligne: <https://www.ioshmagazine.com/article/automation-could-increase-workplace-stress-unions-warn>
- Dreyfus, H. L., 1972, *What computers can't do*, New York: Harper et Row (réimpression par MIT Press 1979, 1992).
- Ernst, E., Merola, R., Samaan, D., 2018, *The economics of artificial intelligence: Implications for the future of work*, séries de documents de recherches sur l'avenir du travail, OIT, Genève: Organisation internationale du travail Disponible en ligne: [https://www.ilo.org/global/topics/future-of-work/publications/research-papers/WCMS\\_647306/lang--en/index.htm](https://www.ilo.org/global/topics/future-of-work/publications/research-papers/WCMS_647306/lang--en/index.htm)
- EU-OSHA (Agence européenne pour la sécurité et la santé au travail), 2018, *Prospective sur les risques nouveaux et émergents en matière de sécurité et de santé au travail liés à la numérisation d'ici à 2025*, Luxembourg: Office des publications de l'Union européenne. Disponible en ligne: <https://osha.europa.eu/en/tools-and-publications/publications/foresight-new-and-emerging-occupational-safety-and-health-risks/view>.
- Commission européenne, 2018, Communication sur l'intelligence artificielle en Europe, Bruxelles: Commission européenne.

- Commission européenne, 2017, *Communication on the mid-term review on the implementation of the digital single market strategy: A connected digital single market for all*. Disponible en ligne: <https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1496330315823&uri=CELEX:52017DC0228>
- Feloni, R., 2017, «J'ai testé le logiciel augmenté IA utilisé par des entreprises comme Goldman Sachs et Unilever pour analyser les candidats à l'embauche, avant de les rencontrer, et cela me semble moins inquiétant qu'il n'y paraît», Business Insider UK, 23 août 2017. Disponible en ligne: <https://www.uk.businessinsider.com/hirevue-ai-powered-job-interview-platform-2017-8?r=US&IR=T/#in-recorded-videos-hirevue-employees-asked-questions-like-how-would-you-describe-your-role-in-the-last-team-you-worked-in-4>
- Forster, E. M., 1928/2011, *La machine s'arrête*, Londres: Penguin Books.
- Frey, C., Osborne, M. A., 2013, *The future of employment: How susceptible are jobs to computerisation?*, Oxford: Université d'Oxford, Oxford Martin School. Disponible en ligne: [https://www.oxfordmartin.ox.ac.uk/downloads/academic/The\\_Future\\_of\\_Employment.pdf](https://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf)
- Houghton, E., Green, M., 2018. *People analytics: Driving business performance with people data*, Chartered Institute for Personnel Development (CIPD). Disponible en ligne: <https://www.cipd.co.uk/knowledge/strategy/analytics/people-data-driving-performance>
- Hutter, M., 2012, 'One decade of universal artificial intelligence', *Theoretical Foundations of Artificial General Intelligence* 4, 67-88.
- Huws, U., 2015, «Réflexion sur l'avenir du travail: Les services de main-d'œuvre en ligne ou l'externalisation ouverte: conséquences pour la santé et la sécurité au travail», Document de réflexion, Bilbao: Agence européenne pour la sécurité et la santé au travail. Disponible en ligne: <https://osha.europa.eu/en/tools-and-publications/publications/future-work-crowdsourcing/view>
- IBM, 2018, 'IBM talent business uses AI to rethink the modern workforce', IBM Newsroom. Disponible en ligne: <https://newsroom.ibm.com/2018-11-28-IBM-Talent-Business-Uses-AI-To-Rethink-The-Modern-Workforce>
- OIT (Organisation internationale du travail), 2019, *Travailler pour bâtir un avenir meilleur: Commission mondiale sur l'avenir du travail*, Genève: Organisation internationale du travail
- OIT (Organisation internationale du travail), 2017, «Mettre fin à la violence et au harcèlement contre les femmes et les hommes dans le monde du travail, Rapport V», Conférence internationale du Travail, 107e session, 2018, Genève. Disponible en ligne: [http://www.ilo.org/ilc/ILCSessions/107/reports/reports-to-the-conference/WCMS\\_553577/lang--en/index.htm](http://www.ilo.org/ilc/ILCSessions/107/reports/reports-to-the-conference/WCMS_553577/lang--en/index.htm)
- OIT (Organisation internationale du travail), 2016, *Rapport final: Réunion d'experts sur la violence contre les femmes et les hommes dans le monde du travail*, MEVWM/2016/7, Genève: OIT. Disponible en ligne: [http://www.ilo.org/gender/Informationresources/Publications/WCMS\\_546303/lang--en/index.htm](http://www.ilo.org/gender/Informationresources/Publications/WCMS_546303/lang--en/index.htm)
- Kaivo-oja, J., 2015, 'A review on the future of work: Robotics', Document de réflexion, Bilbao: Agence européenne pour la sécurité et la santé au travail. Disponible en ligne: <https://osha.europa.eu/en/tools-and-publications/seminars/focal-points-seminar-review-articles-future-work>
- Kar, S., 2018, 'How AI is transforming HR: The future of people analytics', Hyphen, 4 janvier 2018. Disponible en ligne: <https://blog.gethyphen.com/blog/how-ai-is-transforming-hr-the-future-of-people-analytics>
- McCarthy, J., Minsky, M. L., Rochester, N., Shannon, C. E., 1955, 'A proposal for the Dartmouth Summer Research Project on Artificial Intelligence'. Disponible en ligne: <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>

- Moore, P. V., 2018a, *The threat of physical and psychosocial violence and harassment in digitalized work*, Genève: Organisation internationale du travail
- Moore, P. V., 2018b, *The quantified self in precarity: Work, technology and what counts*, Abingdon, Royaume-Uni: Routledge.
- Nield, D., 2014, 'In corporate wellness programs, wearables take a step forward', *Fortune*, 15 avril 2014. Disponible en ligne: <http://fortune.com/2014/04/15/in-corporate-wellness-programs-wearables-take-a-step-forward/>
- Noble, S. A., 2018, *Algorithms of oppression: How search engines reinforce racism*, New York: New York University Press.
- Pasquale, F., 2015, *The black box society: The secret algorithms that control money and information*, Boston (MA): Harvard University Press.
- Prassl, J., 2018, *Humans as a service: The promise and perils of work in the gig economy*, Oxford: Oxford University Press.
- PwC, 2018a, 'Artificial intelligence in HR: A no-brainer'. Disponible en ligne: <https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html>
- PwC, 2018b, 'AI will create as many jobs as it displaces by boosting economic growth'. Disponible en ligne: <https://www.pwc.co.uk/press-room/press-releases/AI-will-create-as-many-jobs-as-it-displaces-by-boosting-economic-growth.html>
- Rani, U., Furrer, M., 2017, 'Work and income security among workers in on-demand digital economy: Issues and challenges in developing economies', Rapport présenté lors d'un atelier de l'Université de Lausanne intitulé 'Digitalization and the Reconfiguration of Labour Governance in the Global Economy', 24-25 novembre 2017 (non publié).
- Simon, H., 1969, *The sciences of the artificial*, Cambridge (MA): MIT Press.
- Taylor, M., 2017, *Good work: The Taylor review of modern working practices*, Londres: Département des Affaires, de l'Énergie et des Stratégies industrielles (BEIS) Disponible en ligne: <https://www.gov.uk/government/publications/good-work-the-taylor-review-of-modernworking-practices>
- The Week, 2015, 'The rise of workplace spying', Disponible en ligne: <http://theweek.com/articles/564263/rise-workplace-spying>
- TNO (Netherlands Organisation for Applied Scientific Research), 2018, *Emergent risks to workplace safety; Working in the same space as a cobot*, Rapport pour le ministère des Affaires sociales et de l'Emploi, La Haye.
- Turing, A. M., 1950, 'Computing machinery and intelligence', *Mind* 49, 433-460.
- Royaume-Uni, département des Affaires, de l'Énergie et des Stratégies industrielles (BEIS), département du Numérique, de la Culture, des Médias et du Sport, 2018, 'AI sector deal policy paper'. Disponible en ligne: <https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal>
- FEM (Forum économique mondial), 2018, *How to prevent discriminatory outcomes in machine learning*, Global Future Council on Human Rights and Technology du Forum économique mondial (FEM 2016-2018, Cologny, Suisse: FEM.
- Weizenbaum, J., 1976, *Computer power and human reason: From judgment to calculation*, San Francisco: W. H. Freeman.
- Bureau de la politique scientifique et technologique de la Maison-Blanche (OSTP), 2018, *Summit on artificial intelligence for American industry*. Disponible en ligne: <https://www.whitehouse.gov/articles/white-house-hosts-summit-artificial-intelligence-american-industry/> Résumé du rapport. En ligne: <https://www.whitehouse.gov/wp.../Summary-Report-of-White-House-AI-Summit.pdf>

- Williams-Jimenez, I., 2016, 'Digitalisation and its impact on psychosocial risks regulation', Rapport présenté lors de la Fifth International Conference on Precarious Work and Vulnerable Workers, Londres, Middlesex University.
- Woodcock, J., 2016, *Working the phones: Control and resistance in call centres*, Londres: Pluto Press.
- Wustemann, L., 2017, 'AI and nanotech risk outpacing our safeguards', *IOSH Magazine*, 25 août 2017. Disponible en ligne: <https://www.ioshmagazine.com/article/ai-and-nanotech-risk-outpacing-our-safeguards>