



HAL
open science

Quelle éthique pour l'Intelligence Artificielle dans les procédures de gestion des ressources humaines ?

Thierry Ménissier

► **To cite this version:**

Thierry Ménissier. Quelle éthique pour l'Intelligence Artificielle dans les procédures de gestion des ressources humaines ?. Table ronde "IA & recrutement", ADIRA & Direction Régionale Pôle Emploi Auvergne-Rhône-Alpes, Oct 2021, Lyon, France. halshs-03372648

HAL Id: halshs-03372648

<https://shs.hal.science/halshs-03372648>

Submitted on 11 Oct 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Thierry Ménissier

Thierry.menissier@univ-grenoble-alpes.fr

« Quelle éthique pour l'Intelligence Artificielle dans les procédures de gestion des ressources humaines ? »

[Version écrite de l'intervention à la table ronde « IA & Recrutement », à l'invitation de l'Association pour le Digital en Rhône-Alpes (ADIRA) et Pôle Emploi, direction régionale Auvergne Rhône-Alpes, Vendredi 8/10/2021, 9-12 h, Lyon, Direction régionale Pôle Emploi]

Résumé : Dans cette contribution, il s'agit de préciser la notion d'éthique la mieux adaptée et la plus pertinente pour statuer sur les procédures de recrutement et de gestion des ressources humaines assistées par l'intelligence artificielle. Cette entreprise s'avère complexe car plusieurs questions fondamentales ne sont pas résolues, concernant notamment le rapport de l'humain aux technologies qu'il déploie pour lui venir en aide dans les activités où sa décision est engagée. Une fois ces questions précisées, la contribution propose que les quatre genres de l'éthique aujourd'hui à l'œuvre collaborent de manière soutenue.

Mon positionnement dans cette rencontre obéit à la « double entrée » qui est la mienne en tant qu'enseignant-chercheur à l'Université Grenoble Alpes : d'une part, je suis responsable d'une chaire de recherche intitulée « éthique & IA », portée par l'institut multidisciplinaire de recherche en IA de Grenoble (MIAI) et abritée par mon équipe de recherche l'Institut de Philosophie de Grenoble (IPhIG) ; de l'autre, je suis Professeur à l'Institut universitaire de management, Grenoble-IAE, où je fais partie de l'équipe pédagogique de deux master (« management de l'innovation » et « design », en particulier le parcours « design de transitions ») qui, l'un et l'autre développent des cursus de formation en « mode projets » à partir de problématiques qui concernent les transformations contemporaines des organisations publiques et privées.

D'une part donc, mon point de vue est celui d'un chercheur en philosophie éthique et politique, qui, fatalement, se présente à vous en regard des images associées à cette discipline, ou en regard de ce que l'on attend généralement de l'éthique. Mais je dirais qu'au-delà des images traditionnellement associées à la figure du philosophe et à l'idée d'éthique, ce qu'on doit entendre par ce dernier terme doit être, relativement à notre objet de réflexion d'aujourd'hui, impérativement mieux défini et établi. Quelle éthique pour l'IA dans son usage lié au recrutement ?

Si une telle question mérite attention, c'est que la réponse ne me semble ni évidente ni claire ni simple. En attendant d'élucider plus amplement pour quelles raisons il en est ainsi, je peux préciser un point, lié à « l'horizon d'attente » habituel vis-à-vis de l'éthique : on attend généralement de l'éthique qu'elle indique les règles (et ce point partagé avec nos collègues les enseignants-chercheurs en droit et avec les juristes) et les valeurs qui permettent un usage des technologies et des pratiques conformes à ces règles et à ces valeurs. Sommairement appréhendée, l'éthique peut donc se définir comme la discipline des bonnes pratiques.

Th. Ménissier, « Quelle éthique pour l'IA dans les procédures de gestion des ressources humaines ? », Lyon, 08102021

De l'autre, mon point de vue est celui d'un enseignant dans les études supérieures qui forme des futurs cadres qui eux-mêmes sont, en formation initiale comme en alternance ou en formation continue, en interaction permanente avec le terrain, au sein d'organisations publiques et privées qui vivent actuellement de nombreuses mutations, notamment les transformations digitales et environnementales. Autant de bouleversements complexes et potentiellement vécus dans l'inquiétude, ainsi qu'on le constate aisément quand on consulte à la fois la littérature académique dédiée (en sciences de l'information et de la communication et en sciences de gestion) [voir *infra* références n°2, 6 et 12] ainsi que les témoignages réflexifs émanant des organisations [1, 4 & 5] : manifestement, il se passe quelque chose de particulier pour le monde du travail avec l'apparition et le développement d'outils d'IA, quelque chose qui se retrouve, ainsi que je l'ai montré ailleurs [8], dans toutes les situations où les organisations sont concernées par la fièvre de l'innovation : un mélange émotionnel complexe fait d'expressions d'enthousiasme, de manifestation de signes d'inquiétude et également d'appels à la prudence.

Il y a aura deux temps à mon intervention, le premier est destiné à préciser, sous la forme de trois questions que je vais poser, des points qui me semblent importants pour une discussion sur notre sujet ; le second consiste en la réponse que j'apporterai aux questions.

Premier moment, le questionnement, destiné, ainsi que l'approche philosophique recommande de le faire depuis sa fondation avec Socrate, de construire les problèmes afin d'en apercevoir clairement les éléments constitutifs. Ces questions ne sont pas destinées, à l'heure actuelle, à recevoir de réponses définitives : elles visent littéralement à ouvrir l'esprit aux champs de questionnements qui sont aujourd'hui les nôtres, humains vivant dans des sociétés technologiquement assistées en voie de transition.

D'abord, la question la plus générale : « Quelle philosophie pour l'adoption des technologies innovantes ? ». Par « philosophie », ici, j'entends le sens commun du mot, je l'entends comme « doctrine de pensée et d'action ». Pourquoi poser cette question ? Parce qu'elle suggère qu'il est nécessaire de disposer d'une telle doctrine, et cela parce qu'on doit toujours avoir en tête la non neutralité de la technique (ainsi que l'a par exemple souligné Habermas [7]) : il convient de se prémunir contre une tendance mentale, une habitude souvent prise en France notamment par les ingénieurs issus des grandes écoles et des grands corps. La technologie n'est jamais neutre, elle obéit à des intérêts et elle constitue par elle-même une « vision du monde » (celle de la raison, de la science et du progrès – et souligner cela ne signifie pas du tout se déclarer contre ces valeurs fondatrices !). En somme, poser la question : « quelle philosophie pour l'adoption des technologies innovantes ? » permet de se prémunir contre l'illusion « objectiviste » qui constitue paradoxalement, me semble-t-il, une des premières causes de rejet des innovations en France (comme on le voit aujourd'hui avec les vaccins, le pass sanitaire dans la lutte contre la COVID, ou encore les éoliennes *offshore*, etc.).

De plus, en affirmant que cette question n'est pas passible d'une réponse définitive, je veux également suggérer un autre point, à savoir, le fait que La technologie ne relève pas d'une seule philosophie : il existe aujourd'hui une belle diversité et une grande richesse des approches possibles, ainsi qu'une recherche collective que j'ai récemment encadrée en atteste – la pensée contemporaine se montre ouverte sur des conceptualisations riches et variées à propos de la réception de la technologie [9]. Et d'ailleurs, un pas de plus s'impose à propos des considérations générales qui visent à permettre à la question de pleinement se poser : en fait de

techniques, il convient plutôt d'évoquer des « dispositifs socio-techniques » (ainsi que le suggèrent de manière judicieuse, parmi les sciences sociales, les sciences de l'information et de la communication). Ce qui permet de contrer efficacement un autre préjugé, le préjugé qui, en appréhendant la réalité à travers le concept illusoire d'« objets techniques », manque complètement la réalité sociale des technologies, qui pourtant confère à ces dernières leur signification pleine et entière.

Ensuite, une question générique elle aussi mais liée à notre objet d'aujourd'hui : quelles sont, sur le plan technique, les limites de la transformation des activités & procédures de la GRH par l'IA, le numérique et les données ? Il me semble que la réponse tient probablement en un mot : « *Aucune !* ». Bien entendu, je ne suis nullement prophète et l'histoire du dialogue entre société et innovation technologique est toujours pleine de surprises, mais il me semble que – si l'on met toutefois de côté les incertitudes fortes liées à la soutenabilité (énergétique, environnementale) du développement technologique au sein de l'économie globalisée de l'innovation – il n'y a aucune raison de ne pas accepter l'hypothèse selon laquelle ce développement ne connaît *a priori* aucune limite. Il faut accepter cette hypothèse car elle a le mérite de créer en nous une sorte de vertige, salutaire. La technologie aujourd'hui déployée en matière de recrutement et plus généralement de gestion des ressources humaines est peut-être encore rudimentaire (par exemple en employant des agents conversationnels classiques, ainsi que le montre Daniel Péliissier [12]), mais beaucoup de choses sont possibles, et c'est cela qui est excitant (pour qui aime le dialogue entre la société et la technologie), fascinant même compte tenu des possibles qui peuvent advenir en matière de transformations sociétales, mais également inquiétant, notamment en ce qui concerne la préservation des métiers et le souci du respect des valeurs éthiques comme le respect des personnes.

Enfin, dernière question, apparemment naïve : « À quoi sert l'IA ? ». En fait, comme ce qu'on appelle génériquement l'IA constitue un outil, comme à propos de tout outil dont les humains se servent, la question doit être reformulée et entendue comme ceci : « À quoi fait-on servir l'IA ? ». Une réponse d'ingénieur et de technicien serait qu'on la fait servir aux fonctions qu'on lui a assignées. Mais ce n'est pas ainsi que l'entend un philosophe. A cette seconde question, il pose une autre question impliquant une autre forme de réponse, une question qui concerne les finalités des usages des objets, des systèmes, des dispositifs socio-technique : « À quelles fins veut-on qu'elle serve ? ». Il est permis de suggérer qu'on connaît la réponse à cette question, même si, bien entendu, elle est toujours à poser de manière particularisée, en fonction des usages spécifiques d'emploi de l'IA. La réponse est que l'essor de l'IA sert une automatisation exponentielle des activités humaines pour tous les domaines où s'exerce la décision humaine sous contrainte de responsabilité forte.

Aujourd'hui, ainsi que j'ai récemment eu l'occasion de le développer ailleurs [10], l'expertise algorithmique et la fiabilité des dispositifs technologiques en informatique et dans le numérique contribuent à la mise en place d'une confiance tellement forte dans les algorithmes que ces derniers peuvent acquérir une forme d'autorité. Cette confiance sous-tend en effet les relations opérationnelles avec l'IA, dans le secteur du recrutement comme dans d'autres qu'on doit légitimement considérer comme fondamentaux tant ils sont « critiques » en ce qu'ils requièrent l'engagement d'une faculté supérieure, le pouvoir de juger : la décision médicale, la conduite stratégique d'une opération de combat ou encore les décisions de justice. Au sein du monde qui est en train de se mettre en place, ce monde où l'activité humaine critique est assisté

par l'expertise algorithmique, aucun de ces secteurs n'ignore aujourd'hui la possible mise en confiance procurée par cette dernière.

Cette confiance, malheureusement pour les usagers et les prescripteurs d'IA, n'est ni possible de manière absolue ou intégrale, ni même entièrement stabilisée : la décision humaine se trouve toujours, bien entendu, quelque chose d'imparfait et d'indéfiniment perfectible. Or, il est permis de supposer que ce qui motive le recours à l'IA, c'est pourtant la recherche d'une *forme d'assurance* (par exemple une forme d'assurance concernant la « performance » ou la performativité d'une activité comme le recrutement). Mon hypothèse est que cette tendance masque un autre élément : ce qui motive le recours à l'IA et l'irrésistible désir de « faire confiance » à l'IA en matière de décision humaine, c'est peut-être une chose moins facilement avouable par les prescripteurs aux usagers. Derrière la recherche de l'assurance, c'est un *désir de sécurité ou de sécurisation* par rapport à la prise de décision comme vis-à-vis des effets de la décision. Or, aujourd'hui ce désir de sécurité peut se faire excessif et dangereux pour l'activité humaine, car l'humain est pour sa part faillible, et s'il est compétent, il n'est jamais parfaitement assuré dans sa démarche. C'est pourtant en vue de telles fins, la mise en confiance et la sécurisation de l'activité, que s'exercent les fortes séductions de l'IA.

Second moment dans ma réflexion, à présent : compte tenu de ce contexte complexe, qu'en est-il de l'éthique ?

Je vais avancer une seule idée, qui me semble importante car de nature à clarifier la tâche qui attend aujourd'hui toutes les personnes, toutes différences sociales et professionnelles confondues, qui s'intéressent à l'essor exponentiel dans les sociétés développées des technologies de l'IA, du numérique et de la donnée. A propos de ces nouvelles réalités socio-techniques, on ne peut parler d'une éthique, mais plutôt on doit évoquer *des éthiques* de l'IA, du numérique, de la donnée, au pluriel. Ce sont en effet 4 genres d'éthiques de l'IA qui se trouvent simultanément requises et d'ores et déjà mobilisées. Ces genres d'éthique m'apparaissent irréductibles ou insubstituables les unes aux autres, mais également parfaitement complémentaires. Pourtant, elles ne dialoguent pas encore ensemble. Ce point est particulièrement sensible dans le contexte français, et cela pour des raisons qui, à mon sens, tiennent pour une grande part à la formation des spécialités dans l'enseignement supérieur, laquelle est fortement clivée : d'un côté, les ingénieurs, radicalement coupés des sciences humaines et sociales dès leurs premières années de formation ; de l'autre, les *social scientist* et plus encore les spécialistes des humanités, qui ne disposent pas de la culture technique indispensable pour contribuer aux débats.

Ces quatre genres sont : l'éthique algorithmique, l'éthique artificielle, l'éthique du numérique et de la donnée, enfin l'éthique des usages de l'IA.

La première, l'éthique algorithmique (ou *computer ethics*) est le produit du travail des informaticiennes et informaticiens, gens de l'art qui entendent que les programmes qu'ils créent répondent à des critères opératoires qui sont en même temps des valeurs éthiques, tels que l'explicabilité, la transparence, la « reddition de comptes » (*accountability*), l'« honnêteté » (*fairness*) etc.

La deuxième, l'éthique artificielle, correspond au souci des roboticiennes et roboticiens d'engendrer des automates qu'on peut qualifier d'éthiques. Bien entendu, pour qu'une telle

qualification puisse être complète, il faudrait que la machine soit réellement autonome, qu'elle ait la possibilité d'effectuer en conscience de véritables choix de valeurs, ce qui n'est pas le cas aujourd'hui, et on ne sait pas si cela sera possible un jour. Mais l'éthique artificielle est importante pour les systèmes qui sont en relation directe avec les humains – par exemple, l'éthique artificielle peut concerner la programmation d'un véhicule autonome ou d'un robot domestique ou ménager. L'objectif est minimal : avec le degré d'autonomie limité qui est actuellement celui des êtres artificiels, il s'agit *a minima* d'en concevoir qui soient non nocifs sur le plan éthique.

La troisième, l'éthique du numérique et de la donnée, affronte les questions éthiques variées qui se posent à travers la conception et les usages des systèmes info-communicationnels, tels que les réseaux et les plateformes. Elle implique toutes les actrices et tous les acteurs de la sphère de l'information et de la communication en matière de nouvelles technologies, des concepteurs des systèmes aux promoteurs de l'information et aux spécialistes de sciences humaines et sociales qui étudient les interactions entre les différentes parties prenantes du numérique et de la donnée.

L'éthique des usages de l'IA, enfin, désigne la démarche éthique concernant les usages des dispositifs socio-techniques d'IA, du numérique et de la donnée, et qui les interroge en fonction de leurs finalités politiques ou profondes. C'est sur ce plan que le travail des philosophes est plus particulièrement requis, en réinvestissant de sens les concepts cardinaux de la tradition de la pensée occidentale, tels que l'autorité, la responsabilité, l'autonomie, etc. Mais, comme la dernière mentionnée, elle ne se définit qu'en débat avec des parties prenantes qui ont des intérêts variés dans des jeux complexes, si bien que la sociologie, anthropologie, ou encore la psychologie sociale constituent de précieuses ressources. Il me semble personnellement que ce quatrième genre d'éthique de l'IA ne peut être pleinement rendu fonctionnel que par la mise en débat public des grandes options qui régissent le débat humains-machines. Ce qui est possible dans le cadre d'instances publiques de niveau national ou régional, qui assument leurs responsabilités en permettant, en plus des discussions techniques entre experts, l'émergence d'une authentique « conversation éthique » (ainsi que le dit par exemple la philosophe Martha Nussbaum [11]). Ce qui permet à la réflexion commune d'acquérir l'ampleur qu'elle doit nécessairement prendre sur toutes questions où le devenir technologique risque de modifier sensiblement l'existence humaine.

Bien entendu, de tels débats sont sensibles, et du fait de leur importance ils favorisent l'expression d'émotions confuses, contradictoires, et parfois violentes. Dans l'espoir de s'économiser de telles effusions sociales, il existe une tendance à vouloir faire reposer la dimension éthique sur la seule régulation juridique, ce qui reviendrait à faire endosser aux juristes le costume de seuls éthiciens patentés. Cette tendance est illusoire, les juristes eux-mêmes le reconnaissent, car c'est seulement lorsque la conscience commune se trouve, autant qu'il est possible, clarifiée et apaisée que la loi peut exercer son pouvoir de régulation. Concernant les procédures de recrutement assistées par l'IA, la situation actuelle est précisément celle-ci, le droit a besoin du « secours » de l'éthique [3].

Pour achever ce développement, je voudrais revenir au sujet qui nous réunit ce matin, en l'envisageant dans sa généralité avec cette dernière question : en dehors même de l'apport des technologies d'IA (et même au-delà de toute dimension technique), en quoi consiste

exactement une opération de recrutement réussie, et plus généralement une gestion des ressources humaines, qu'on peut qualifier d'« efficace » ?

Attentif à la manière dont l'humain constitue sa propre existence, le philosophe aurait tendance à répondre en soulignant la dimension *d'institution symbolique* de la procédure de recrutement, compte tenu du caractère extrêmement important et valorisant de l'entrée dans un emploi pour une vie humaine ordinaire. Et de ce fait, il convient également de rappeler *le rôle structurant de l'organisation professionnelle* pour une vie humaine décente : étant donné le poids du travail salarié dans une existence socialisée, recruter puis gérer professionnellement la carrière d'une personne, cela revient à lui accorder un cadre pour une vie normale et heureuse.

C'est en vertu de l'importance de cette procédure d'institution qu'il est important de doter les services de recrutement et de GRH d'une dimension éthique. Et le dialogue des 4 éthiques constitue ici une nécessité, et telle est la tâche de notre temps : dans le contexte de la révolution de l'IA aujourd'hui en cours, les nouveaux outils qui s'offrent aujourd'hui aux services de gestion des ressources humaines doivent et peuvent être accompagnés par un travail de réflexion soutenu, complexe et nuancé. C'est pourquoi je milite pour des comités d'éthique pluriels, ancrés dans l'analyse des situations particulières de chaque type d'organisation, certes, mais accompagnée de débats publics sur les options sociétales qui se verront promues avec l'adoption de tel ou tel système d'intelligence artificielle.

Je vous remercie de votre attention.

Références :

1. APEC, « Etat des lieux et prospective : l'IA dans le recrutement » : <https://corporate.apec.fr/home/nos-etudes/toutes-nos-etudes/etat-des-lieux-et-prospective--1.html> , mis en ligne le 26/10/2020.
2. Françoise Chevalier, Cécile Dejoux, « Intelligence artificielle et Management des ressources humaines : pratiques d'entreprises », *Enjeux numériques*, n°15, septembre 2021, Annales des Mines, p. 95-105 : <http://anales.org/enjeux-numeriques/2021/en-15-09-21.pdf#page=95>
3. Isabelle Desbarats, « Le recrutement à l'ère de l'IA : l'éthique au secours du droit ? », *Revue Lamy Droit des affaires* (n°153). p. 37-49.
4. « Intelligence artificielle et recrutement : quelles applications concrètes ? », *Blog Digital Recruiters* <https://www.digitalrecruiters.com/blog/intelligence-artificielle-recrutement-applications-concretes>, mis en ligne le 15/03/2021.
5. Julie Krassovsky, « Recrutement : quelles sont les limites de l'intelligence artificielle ? », *Revue Capital*, <https://www.capital.fr/votre-carriere/recrutement-queelles-sont-les-limites-de-lintelligence-artificielle-1368783>, mis en ligne le 29/04/2020.
6. Alain Lacroux, Christelle Martin-Lacroux, « L'intelligence artificielle au service de la lutte contre les discriminations dans le recrutement : nouvelles promesses et nouveaux risques », *Management & Avenir*, 2021/2, n°122, pages 121-142.
7. Jürgen Habermas, *La technique et la science comme « idéologie »* (1968), trad. J.-R. Ladmiral, Paris, Gallimard, 1990.
8. Thierry Ménissier, *Innovations. Une enquête philosophique*, Paris, Hermann, 2021.
9. Thierry Ménissier, dir., « Ethique, politique, philosophie des techniques », *Pistes. Revue de philosophie contemporaine*, n°1-2021, Librairie Philosophique J. Vrin.

10. Thierry Ménissier, « Confiance en l'intelligence artificielle et autorité des machines », *Storia e politica*, tome XIII-2, 2021, p. 264-287, accessible à l'URL : <https://hal.archives-ouvertes.fr/halshs-03341931/>
11. Martha Nussbaum, *L'Art d'être juste. L'imagination littéraire et la vie publique* (1995), trad. S. Chavel, Flammarion, 2015.
12. Daniel Péliissier, « La coconstruction ambiguë de l'intelligence artificielle, analyse de la conception de l'intervention d'ouverture des chatbots de recrutement », *Communication & management*, 2020/2, Vol. 17, pages 67-82.