



# L'éthique dans l'emploi à l'ère de l'intelligence artificielle

Contribution de Renaissance Numérique & du groupe Randstad en France au débat public lancé par la CNIL sur les enjeux éthiques soulevés par les algorithmes

*Coordination :*

*Henri Isaac, président de Renaissance Numérique*

*Jennyfer Chrétien, déléguée générale de Renaissance Numérique*

## Table des matières

Introduction - <i>Les prémices d'un débat continu</i> .....	2
I. Réalité de l'intelligence artificielle .....	4
a- Les intelligences artificielles .....	4
b- Enjeux éthiques de l'intelligence artificielle .....	6
II. Vers la transformation des rapports au travail .....	7
a- Une profonde incertitude quant à l'impact réel sur le marché du travail.....	7
b- Redéfinir notre rapport au travail et la notion de travail .....	11
III. Les Ressources Humaines (RH) : la transformation d'une filière complexe aux enjeux profondément sociétaux .....	14
a- De l'optimisation du processus RH à l'amélioration de la politique de l'emploi	14
b- Des enjeux éthiques prégnants .....	16
1. Un secteur directement lié aux droits fondamentaux de la personne.....	16
2. Entre ouverture des process humains et risques de renforcement des biais existants .....	18
Conclusion – <i>Pour une approche « human-in-command »</i> .....	20
Synthèse des pistes de recommandations .....	21
Remerciements .....	23

## Introduction - *Les prémices d'un débat continu*

Les technologies d'intelligence artificielle (IA) ont une influence croissante sur la société. Qu'il s'agisse d'applications GPS, d'outils de traduction, d'agents conversationnels, de recommandation de contenus ou demain de voitures autonomes, leurs applications n'ont de cesse de se développer dans notre quotidien. Au regard de cette accélération et de la nature de ces innovations, il est légitime de s'interroger, en tant que citoyens et acteurs de la société numérique, sur les impacts de cette transformation, pour l'individu et la collectivité.

C'est dans cet esprit que la Commission nationale de l'informatique et des libertés (CNIL) a lancé en janvier 2017, un débat public sur les enjeux éthiques soulevés par les algorithmes et l'intelligence artificielle. Ce dernier vise à dresser un panorama général des problématiques éthiques soulevées par ces technologies et à porter des premières recommandations pour accompagner leur développement dans un cadre éthique. C'est dans ce contexte que s'inscrit cette contribution. Au regard de sa dimension sociétale forte, elle s'attache tout particulièrement à interroger les enjeux dans le secteur de l'emploi et des ressources humaines. Elle est issue des débats qui se sont tenus lors de la matinée-débat organisée par le think tank Renaissance Numérique et le groupe Randstad en France, « Le travail à l'ère des algorithmes : quelle éthique pour l'emploi ? », le 5 septembre 2017. Ces réflexions font également écho aux premiers travaux du think tank, sur le travail à l'ère digitale<sup>1</sup>.

Il s'agit de porter ici une vision réconciliatrice autour de l'intelligence artificielle. Loin d'un angélisme aveugle, il devient essentiel d'instaurer un cadre de confiance autour de ces technologies, afin d'en tirer tous les bénéfices pour la société. Cette note vise à éclairer le débat, à donner réalité à la place qu'occupe ces technologies dans notre société, et ainsi dépasser des fantasmes sur un avenir empreint par l'affranchissement du robot par rapport à l'homme, qui demeurent prégnants dans le débat public et répondent bien souvent à des intérêts particuliers<sup>2</sup>.

Les progrès permis par les technologies de l'IA sont multiples et se sont développés dans de nombreux secteurs de la société, à l'instar de la santé ou de l'énergie. D'un point de vue économique, elles constituent un axe incontournable d'une politique industrielle future, conquérante de nouveaux marchés. En témoigne la masse des investissements sur ces questions aux Etats-Unis et en Chine qui laisse peu de doute sur cette question. Selon le Comité économique et social européen, le marché de l'IA s'élève aujourd'hui à environ 664 millions de dollars et devrait atteindre 38,8 milliards en 2025<sup>3</sup>. Bien qu'il soit difficile d'avancer un chiffre avec certitude étant donné le caractère diffus des technologies de l'IA, l'essor de ces technologies reste quasi-certain

---

<sup>1</sup> « La question du travail à l'ère digitale », Henri Isaac, Renaissance Numérique, juillet 2016.

<sup>2</sup> Olivier Ezratty rattache cela aux « *mécanismes de propagande de l'innovation* ». « Douze mythes de l'intelligence artificielle », Olivier Ezratty, Billet, Opinion Libres – Le blog d'Olivier Ezratty, 29 septembre 2017.

<sup>3</sup> « Les retombées de l'intelligence artificielle pour le marché unique (numérique), la production, la consommation, l'emploi et la société », Comité économique et social européen, Avis d'initiative, 31 mai 2017.

et la France doit adopter une approche résolue pour en faire un axe européen majeur de croissance économique.

Or, notre pays pâtit aujourd'hui d'un défaut généralisé de connaissance technique de ces technologies, qui risque de freiner les choix d'investissement et les avancées scientifiques en la matière. Un travail de pédagogie et de sensibilisation autour de ces enjeux est donc nécessaire, tant auprès des décideurs publics et privés que des citoyens. Il s'agit non seulement d'un enjeu d'acceptabilité sociale mais également d'égalité, si l'on ne veut pas augmenter la fracture déjà existante entre ceux qui ont le savoir et la maîtrise des usages et les autres.

Ce débat n'est d'ailleurs pas nouveau. Il traverse le secteur du numérique depuis ses débuts. Si les algorithmes ne sont eux-mêmes pas récents, c'est leur utilisation nouvelle et plus complexe qui questionne aujourd'hui. L'accélération massive de l'IA, grâce à l'augmentation des capacités de calcul, de stockage et de traitement des données, interroge en effet le cadre actuel. Il convient à ce propos de remarquer le lien étroit qui existe entre l'éthique et le droit autour de ces enjeux, en particulier concernant le traitement et la protection des données. Des principes établis par le cadre juridique existant et à venir avec la mise en application en mai 2018 du Règlement général sur la protection des données européen, tels la transparence des systèmes, le droit à l'oubli, à l'adaptation de ses données ou encore le droit au consentement des personnes, ne tiennent pas aujourd'hui compte des spécificités des technologies d'intelligence artificielle. La STOA (*Scientific Foresight Unit*) du Parlement européen, a ainsi identifié en 2016 six domaines dans lesquels l'Union européenne est susceptible de devoir réviser ou adapter ses textes du fait de l'évolution des domaines de la robotique, des systèmes cyber-physiques et de l'intelligence artificielle, à savoir : les transports, les systèmes de double utilisation, les libertés civiles, la sécurité, la santé et l'énergie<sup>4</sup>.

Il importe ainsi de préciser les termes du débat en introduction. Tout d'abord, il est intéressant de noter la variation sémantique entre acteurs pour désigner ces technologies – intelligence artificielle, algorithme, machine, système... - qui traduit le caractère extrêmement large que revêt le champ de l'IA et la diversité d'usage, si ce n'est confusion, selon la discipline ou profession. Dans cette contribution, nous privilégierons l'emploi de l'expression « technologies d'intelligence artificielle » qui traduit sa nature diverse. Rappelons par ailleurs que l'éthique revêt avant tout une dimension humaine. Il ne s'agit donc pas dans cette réflexion d'intégrer des dilemmes éthiques dans les technologies et de leur attribuer une personnalité juridique, au risque d'affecter le droit de la responsabilité civile<sup>5</sup>. Il convient également de distinguer l'éthique individuelle de l'éthique collective. Par exemple, concernant le droit au consentement, l'IA « *contribue à déplacer le consentement du niveau individuel de*

---

<sup>4</sup> "Ethical Aspects of Cyber-Physical Systems", Scientific Foresight Unit, Parlement européen, juin 2016.

<sup>5</sup> Le Comité économique et social européen évoque également « *des possibilités d'abus de même qu'un risque moral* ». « Les retombées de l'intelligence artificielle pour le marché unique (numérique), la production, la consommation, l'emploi et la société », Comité économique et social européen, Avis d'initiative, 31 mai 2017.

*l'usage de ses données personnelles, à un niveau collectif du consentement à ce que des systèmes informatiques puissent servir à orienter la société à partir d'observations globales de cette société* »<sup>6</sup>. Par ailleurs, si l'éthique n'est pas universelle, mais dépendante du contexte social et culturel dans lequel elle s'inscrit, les technologies d'intelligence artificielle ne connaissent, elles, pas de frontières. Aussi, il s'agit de s'interroger sous le prisme de nos valeurs, sur ce que devrait être une éthique de l'IA portée par l'Europe au niveau international. Enfin, notons qu'il ne s'agit d'un débat définitif. Il s'inscrit nécessairement dans le temps, au regard de l'augmentation continue de nos connaissances en matière d'IA et de l'analyse de ses effets. Par cette approche, il s'agit donc d'établir une « *vigilance éthique* », en instaurant les conditions propices au bon développement de ces technologies, plutôt que des normes trop restrictives qui pourraient vite devenir obsolètes<sup>7</sup>.

## I. Réalité de l'intelligence artificielle

### a- Les intelligences artificielles

Les applications basées sur de l'IA (moteurs de recherche, traductions automatiques, recommandations de contenus, assistants virtuels, etc.) sont déjà très présentes dans notre quotidien et se développent dans un grand nombre de secteurs à l'instar de la santé, de la banque ou des transports, sans nécessairement que nous en ayons conscience ou comprenions ses principes.

Un préalable à toute réflexion sur l'intelligence artificielle est donc de comprendre à quoi se réfère ces termes et son périmètre. Comme nous l'évoquions en introduction, il serait d'ailleurs plus juste de parler « des » intelligences artificielles, au regard de la diversité des technologies que couvre ce champ.

Une confusion est également souvent faite avec les notions de *big data* et d'algorithme. Le *big data* décrit l'ensemble des données qui se sont massivement accrues ces dernières années, dépassant les capacités de traitement humaines et offrant de nouvelles opportunités technologiques. Outre la volumétrie, le *big data* manifeste également le fait que les données sont de plus en plus variées en nature (données structurées, multimédias, non-structurées, etc.) et disponibles à des vitesses de plus en plus rapides (temps réel)<sup>8</sup>. Le terme algorithme est lui ancien et consiste au traitement des données, c'est-à-dire à une suite d'opérations permettant d'obtenir un résultat sur la base de ces dernières<sup>9</sup>.

---

<sup>6</sup> « Ethique de la recherche en apprentissage machine », CERNA, Rapport, juin 2017.

<sup>7</sup> « Ethique de la recherche en apprentissage machine », CERNA, Rapport, juin 2017.

<sup>8</sup> « Big Data : l'accélérateur d'innovation », Renaissance Numérique et G9+, décembre 2014.

<sup>9</sup> « Algorithmes : dans quelle société souhaitons-nous vivre ? », 14<sup>ème</sup> édition du cycle d'événements « Aux sources du numérique », organisé par Renaissance Numérique et Spintank, en présence de Serge Abiteboul et Gilles Dowek, auteurs du livre *Le Temps des Algorithmes* (Editions Le Pommier, 2017).

« Un algorithme est un procédé qui permet de résoudre un problème, sans avoir besoin d'inventer une solution à chaque fois. »

Serge Abiteboul et Gilles Dowek<sup>10</sup>

Enfin, l'intelligence artificielle revient à chercher à reproduire ou à imiter les capacités de l'homme, que ce soit en termes de perception, en raisonnement ou en génération d'informations.

Il existe deux grandes catégories d'IA :

- L'IA faible. Il s'agit d'une intelligence artificielle non sensible, qui se concentre sur une tâche précise (reconnaissance de formes, d'images, jouer aux échecs, jouer une partie de Go, etc.). Tous les systèmes existants utilisant de l'IA actuellement sont considérés comme des intelligences artificielles faibles.
- L'IA forte ou l'intelligence artificielle générale. Elle correspond à une machine capable d'appliquer l'intelligence à tout problème plutôt qu'un problème spécifique.

L'IA comporte par ailleurs plusieurs champs qui utilisent des approches et des méthodes différentes. On parle ainsi d'apprentissage automatique, apprentissage statistique ou encore d'apprentissage machine (*machine learning*).

La Commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique d'Allistène (CERNA) distingue trois grandes catégories d'intelligences artificielles, selon leur mode d'apprentissage de la machine : « supervisé », « non supervisé » et « par renforcement ». Dans l'apprentissage dit « supervisé », les données exploitées sont annotées préalablement par les concepteurs. En apprentissage « non supervisé », l'algorithme classe par lui-même les données. Enfin, l'apprentissage « par renforcement » consiste à ce que la machine apprenne par elle-même ce qu'il convient de faire en différentes situations, à partir d'un cumul d'expériences. Ces différents modes sont souvent combinés dans un même système.

Notons qu'au regard de la complexité de l'intelligence humaine, la recherche en IA est encore très loin d'avoir atteint des capacités équivalentes. Par ailleurs, ce développement n'est absolument pas linéaire et chacune des IA se développe à une vitesse différente. Si nous avons connu des évolutions récentes majeures en la matière, grâce notamment à une augmentation massive des données disponibles et à des outils de calcul beaucoup plus performants, nous n'en sommes par exemple qu'au début du *deep learning*<sup>11</sup>.

---

<sup>10</sup> *Le Temps des Algorithmes*, Serge Abiteboul et Gilles Dowek (Editions Le Pommier, 2017).

<sup>11</sup> Le *deep learning*, ou apprentissage profond, utilise des structures, appelées « réseaux neuronaux », fondées sur des fonctions du cerveau humain qui tirent leur apprentissage de l'expérience et du retour d'informations.

## b- Enjeux éthiques de l'intelligence artificielle

Le développement de ces nouvelles technologies présage de progrès considérables dans de nombreux secteurs, à l'instar des transports ou de l'éducation. Il offre de nouvelles capacités sociétales. Prenons l'exemple des véhicules autonomes qui pourront faciliter la mobilité de certaines catégories de la population aujourd'hui en difficulté (personnes âgées, personnes en situation de handicap, personnes isolées n'ayant pas accès à la mobilité en zone rurale, etc.). Au regard de la spécificité de ces technologies, il n'en demeure pas moins nécessaire de penser au cadre dans lequel elles se développent, en particulier à sa dimension éthique.

Certains de ces enjeux sont à rattacher de manière plus générale aux systèmes numériques, tels les questions de loyauté, équité, transparence, traçabilité, explicabilité, sécurité, contrôlabilité, conformité, responsabilité ou encore le respect de la vie privée. Dans son rapport de juin 2017, la CERNA établit un panorama assez précis de ces principes.

Toutefois, les technologies d'intelligence artificielle apportent également de nouveaux questionnements, notamment autour de la responsabilité des potentiels préjudices. Ainsi, si la responsabilité d'un mauvais entraînement de la machine revient à l'agent entraîneur, dans un système qui apprend en continu, plusieurs agents deviennent entraîneurs, à côté de l'entraîneur initial, les utilisateurs participent également à cet entraînement. De même, il devient plus difficile d'assurer le principe d'explicabilité et de transparence. Cette difficulté s'accroît avec la complexité des systèmes, à l'instar du *deep learning* qui laisse moins de traces. Le caractère évolutif de ces modèles et leur cumul potentiel entre eux et avec d'autres technologies vient également perturber leurs résultats et peut provoquer une certaine instabilité, si ce n'est des risques de sécurité en fonction des applications. En IA, il convient ainsi de distinguer les processus des algorithmes, dits « systèmes », des données. Il est difficile de savoir ce que font les systèmes. Ce que l'on peut savoir c'est quelles données ont été utilisées pour faire apprendre ce système.

*« Vous pouvez très bien regarder l'algorithme. C'est une série de lignes, de codes. Pour autant, ce n'est pas pour cela que vous allez comprendre ce qu'il fait. »*

Laurence Devillers, Chercheuse au LIMSI-CNRS – Membre de la CERNA<sup>12</sup>

Cela oblige le concepteur du code non seulement à comprendre ce que l'on fait et notamment à construire les technologies de façon éthique (« *ethics by design* »), mais également à instaurer des contrôles tout au long du processus. Une démarche éthique devrait ainsi permettre d'expérimenter ces technologies afin d'évaluer a posteriori leur efficacité et les risques potentiels. Des benchmarks peuvent être établis par le concepteur pour permettre d'anticiper ce que produira l'algorithme. Il s'agirait ainsi

---

<sup>12</sup> Matinée-débat « Le travail à l'ère des algorithmes : quelle éthique pour l'emploi », organisée par Renaissance Numérique et le groupe Randstad en France, le 5 septembre 2017 à Paris.

d'examiner les données en entrée et en sortie du système et mettre les technologies dans des situations extrêmes, afin de construire les frontières de réponses du système.

On peut ainsi distinguer deux types d'IA selon leur sensibilité éthique :

- Une technologie mûre, qui fait des tâches avec une fiabilité statistique de 95% et pour laquelle l'enjeu est qu'elle se diffuse le plus largement dans de nombreux secteurs pour réaliser des gains de productivité. Ici, nous devons nous demander si nous avons nécessairement besoin de comprendre l'ensemble du processus interne, dès lors que l'outil fonctionne pour des tâches de base pour lesquelles il n'y a pas d'aspect critique.
- Les autres IA concernent des fonctions plus complexes, qui ont trait à la prise de décision. Il sera alors nécessaire de prouver par des études suffisamment fiables que la décision automatique est suffisamment meilleure que la décision humaine.

*« Un compromis entre les capacités d'apprentissage et l'explicabilité s'avère nécessaire. Ce compromis est à apprécier en fonction des domaines d'applications : si l'explicabilité n'est a priori pas une exigence dans des applications telles le jeu, elle est impérative dès lors que l'on touche aux intérêts et droits des personnes ou à leur sécurité. »*

*« Ethique de la recherche en apprentissage machine », CERNA (juin 2017)*

## II. Vers la transformation des rapports au travail

Les défis posés par les technologies d'intelligence artificielle concernant l'emploi rejoignent des enjeux plus globaux, qui dépassent ce seul secteur. Avec 17,2% des jeunes à la recherche d'un emploi (19,4% dans la seule zone euro), les taux de chômage des moins de 25 ans demeurent très élevés dans l'Union européenne<sup>13</sup>. Ces chiffres augurent une génération « sacrifiée » si rien n'est fait et invitent à repenser de manière plus profonde à notre approche du travail.

### a- Une profonde incertitude quant à l'impact réel sur le marché du travail

Aujourd'hui, personne ne sait évaluer de manière précise quelles seront les conséquences de l'IA sur le marché du travail. Les derniers rapports sur la question révèlent en réalité une incertitude sur leurs impacts en termes de destruction et de transformation d'emploi : l'Organisation de coopération et de développement économiques (OCDE), 9% en moyenne des emplois menacés dans les 21 pays de l'OCDE<sup>14</sup> ; le Conseil d'orientation pour l'emploi (COE), 10% des emplois susceptibles d'être supprimés et 50% significativement transformés en France ; Roland Berger prévoit 3 millions d'emplois supprimés à horizon 2025 en France ; Frey et Osborne,

---

<sup>13</sup> Source Eurostat.

<sup>14</sup> Un emploi est dit « menacé » dès lors qu'au moins 70% des tâches peuvent être automatisées.



47 % des emplois américains profondément modifiés voire supprimés en 10 à 20 ans ; Bruegel, selon la méthode Frey et Osborne, la moitié des emplois en Europe ; etc.

Ces écarts s'expliquent par des méthodes d'estimation différentes. En particulier, de nombreuses études ne considèrent pas l'aspect multi-tâches d'un emploi : la suppression d'une tâche est alors considérée comme la suppression d'un emploi. En réalité, les métiers sont amenés à évoluer. Si certaines tâches disparaissent, il est probable que cet emploi reste pertinent au regard des autres tâches non automatisables.

Toutefois, les prédictions restent délicates car les intelligences artificielles sont difficilement assimilables aux modèles économiques traditionnels de l'innovation (habituellement linéaire et mono-produit), pour plusieurs raisons, et notamment :

- 1- Les technologies regroupées sous le nom IA sont en réalité multiples (*machine learning, deep learning, etc.*) et leur développement n'est pas endogène (les avancées dans un champ donné de l'IA peuvent alimenter les autres champs de l'IA).
- 2- Le développement d'une IA est tributaire à la fois des jeux de données détenus et des compétences humaines en la matière.
- 3- Elle peut également varier selon la finalité stratégique du développement de l'IA par un acteur : s'agit-il de se substituer purement et simplement à une fonction, à une activité voire un secteur économique (voitures autonomes vs. taxis, transports en commun, fret, etc.) ?
- 4- L'impact économique et social d'une IA peut à tout moment changer si elle s'articule avec le développement d'une autre technologie (ex. : *blockchain, ordinateur quantique, etc.*).

Il est donc impossible de quantifier de manière fiable le risque pour l'emploi dans 5 ans, 10 ans ou encore 20 ans.

Cette transformation n'est toutefois pas juste celle des technologies d'intelligence artificielle, mais concerne la numérisation de secteurs d'activité qui est plus large et a débuté depuis longtemps. L'incertitude en matière du nombre d'emplois détruits par l'IA doit ainsi être mise en perspective avec le retard qu'accuse la France en matière d'adoption et de transferts des technologies. Déjà dans les précédentes vagues d'innovation technologique, les entreprises françaises ont connu un certain retard dans l'adoption. Si l'industrie numérique française est bien positionnée dans la compétition internationale, l'économie française dans son ensemble est, elle, à la traîne. À titre d'exemple, la situation des PME françaises, qui représentent aujourd'hui 49% de l'emploi en France, est criante. Le *Digital Economy and Society Index 2017* de la Commission européenne classe ainsi le pays au 16<sup>ème</sup> rang en termes de numérisation de l'économie. Le pays se classe également 16<sup>ème</sup> sur l'intégration du *cloud* dans ses entreprises, avec seulement 12% de ses entreprises utilisant l'informatique en nuage pour optimiser leurs coûts. L'adoption de l'IA peut donc paraître prématurée et difficile à appréhender à des acteurs qui n'ont pas su forcément s'adapter aux précédentes vagues d'innovations technologiques.



Mais il est à noter qu'à la différence d'autres technologies, les technologies de l'IA sont diverses et ne sont pas conçues pour une seule activité particulière. Elle devrait donc avoir un impact dans de nombreux secteurs. Afin d'encourager ce déploiement, il convient toutefois de faciliter son transfert technologique, en la rendant largement accessible aux acteurs privés comme publics dans des logiques d'*Open source* et d'*Open data*. A cet égard, des initiatives collaboratives privées à l'instar de *TensorFlow* (Google), *Big Sur* (Facebook) ou encore *Open LA* (Elon Musk), contribuent à cette démarche. Cependant, si cette ouverture est clé et participe de la transparence de ces technologies, elle demeure contrainte par l'accessibilité aux données nécessaires pour entraîner ces systèmes.

*« Les technologies d'intelligence artificielle peuvent être appliquées à énormément de champs différents et c'est cela aussi qui en fait la force par rapport à d'autres technologies. Elles vont être extrêmement généralisées, c'est-à-dire qu'on les utilisera dans l'agriculture, dans la santé, dans les services et dans l'industrie. »*

Lionel Janin,  
Adjoint au directeur du Département Développement durable et Numérique  
de France Stratégie<sup>15</sup>

On devrait ainsi assister à une transformation massive des emplois plutôt qu'à une destruction. Un emploi couvrant souvent plusieurs activités, ces technologies devraient venir le plus souvent se substituer à certaines d'entre elles, des tâches souvent répétitives et facilement automatisables, avec pour conséquence de faire évoluer le contenu du travail, voire de l'enrichir. L'exemple de l'imagerie médicale est à ce sujet parlant : si la machine aide au diagnostic, elle ne pourra pas en effet faire la prescription qui comporte des données complexes qui nécessitent une supervision humaine. Il convient également de noter un mouvement de transfert d'un pan du travail vers les clients, comme le montrent les exemples de la caisse automatique ou encore le secteur de l'économie collaborative<sup>16</sup>.

La spécificité de cette révolution sur l'emploi est aussi liée au fait qu'elle concerne toutes les catégories de travailleurs. Comme l'ont démontré Erik Brynjolfsson et Andrew McAfee (MIT), nous avons atteint la « *deuxième ère du machinisme* ». Alors que le premier temps a consisté essentiellement au remplacement de la force de travail musculaire, le deuxième temps consiste aujourd'hui au remplacement de prestations intellectuelles et compétences cognitives, impactant ainsi à la fois les travailleurs peu qualifiés et les travailleurs diplômés<sup>17</sup>.

---

<sup>15</sup> Matinée-débat « Le travail à l'ère des algorithmes : quelle éthique pour l'emploi », organisée par Renaissance Numérique et le groupe Randstad en France, le 5 septembre 2017 à Paris.

<sup>16</sup> « Le travail à l'ère digitale », Henri Isaac, Renaissance Numérique, 16 juillet 2016.

<sup>17</sup> *The Second Machine Age - Work, Progress, and Prosperity in a Time of Brilliant Technologies*, Erik Brynjolfsson et Andrew McAfee, W. W. Norton & Company (2014).

Parallèlement, on ignore le nombre d'emplois que créeront les nouveaux secteurs. S'il est partagé que de nouveaux besoins et de nouveaux métiers devraient apparaître, il est difficile de prédire en quoi ils consisteront, leur nombre et sur quelle temporalité ils apparaîtront. Toutefois, il convient de préparer au mieux les conditions de ce déploiement. Or, le pays pâtit d'une vraie lacune dans les compétences de ses travailleurs.

Selon les statistiques de la Commission européenne basées sur le référentiel européen *DigComp*, 8% de la population active française n'aurait ainsi aucune compétence numérique et 27% un niveau très faible. Dans son dernier rapport, le Conseil d'orientation de l'emploi rapporte de son côté la pénurie croissante de compétences expertes dans les nouvelles technologies, qu'il s'agisse du secteur du numérique ou des secteurs utilisateurs<sup>18</sup>. Parmi les compétences qui seraient les plus demandées, on retrouve celles liées à la gestion et l'exploitation des données, à la programmation et celles liées à la conception et la maintenance des logiciels. Le COE montre également que les technologies seront à l'origine d'une « *demande accrue généralisée* » pour des compétences transversales, qu'elles soient sociales (travail en équipe, intelligence sociale) ou situationnelles (autonomie, adaptabilité). Or, le secteur du numérique connaît déjà une véritable pénurie de talents en France pour répondre aux offres d'emplois : développeurs, intégrateurs, *data scientists*, ... sans compter ces métiers qui n'existent pas encore. Cette situation porte nécessairement atteinte au développement de l'IA dans notre pays.

Il convient enfin de noter que cet élan n'est pas uniforme selon les secteurs et acteurs. Au-delà de la question des moyens et de la mise en capacité, il est dépendant de deux facteurs. Le premier est comment l'entreprise ou l'organisation va s'emparer de ces outils. Ce n'est pas parce qu'une tâche est automatisable que l'on va nécessairement l'automatiser. Cela renvoie à des questions stratégiques pour l'organisation et à sa culture, des valeurs, des questions éthiques qui ne sont pas liées qu'aux algorithmes. Cela peut par ailleurs ne pas avoir lieu si ce n'est pas économiquement utile et/ou si cela n'est pas socialement accepté. Cette transformation dépendra également de la nature des métiers. Sur ce point, il est essentiel d'identifier dès à présent quels secteurs seront influencés, à quel degré et quelle échéance, pour anticiper les conséquences potentielles sur l'emploi (contenu du travail, compétences nécessaires, régimes sociaux, etc.) et apporter les solutions adéquates. Ce travail doit s'inscrire dans une démarche collaborative avec toutes les parties prenantes - pouvoirs publics, partenaires sociaux, organismes de formation et experts de l'IA et du numérique – et participer d'une GPEC renouvelée.

---

<sup>18</sup> « Automatisation, numérisation et emploi – Tome 2 : l'impact sur les compétences », Conseil d'orientation pour l'emploi, Rapport, septembre 2017.

## **b- Redéfinir notre rapport au travail et la notion de travail**

Au-delà de la question de l'emploi, cette évolution nous invite à repenser notre propre rapport au travail et à la définition même de ce dernier. Cela interroge la manière dont nous qualifions la valeur du travail. Qu'allons-nous rémunérer demain ? Quelle juste répartition d'une richesse produite par l'IA, entre le travail et le capital ? Comment mesurer la contribution d'un salarié dans un monde de plus en plus orienté vers les services ? Comment mesurer demain sa contribution s'il travaille majoritairement avec un robot ?

Nous le savons, les technologies d'intelligence artificielle vont jouer un rôle croissant dans le travail. À côté des applications que l'on demande déjà aux travailleurs d'utiliser au quotidien (sans qu'ils aient parfois conscience d'utiliser cette technologie), ils en utiliseront de plus en plus pour bien faire leur travail. Ces applications ont ainsi une influence croissante sur le cadre du travail. Si elles sont la promesse de l'amélioration de certaines conditions de travail, avec notamment la suppression de tâches pénibles et/ou répétitives, il convient de rester vigilant à ce que certaines de ces applications ne se traduisent par une surveillance accrue des travailleurs, au risque d'une perte d'autonomie et de créativité, et de déshumaniser notre rapport au travail.

*« Le chemin qui s'ouvre doit avoir pour but de préserver et d'augmenter la valeur du travail humain, et de concilier de façon innovante une économie compétitive et une amélioration du bien-être dans une société inclusive. »*

Conseil national du numérique<sup>19</sup>

Par ailleurs, il convient de remarquer que nous ne sommes pas tous égaux dans le travail face à ces technologies. Déjà, nous ne sommes pas tous touchés de la même façon selon la nature de notre métier. Ensuite, nous ne partageons pas le même niveau de compétences. Une étude du COE montre ainsi que les emplois les plus exposés au risque d'automatisation sont le plus souvent des métiers peu qualifiés et que la transformation en cours de la structure des emplois favorise les emplois les plus qualifiés.<sup>20</sup> Avec l'intégration croissante de ces technologies dans les différents secteurs d'activité, il va devenir nécessaire d'avoir la capacité de s'adapter et de comprendre ces dernières, la façon dont les algorithmes « travaillent » avec les données, les biais possibles dans les résultats obtenus. Une attention particulière doit ainsi être portée aux travailleurs peu qualifiés, si l'on ne souhaite pas accroître une fracture déjà prégnante. Cette évolution n'est donc pas seulement technique, elle est également sociale, sociétale et philosophique. Elle nous invite à penser la collaboration de l'humain avec la technologie afin de ne pas être réduit à la surveillance des machines.

Cette adaptation passe nécessairement par une transformation profonde, si ce n'est révolution, de notre appareil de formation actuel. Espérons que la réforme à venir du

---

<sup>19</sup> Rapport de synthèse France Intelligence Artificielle, mars 2017.

<sup>20</sup> « Automatisation, numérisation et emploi – Tome 1 : les impacts sur le volume, la structure et la localisation de l'emploi », Conseil d'orientation pour l'emploi, Rapport, septembre 2017.

gouvernement français portera une ambition suffisante pour répondre à l'urgence et anticiper les évolutions à venir.

« *Nous devons infuser une meilleure connaissance de l'intelligence artificielle à tous les niveaux de la société, à des degrés différents d'expertise, de connaissance et de compréhension.* »

Amal Taleb,  
Directrice adjointe des affaires publiques de SAP France<sup>21</sup>

La défiance vis-à-vis de l'IA est aujourd'hui alimentée par une méconnaissance de ces technologies, qui laisse toute leur place aux discours les plus anxiogènes. Sans maîtrise, pas de confiance, cette dernière passe nécessairement par une meilleure compréhension technique. Si nous souhaitons nous saisir des opportunités offertes par ces technologies, il convient donc d'intégrer le champ de l'IA dans toutes les étapes de l'éducation et la formation continue.

Cette problématique se retrouve d'ailleurs sur le numérique de manière plus globale. Il s'agit de travailler tant sur la montée en qualification générale que d'agir sur les compétences. La formation aux STIM (science, technologie, ingénierie et mathématiques) doit être démocratisée, afin de permettre aux travailleurs de demain de mieux appréhender leurs rapports avec ces technologies, de préserver leur autonomie et capacité de contrôle vis-à-vis de ces dernières. Cette formation n'est pas nécessairement que technique, mais doit également s'intéresser aux usages qu'il en est fait : comprendre et savoir interpréter les résultats, expliquer et communiquer le sens des décisions.

Le caractère évolutif de ces technologies oblige à dépasser la question des seules compétences techniques, pour lesquelles nous avons de véritables lacunes en France comme nous l'avons vu précédemment. Les attentes en compétences transversales sont aujourd'hui croissantes. Alors que les besoins en compétences « physiques » se réduisent, les attentes de compétences cognitives (littératie, numératie, résolution des problèmes) et de compétences sociales et situationnelles, se multiplient.

Pour le COE, il existe deux raisons à cette évolution. La première est que les tâches nécessitant ces compétences sont difficilement automatisables, et donc leur valeur augmente avec le déplacement de la « *frontière technologique* ». Par ailleurs, cette mutation technologique, en transformant profondément l'organisation du travail, participe à accroître la demande générale de ces compétences, qui jusqu'alors concernaient plutôt les professions les plus qualifiées. Or, 13% des actifs en emploi en France (soit 3,3 millions de personnes), ont aujourd'hui un niveau de maîtrise en numératie et littératie qui pourrait les mettre en difficulté dans leur emploi actuel ou dans la recherche d'un nouvel emploi<sup>22</sup>. Face à des métiers qui seront en perpétuelle mutation et au regard de la vitesse de cette dernière, nous devons apprendre aux personnes à apprendre, c'est-à-dire à s'auto-former, à être capable de mobiliser leurs

---

<sup>21</sup> Matinée-débat « Le travail à l'ère des algorithmes : quelle éthique pour l'emploi », organisée par Renaissance Numérique et le groupe Randstad en France, le 5 septembre 2017 à Paris.

<sup>22</sup> « Automatisation, numérisation et emploi – Tome 2 : l'impact sur les compétences », Conseil d'orientation pour l'emploi, Rapport, septembre 2017.

savoirs et en apprendre d'autres. Au regard de la teneur multidimensionnelle des applications de ces technologies (technique, éthique, juridique...), la formation se doit également d'être pluridisciplinaire, par l'intégration notamment des humanités. Il ne s'agit pas seulement de savoir les utiliser, mais de savoir les utiliser de manière responsable, avec un esprit critique. Il s'agit également de ne pas perdre en créativité humaine face à un risque de « standardisation ».

Au-delà de la formation des travailleurs de demain, il y a urgence à adapter les politiques de formation, au risque d'exclure un grand pan des travailleurs actuels de cette transformation. La réforme du gouvernement à venir doit porter une vision beaucoup plus agile des systèmes de formation, afin de permettre de requalifier, d'aider les travailleurs à évoluer dans leurs compétences au gré de l'évolution de leurs métiers. Ce nouveau rapport au travail, et notamment la collaboration avec les technologies, doit également nous inviter à repenser les métriques d'évaluation des personnes dans leur travail. Les missions et objectifs seront-ils les mêmes ?

*« L'IA est un chantier de formation professionnelle essentiel qui peut servir de cas d'école pour les nouvelles visions de l'éducation et de la formation qui sont depuis plusieurs années débattues dans la société : importance des compétences transversales (soft skills), littératie numérique, à laquelle il faut adjoindre une littératie de l'IA. Une telle formation passe par une éducation informatique générale et une mobilisation de la société sur les formations initiales (entreprises, secteur de l'éducation populaire) mais aussi une inventivité dans les formations sur poste, les formations en conception, ce qui implique une forte mobilisation des entreprises elles-mêmes, de la recherche et des partenaires sociaux pour aboutir à de nouvelles voies. »*

Rapport de synthèse France Intelligence Artificielle (mars 2017)

Enfin, cette adoption ne peut être *top-down*. Des dispositifs de test en situation réelle sont aujourd'hui déployés dans certaines entreprises. Cette approche, à la manière de *living labs*, doit être encouragée et généralisée dans les différents métiers. La Direction des Ressources Humaines de SNCF Réseau et la CFDT-Cadres ont à ce titre lancé une initiative intéressante avec Sciences Po et le Cabinet Ytae : un laboratoire d'idées, le Digilab Social, visant à anticiper ce que sera demain le périmètre des fonctions de manager et de représentants du personnel dans différents secteurs d'activité. Il propose ainsi de nouvelles pratiques à expérimenter. Cette mise en situation permet d'anticiper la réalité sur le terrain, en fonction de la spécificité des filières et du rôle des différents acteurs. Ce sont les mêmes acteurs qui seront amenés à utiliser ces technologies, qui participent à ces expériences. Cette démarche contribue à l'appropriation par chacun de ces outils et usages et ainsi au succès de l'introduction de ces technologies dans les entreprises.

Le débat est en cours. Interroger notre rapport au travail pose de manière plus générale la question de nos institutions traditionnelles. Quand notre modèle de

protection sociale est aujourd'hui basé majoritairement sur le salariat, que faire demain avec des indépendants et des robots ?<sup>23</sup>

### III. Les Ressources Humaines (RH) : la transformation d'une filière complexe aux enjeux profondément sociétaux

Poser la question de l'emploi revient également à s'interroger sur la mutation de ses acteurs. La filière des RH est un cas d'étude particulièrement intéressant, au regard des dimensions sensibles qu'elle couvre, qu'elles soient économiques, sociales, éthiques ou encore politiques. C'est par ailleurs une filière qui a déjà entamé une réflexion profonde sur sa transformation liée à l'essor des technologies d'intelligence artificielle, en particulier du côté des porteurs de solutions.

#### a- De l'optimisation du processus RH à l'amélioration de la politique de l'emploi

Si les solutions RH intégrant des technologies d'intelligence artificielle se multiplient, nous n'en sommes qu'aux prémices de leur déploiement. En effet, le nombre d'entreprises et organisations qui ont aujourd'hui les capacités réelles d'intégration et d'utilisation de ces applications est encore limité.

Les acteurs du secteur s'accordent toutefois sur le fait que ces technologies prendront une place croissante dans le processus RH, au regard de leur contribution pour améliorer et optimiser ce dernier, et ce dans toutes ses dimensions : recrutement, engagement des salariés, développement de modèles de rétention, parcours de mobilité, rémunération, gestion des compétences. Parmi les solutions les plus récentes, on peut noter l'essor du concept de « graph d'entreprise », porté par les éditeurs de logiciel, à l'instar de Microsoft ou de Slack, ou de nouveaux arrivants sur le marché comme Facebook. Par l'analyse des interactions digitales des salariés avec leur écosystème, ces technologies permettent d'obtenir une cartographie fine de leur profil, domaines d'expertise et d'intervention, et ainsi de personnaliser et prioriser leur contenu. Avec leur ouverture croissante à des applications tiers, ces solutions permettent également d'établir une connaissance beaucoup plus approfondie de l'écosystème de l'entreprise et de l'emploi dans sa globalité. Ainsi, Microsoft, au travers du Microsoft 365 Graph, réunit un nombre croissant de données d'usage liées à ses applications, telles Cortana ou LinkedIn qu'ils ont récemment intégré.

*« Microsoft 365 Graph fait désormais partie des fondations sur lesquelles nous bâtissons nos applications. C'est le cas avec Cortana, mais aussi avec Bing for Business et Microsoft Teams, et maintenant avec LinkedIn dont les données ont été récemment intégrées au Microsoft Graph. »*

---

<sup>23</sup> « Le travail à l'ère digitale », Henri Isaac, Renaissance Numérique, 16 juillet 2016.



Harry Shum,  
Vice-président exécutif du groupe IA et recherche de Microsoft<sup>24</sup>

Comme dans un certain nombre de secteurs, l'IA devrait apporter dans ce domaine des gains de productivité importants, notamment en termes de gain de temps. Sont également attendus des effets positifs quant à « l'ouverture » du processus RH, en particulier dans le recrutement où les idées préconçues sont souvent très présentes. Un éclairage par la donnée devrait ainsi contribuer à élargir le spectre du recruteur et donc les profils retenus. Du côté des candidats, ces solutions devraient leur permettre de mieux identifier les besoins et les attentes de leur futur employeur, à condition bien entendu que ce processus soit suffisamment explicité.

*« A côté des algorithmes classiques d'analyse des données existantes, les entreprises ont encore à investir le terrain des algorithmes prédictifs pour alimenter leur vision stratégique, améliorer leur politique RH, et plus particulièrement de recrutement et de rétention des talents, ou faire de la vraie GPEC par exemple. »*

Jérôme Chemin,  
Secrétaire national de la CFDT-Cadres<sup>25</sup>

Au niveau macroéconomique, le *big data* et l'IA offrent aujourd'hui des solutions nouvelles de cartographie de l'emploi, à l'instar des graphs d'entreprise. En cela, ces technologies constituent un levier important pour les politiques de l'emploi et la formation. Elles permettent notamment d'améliorer la gestion des transitions professionnelles et donc d'éviter les ruptures professionnelles. Elles devraient ainsi participer à orienter notre modèle social de l'emploi vers un modèle préventif, économiquement plus viable qu'un modèle curatif.

*« Il y a un potentiel important au travers des plateformes RH quand elles sont mises au service de l'intérêt commun et notamment des pouvoirs publics pour mieux comprendre un bassin d'emploi. Un bassin d'emploi, cela peut être au niveau mondial puisque la plateforme est au niveau mondial, au niveau d'un État ou d'une région. »*

Laurence Lafont,  
Directrice Marketing & Opérations de Microsoft France<sup>26</sup>

Les plateformes possèdent de vastes bases de données sur l'état du marché du travail. Cela leur permet d'identifier des tendances macroéconomiques, telles que les besoins de compétences spécifiques ou la nécessité de formation dans certains secteurs ou bassins d'emploi. Elles permettent ainsi d'accroître l'efficacité et le ciblage des politiques publiques, en complémentarité avec les autres acteurs de l'emploi. A titre d'exemple, LinkedIn a créé un outil appelé "Economic Graph" visant à pourvoir les

---

<sup>24</sup> « 2017, l'année des graphs d'entreprise », Antoine Crochet-Damais, JDN, 20 octobre 2017.

<sup>25</sup> Matinée-débat « Le travail à l'ère des algorithmes : quelle éthique pour l'emploi », organisée par Renaissance Numérique et le groupe Randstad en France, le 5 septembre 2017 à Paris.

<sup>26</sup> Matinée-débat « Le travail à l'ère des algorithmes : quelle éthique pour l'emploi », organisée par Renaissance Numérique et le groupe Randstad en France, le 5 septembre 2017 à Paris.



collectivités en données sur les différents marchés du travail. Il est évident que ces documents ne représentent pas l'ensemble des travailleurs, et concernent particulièrement les « cols blancs » qui forment à ce jour la majorité des utilisateurs de la plateforme. Cependant, les travailleurs sont de plus en plus nombreux à utiliser ces applications et il est nécessaire d'assurer un suivi de ces travailleurs pour comprendre leurs trajectoires et leur proposer des formations adaptées. Il est à noter qu'au niveau microéconomique, les données pourraient elles-mêmes être mieux utilisées dans le domaine RH. Les parties s'accordent sur une grande déperdition entre la matière.

De manière générale, on peut distinguer trois grandes catégories de données RH :

- Des données tangibles, comme les compétences techniques. Elles sont facilement mesurables et intégrables. Toutefois, elles ne suffisent pas pour réussir un recrutement et expliquent sans doute la majorité des échecs en ce domaine. Imaginez par exemple un chauffagiste incapable de parler correctement à son client ou un cadre au tempérament peu coopératif plongé dans une start-up...
- Des données statistiques, non nominatives, qui n'entrent pas dans le champ de la réglementation relative à la protection des données, au même titre que certaines données publiques comme celles de l'INSEE ou de l'ACROSS.
- Des données plus complexes et plus délicates à manipuler, d'ordre comportemental et cognitif. Elles permettent de juger du savoir-être d'un candidat en modélisant ses modes d'interaction humaine.

Ces données sont plus ou moins complexes à collecter et à traiter. Les technologies d'intelligence artificielle tendent à les associer afin d'affiner toujours plus leurs résultats. Selon leur nature ou leur combinaison, la question éthique se posera alors à des degrés différents.

## **b- Des enjeux éthiques prégnants**

### **1. Un secteur directement lié aux droits fondamentaux de la personne**

La question éthique est étroitement liée à celle du droit. Cela est particulièrement vrai dans le champ des ressources humaines, tant ce secteur est lié à nos droits fondamentaux.

*« Les droits des personnes ne sont pas virtuels dans le monde virtuel. Ce qui s'applique en matière de recrutement doit s'appliquer, y compris quand le recrutement utilise des aspects numériques, quand il est éventuellement dématérialisé. »*

Sarah Bénichou,  
Cheffe de pôle au Département Promotion de l'égalité et de l'accès au droit  
du Défenseur des Droits<sup>27</sup>

---

<sup>27</sup> Matinée-débat « Le travail à l'ère des algorithmes : quelle éthique pour l'emploi », organisée par Renaissance Numérique et le groupe Randstad en France, le 5 septembre 2017 à Paris.

Le Défenseur des droits relève ainsi un certain nombre de droits fondamentaux qui s'appliquent au secteur de l'emploi et rejoignent des enjeux éthiques liés aux technologies d'intelligence artificielle, parmi lesquels :

- Le droit à la vie privée. Un article dans le Code du travail prévoit en effet qu'il est interdit de collecter des informations dites « non pertinentes » dans le recrutement. Il indique par ailleurs que les méthodes de recrutement doivent être pertinentes et connues des candidats. Concernant l'usage de l'IA, cela impose ainsi un partage de l'information sur la technologie utilisée et un langage commun entre les parties (DRH, *data scientist*, candidat, etc.), afin que celle-ci soit intelligible pour tous et que chacun puisse participer de manière éclairée au processus.
- L'interdiction de la discrimination. Sur ce point, on distingue deux types de discriminations. La première est la discrimination dite « directe ». Elle revient à traiter différemment une personne (ne pas la recruter, ne pas la promouvoir, ne pas la former, etc.) parce qu'elle relève d'une catégorie en lien avec une caractéristique protégée par le droit français<sup>28</sup>. Dans ce cadre, une démarche éthique reviendrait à s'assurer que les données transmises à l'algorithme ne relèvent pas strictement d'une discrimination préconçue. Le risque de discrimination directe des technologies d'intelligence artificielle est donc ici moindre. Il n'en demeure pas moins possible malgré une vigilance, tant la liste des motifs discriminatoires est large. L'autre dimension de la discrimination est la discrimination dite « indirecte ». Il s'agit de critères ou de mesures neutres en apparence, mais qui auront un impact particulièrement désavantageux pour une catégorie de personnes protégées par le droit (par exemple, le critère « grandes écoles »). L'éthique des algorithmes consiste ici à s'intéresser tant à la nature des données qu'à leur combinaison et aux résultats au sortir de leur traitement par les algorithmes.

L'essor des technologies d'intelligence artificielle questionne ainsi le droit dans le travail. Une difficulté relève notamment de l'explicabilité. En droit, la discrimination sera prouvée à partir du moment où la personne présentera des faits laissant supposer qu'effectivement elle a été victime d'une sélection discriminatoire. Puis, c'est au tour de l'employeur de démontrer qu'il n'a pas discriminé, en justifiant ses choix de critères qui ont un impact discriminatoire. Le droit repose ainsi sur des principes de transparence, d'objectivité et de traçabilité. Or, les intelligences artificielles, et en particulier les systèmes auto-apprenants, limitent la transparence des processus.

---

<sup>28</sup> On compte aujourd'hui plus de vingt-cinq critères discriminatoires en France : âge, sexe, origine, handicap, etc. : <https://www.defenseurdesdroits.fr/fr/institution/competences/lutte-contre-discriminations>

## 2. Entre ouverture des process humains et risques de renforcement des biais existants

*« Nous sommes sur un secteur éminemment complexe, qui se complexifie d'autant plus qu'il y a de plus en plus de multi-activité, des vœux de transition professionnelle. Ce ne sont pas des trajectoires linéaires, quels que soient d'ailleurs les niveaux de compétences et d'expérience. Cette richesse aujourd'hui elle ne peut pas se résumer comme étant un traitement algorithmique. »*

Bernard Lamberti,  
Directeur de la stratégie de l'APEC<sup>29</sup>

En permettant de corriger les biais inhérents à la société, l'utilisation de l'IA participe à renforcer l'éthique dans le champ des RH. La distanciation permise par cette technologie permet en effet d'éliminer les biais de discrimination dont tous les humains sont porteurs, préjugés que l'on mobilise souvent de manière inconsciente et qui font que l'on discrimine et que l'on recrute à peu près à son image la plupart du temps. Les technologies d'intelligence artificielle permettent de rapprocher des succès, des savoir-faire et savoir-être fondamentaux pour un poste. Le déploiement des IA conduit à repenser la manière dont on recrute, dont on fait le *sourcing* et pense les profils. Certains acteurs réfléchissent à ce propos à leur donner un rôle même plus important dans la lutte contre les discriminations, en leur faisant faire de « l'action positive », à condition bien entendu de respecter le droit français à cet égard, c'est-à-dire sur le principe de « à compétence égale ». La technologie d'intelligence artificielle pourrait ainsi prévoir qu'à compétence égale, elle proposera plus de femmes, de personnes en situation de handicap, etc.

Toutefois, les IA peuvent elles-mêmes être porteuses de biais. D'autant que le prédictif c'est aussi ce que fait la majorité, d'où le risque de discrimination d'individus par un recrutement homogène, peu diversifié. Le premier risque concerne ainsi la collecte, la qualité et la pertinence des données analysées. Il s'agit ici essentiellement de relais de biais existants chez l'homme. Il convient en amont du processus de veiller à la nature des données et aux conditions de collecte. Cette information doit elle-même être partagée avec les parties prenantes, soit directement, soit par l'intermédiaire d'un tiers de confiance afin de limiter le risque de concurrence dans certains cas.

*« Il convient à cet égard de noter qu'à l'heure actuelle, le développement de l'IA se déroule dans un environnement homogène, composé principalement de jeunes hommes blancs, ce qui contribue à ancrer (sciemment ou non) les disparités culturelles et de genre dans l'IA, notamment du fait que les systèmes d'IA apprennent sur la base de données de formation. »*

Comité économique et social européen<sup>30</sup>

---

<sup>29</sup> Matinée-débat « Le travail à l'ère des algorithmes : quelle éthique pour l'emploi », organisée par Renaissance Numérique et le groupe Randstad en France, le 5 septembre 2017 à Paris.

<sup>30</sup> « Les retombées de l'intelligence artificielle pour le marché unique (numérique), la production, la consommation, l'emploi et la société », Comité économique et social européen, Avis d'initiative, 31 mai 2017.

L'autre risque concerne le traitement qu'il est fait de ces données, qu'il s'agisse de relais de biais humains par le choix initial des modes de calcul, ou de biais créés directement par les technologies d'intelligence artificielle. Notons que les premiers, sur lesquels l'homme a directement la main, sont plus faciles à corriger que les seconds.

Le préalable à une démarche éthique est ainsi de s'assurer du caractère non discriminatoire des données directement, mais également indirectement, dans leur corrélation et les corrélations que l'on va mettre en œuvre. Des données qui nous semblent objectives en tant que telles peuvent traduire des inégalités qui existent dans la société. Il convient alors de s'attacher à conserver des signaux faibles. De plus, ces biais discriminatoires ne vont pas toujours se voir. Il est donc nécessaire de maintenir une supervision humaine. Il s'agirait de mettre en place au sein des entreprises et des organisations des services de certification pluridisciplinaires (RH, juristes, ingénieurs, représentants des salariés, etc.) qui s'assureraient de la pertinence des données, de leur caractère non discriminatoire et de leur pondération dans le résultat général.

Cependant, dans des applications d'intelligence artificielle, il n'est pas toujours possible de détecter directement les biais discriminatoires. Elles peuvent engendrer des corrélations à l'impact discriminatoire auxquelles on n'a pas pensé et que l'on ne peut pas percevoir en avance. Si le droit interdit certains critères discriminants dans l'emploi, la technologie peut dans son processus reconstruire les valeurs de ces variables. Une évaluation doit donc se faire tout au long du processus. Il est par ailleurs indispensable d'avoir une compétence métier (RH), pour confronter les résultats à la sortie de ce dernier.

La technologie d'intelligence artificielle donne une prédiction. Il faut savoir après comment cette prédiction est utilisée. D'autant que, comme le soulignent les chercheurs de la CERNA, des principes sont particulièrement difficiles à spécifier, à l'instar de l'équité qui nécessite une appréciation humaine. Il y a également des compétences émotionnelles, de l'empathie par exemple, que l'on ne sait pas mettre dans des modèles. Cela pose la limite de l'usage de l'IA dans le processus RH. Doit-il être limité à des phases spécifiques, par exemple à la présélection dans le cadre d'un recrutement ?

*« Différents types d'agents pourraient apparaître dans la gestion des systèmes apprenants : des agents « interpréteurs » qui aident à comprendre le comportement de la machine à partir de jeux de test, des agents « évaluateurs » ou « vérificateurs » des algorithmes d'apprentissage afin de vérifier que les systèmes apprenants restent loyaux et équitables, des agents « juges » qui vérifient que ces systèmes se comportent conformément à la loi lors de leur utilisation. »*

*« Éthique de la recherche en apprentissage machine », CERNA (juin 2017)*

Par ailleurs, se pose la question de la responsabilité sur le choix des critères, en particulier dans la décision de pondération. Si en tant qu'experts et intermédiaires, les plateformes porteuses de ces solutions ont une responsabilité dans ce développement, cette décision relève d'un choix politique. C'est une orientation politique qui va se traduire dans les outils. Il s'agit de poser un cadre pour corriger certains risques de biais. Ce travail de définition doit donc être effectué par les pouvoirs publics, en collaboration avec les parties prenantes, partenaires sociaux et experts.

Ces technologies soulèvent également la question des données personnelles dans le domaine de l'emploi, et de la frontière en notre identité professionnelle et notre identité personnelle. Cette frontière est aujourd'hui faible. Jusqu'où doit-on, veut-on accepter de donner ses données pour améliorer notre bonheur au travail ? À qui va-t-on les donner, notre (futur) employeur, un tiers de confiance ? Par ailleurs, ce principe est-il réciproque ? Comment faire en sorte que le candidat ait accès aux mêmes types d'information sur l'entreprise ?

*« La prochaine frontière pour conditionner la réussite d'un recrutement, c'est de permettre aux candidats eux-mêmes de disposer d'une fenêtre sur l'entreprise dans laquelle ils pourraient être amenés à travailler. »*

Christophe Montagnon,  
Directeur de l'organisation, des systèmes d'information et de la qualité  
du groupe Randstad en France<sup>31</sup>

### *Conclusion – Pour une approche « human-in-command »*

Le développement en matière d'IA relève ainsi grandement d'un choix politique, de l'ambition et du cadre propice qu'on lui donnera. À cette fin, s'interroger sur les finalités, les usages et les impacts de ces technologies est nécessaire pour leur acceptation et leur encouragement. Mais si nous devons être vigilants en la matière, nous devons l'être également vis-à-vis du cadre que nous lui construisons. Evitons d'enfermer l'innovation dans un cadre réglementaire trop strict. Le caractère fortement évolutif de ces technologies nécessite de les tester, expérimenter en continu, afin d'être en mesure de l'adapter en permanence. L'ère où les technologies d'intelligence artificielle prendront le pas sur l'homme n'est pas encore arrivée, et au regard de l'état de la recherche actuel on ne sait dire si ce sera le cas un jour. Plutôt que de penser en opposition, il convient donc de penser à une nouvelle forme de collaboration, de l'homme avec la technologie. Ceci est particulièrement vrai dans l'organisation du travail. Le métier est avant tout celui de l'homme. Des mécanismes de contrôle et de surveillance humains permanents sont ainsi nécessaires, en

---

<sup>31</sup> Matinée-débat « Le travail à l'ère des algorithmes : quelle éthique pour l'emploi », organisée par Renaissance Numérique et le groupe Randstad en France, le 5 septembre 2017 à Paris.

particulier dans le cadre de la prise de décision. Il convient également de se demander quelles sont les tâches que l'on souhaite déléguer à la technologie, jusqu'à quel degré de décision.

*« Les solutions les plus innovantes et les plus créatives sont et seront encore demain obtenues par une collaboration étroite entre les hommes et les machines, plutôt que par les machines et les algorithmes seuls. »*

François Béharel,  
Président du groupe Randstad en France<sup>32</sup>

Cette vigilance éthique doit être pluridisciplinaire, au regard des dimensions multiples que couvre l'IA (numérique, éthique, économie, droit, philosophie, sociologie du travail, psychologie, etc.).

Nous devons également nous interroger sur la portée de ce cadre. Le caractère transfrontalier de ces technologies nous impose de penser à un cadre qui ne soit pas strictement national. L'Europe doit ainsi jouer un rôle moteur pour porter ses normes éthiques au niveau international. Nous devons tirer bénéfice de l'ambition portée actuellement par le gouvernement sur le numérique au niveau européen, pour favoriser le développement d'une école européenne de l'IA, une approche qui inscrirait ces technologies dans un cadre de débat qui serait le nôtre et respecterait les droits fondamentaux et nos valeurs.

## Synthèse des pistes de recommandations

### Sur la recherche et développement

1. Construire les technologies d'intelligence artificielle de façon éthique (« *ethics by design* »), afin d'identifier dès la conception les problèmes potentiels liés à l'équité et aux biais au sein de ces technologies.
2. Eviter d'enfermer ces innovations dans un cadre réglementaire strict, pour permettre leur expérimentation et de tester leur efficacité et risques. Il s'agit d'encourager l'évaluation de ces technologies, grâce notamment à des benchmarks, afin de savoir comment elles se conduiront : regarder ce que l'on a en entrée et en sortie, les mettre dans des situations « extrêmes » pour construire les frontières de réponses de leurs systèmes.
3. Encourager la recherche sur la transparence et l'explicabilité des technologies d'intelligence artificielle.

---

<sup>32</sup> Matinée-débat « Le travail à l'ère des algorithmes : quelle éthique pour l'emploi », organisée par Renaissance Numérique et le groupe Randstad en France, le 5 septembre 2017 à Paris.

4. Instaurer des mécanismes de contrôle et d'amélioration permanents pilotés par l'humain, tout particulièrement lorsque le système apprend par lui-même.
5. Encourager le développement d'une école européenne de l'IA. Il s'agit de développer une approche européenne de ces technologies pour les inscrire dans un cadre qui respecte les droits fondamentaux et nos valeurs, et pourrait être porté au niveau international.

### **Sur les nouveaux rapports au travail**

6. Faciliter le transfert des technologies d'intelligence artificielle aux acteurs privés et publics (*Open source, Open data*).
7. Identifier les secteurs qui connaîtront les impacts les plus forts, à quel degré et quelle échéance, pour anticiper les conséquences sur l'emploi (contenu du travail, compétences nécessaires, régimes sociaux, etc.). Ceci doit être effectué en collaboration avec les parties prenantes : pouvoirs publics, partenaires sociaux et experts pluridisciplinaires.
8. Adapter la politique de formation à la transformation induite par les technologies d'intelligence artificielle et le numérique dans son ensemble :
  - Démocratiser la formation continue tout au long de la vie, en particulier pour les travailleurs peu qualifiés et/ou dont les expertises ne sont plus requises.
  - Infuser une meilleure connaissance de l'IA, à tous les niveaux de la société et selon des degrés différents, de l'éducation à la formation continue.
  - Augmenter massivement les compétences transversales des travailleurs et futurs travailleurs, qu'elles soient sociales (travail en équipe, intelligence sociale) ou situationnelles (autonomie, adaptabilité).
  - Mettre en place des systèmes de formation plus agiles, afin d'aider les travailleurs à évoluer dans leurs compétences au gré de l'évolution de leurs métiers.
  - Repenser les métriques d'évaluation, les missions et objectifs des travailleurs, au regard de leur collaboration avec les technologies d'intelligence artificielle.
9. Mettre en place des *living labs* par secteur ou entreprise, afin de tester et anticiper les conséquences de la mise en application de ces technologies sur le terrain, en fonction des spécificités des filières et des acteurs.

### **Sur le secteur des ressources humaines**

10. Veiller à la qualité, la pertinence et aux conditions de collecte des données analysées dans les technologies d'intelligence artificielle utilisées dans le processus RH. Ces critères doivent être partagés avec les parties prenantes, soit directement, soit par l'intermédiaire d'un tiers de confiance.



11. Conserver des signaux faibles dans la pondération des critères initiaux des technologies d'intelligence artificielle, afin de se prémunir des risques de discrimination indirecte.
12. Mettre en place de chartes de bonnes pratiques sur l'utilisation des technologies d'intelligence artificielle, par secteur.
13. Instaurer des services de certification pluridisciplinaires (RH, juristes, ingénieurs, représentants syndicaux, etc.), afin de s'assurer de la pertinence des données, de leur caractère non discriminatoire et de leur pondération dans le résultat général des technologies d'intelligence artificielle.
14. Les pouvoirs publics doivent définir le cadre du choix des critères et de leur pondération, en collaboration avec les parties prenantes (partenaires sociaux et experts pluridisciplinaires).
15. Définir les tâches que l'on souhaite déléguer dans le processus RH et à quel degré de décision, et instaurer un contrôle humain permanent de la décision.

## Remerciements

Nous tenons à remercier les intervenants de la matinée-débat « Le travail à l'ère des algorithmes : quelle éthique pour l'emploi », organisée en partenariat le 5 septembre 2017 à Paris, par Renaissance Numérique et le groupe Randstad en France :

- Sarah Bénichou, cheffe de pôle au Département Promotion de l'égalité et de l'accès au droit – Défenseur des Droits
- Christophe Bys, journaliste – Usine Nouvelle
- Jérôme Chemin, secrétaire national – CFDT-Cadres
- Laurence Devillers, chercheuse au Laboratoire d'informatique pour la mécanique et les sciences de l'ingénieur – CNRS, membre de la CERNA et du rapport France IA
- Wafae El Boujemaoui, cheffe du service des affaires sociales, du travail et du sport – CNIL
- Lionel Janin, adjoint au directeur du Département Développement durable et Numérique – France Stratégie
- Laurence Lafont, directrice Marketing & Opérations – Microsoft France
- Bertrand Lamberti, directeur de la stratégie – APEC
- Christophe Montagnon, directeur de l'organisation, des systèmes d'information et de la qualité – Groupe Randstad en France
- Amal Taleb, directrice adjointe des affaires publiques – SAP France