



HAL
open science

Intelligence artificielle et recrutement : typologie, controverses et pratiques responsables

Clothilde Legros, Christine Balagué

► **To cite this version:**

Clothilde Legros, Christine Balagué. Intelligence artificielle et recrutement : typologie, controverses et pratiques responsables. Good In Tech Rapport, Equipe de recherche Good in Tech; Institut Mines Télécom Business School; Sciences Po; Fondation du risque, Institut Louis Bachelier. 2023, pp.1-77. hal-04705786

HAL Id: hal-04705786

<https://hal.science/hal-04705786v1>

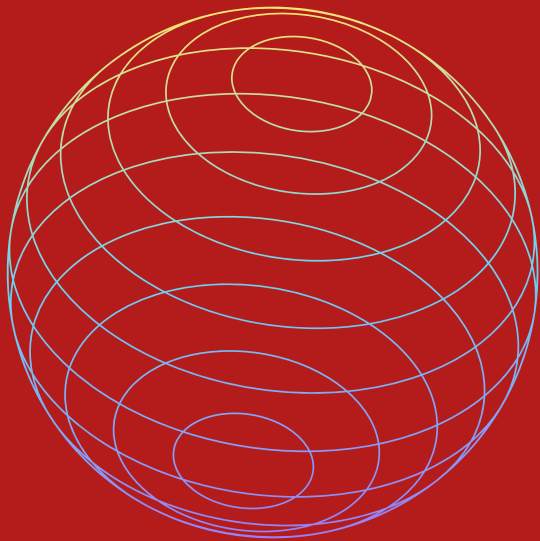
Submitted on 23 Sep 2024

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.

Good In Tech Rapport

Repenser l'innovation et la technologie comme le
moteur d'un monde meilleur pour et par les humains



Intelligence artificielle et recrutement : Typologie, controverses et pratiques responsables

**Clothilde Legros, sous la coordination de
Christine Balagué**

Juillet 2023

PARTIE 1 - TYPOLOGIE DES OUTILS DE RECRUTEMENT BASÉS SUR L'INTELLIGENCE ARTIFICIELLE SELON LES PHASES DU PROCESSUS DE SÉLECTION

I. Première phase du recrutement: Le sourcing, la diffusion de l'offre et la communication avec les candidats potentiels.

1. Le sourcing

- 1.1. Constitution d'une base de données par l'IA
- 1.2. Mise à disposition des recruteurs des bases de données de candidats

2. La rédaction des offres de recrutement

- 2.1. Description des outils
- 2.2. Fonctionnement

3. Les chatbot de recrutement

- 3.1. Description des outils
- 3.2. Fonctionnement

II. Seconde phase du recrutement: L'évaluation des candidats ou screening.

1. Le screening du CV

2. Le screening de l'activité des candidats sur les réseaux sociaux

- 2.1. Objectifs du screening et méthodes
- 2.2. Limites du screening des réseaux sociaux

3. Les tests préalables à l'entretien

- 3.1. Les tests cognitifs
- 3.2. Les tests culturels

III. Troisième phase du recrutement: L'entretien vidéo

1. Description des outils et avantages attendus

2. Fonctionnement

IV. Dernière phase du recrutement : La notation finale des candidats et la prise de décision.

1. La classification des candidats par traitement de données

2. La prise de décision par l'employeur

PARTIE 2. LES CONTROVERSE DES OUTILS DE RECRUTEMENT BASÉS SUR L'INTELLIGENCE ARTIFICIELLE

I. Les tensions entre fluidification du processus de recrutement et risques d'opacification de ses modalités.

1. L'impact de la digitalisation des ressources humaines sur le ressenti des candidats

2. La tension entre efficacité des outils de recrutement et transparence du processus

2.1. L'opacité concernant la présence d'IA

2.2. Les difficultés d'explicabilité posés par les outils basés sur l'IA

II. La tension entre la précision des données collectées durant le processus et les limites légales à la collecte et au traitement

III. La tension entre optimisation du processus de recrutement et risque d'endogénéité

IV. La tension entre l'efficacité prédictive de l'algorithme et les risques de discrimination induits par l'utilisation de l'IA dans le processus de recrutement

1. L'impossibilité de contrôler les discriminations reproduites ou créées par l'algorithme.

2. **Le risque d'accentuation et de créations de discriminations par l'algorithme**

3. **L'accumulation des outils de recrutement : source de multiplication des biais**

V. **Les questions juridiques inédites posées par les algorithmes de recrutement**

1. **Le glissement de la responsabilité : la déresponsabilisation de l'humain par la machine ?**

2. **La complexification de la charge de la preuve : l'exemple français du droit de la non discrimination**

3. **La confrontation complexe des cadres juridiques applicables aux outils de recrutement basés sur l'IA.**

4. **Etude de cas d'espèce**

PARTIE 3 - RECOMMANDATIONS POUR UN USAGE RESPONSABLE DES OUTILS DE RECRUTEMENT ET UNE IA DE CONFIANCE

I. **Prévenir l'introduction de biais algorithmiques dès la conception des outils de recrutement.**

1. **La diversification et la sensibilisation des concepteurs des algorithmes.**

2.1. La création d'une base de données plus représentative

2.2. Processus envisagés pour élargir les bases de données des algorithmes de recrutement

2. **L'intégration d'une plus grande variété de données d'entraînement**

II. **L'instauration d'une vigilance constante dans l'usage des algorithmes de recrutement.**

1. **La sensibilisation du personnel de recrutement**

2. **La sensibilisation et l'information des cibles des algorithmes de recrutement**

3. Le renforcement des obligations de transparence

4. L'audit des algorithmes

4.1. L'audit par les concepteurs/ utilisateurs des algorithmes de recrutement.

4.2. L'audit public des algorithmes de recrutement

4.3. L'élaboration d'une labellisation des algorithmes

I. La spécification des règles juridiques encadrant les algorithmes de recrutement

1. La nécessité de réguler les algorithmes de recrutement

1.1. La mise en place d'un cadre de régulation clair

1.2. Régulation "molle" : La mise en place de lignes directrices "éthiques"

2. Sur l'élaboration d'une chaîne de responsabilité claire en présence d'un algorithme de recrutement.

II. Sur le besoin de recherche spécifique en matière d'algorithmes de recrutement

1. La nécessité d'une recherche transversale

2. La constitution de bases de données

3. La nécessité d'investir dans la recherche à l'échelle européenne pour une IA éthique

Introduction:

En 2019, une enquête menée par l'ONG Algorithmwatch démontrait que les algorithmes de Facebook diffusaient de manière discriminatoire des offres d'emploi en fonction du

genre des utilisateurs. Se faisant passer pour des recruteurs, l'équipe de l'ONG avait acheté des encarts publicitaires simultanément en Suisse, en Allemagne, en Pologne, en France et en Espagne pour poster des offres d'emploi dans différents secteurs (chauffeur.euse poids-lourd, éducateur.trice de jeunes enfants, infirmier.e). Les résultats de ce travail démontraient clairement un biais de genre dans la diffusion des offres : à titre d'illustration, en Allemagne, l'annonce pour les chauffeurs de poids-lourds était diffusée 4864 hommes contre seulement 386 femmes¹.

Si les outils de recrutement basés sur l'intelligence artificielle font encore rarement l'objet de contentieux, un tel cas d'espèce témoigne des difficultés qu'ils peuvent représenter dans un contexte où ils prennent une importance grandissante dans les processus de sélection des candidats.

La notion d'intelligence artificielle est extrêmement complexe à définir. Comme le constate Cédric Villani, en introduction de son rapport au Gouvernement, il serait même "illusoire de chercher une définition claire"².

Il n'existe pas de définition universellement admise de l'IA. Cependant, l'une des plus validées dans le monde de la recherche est celle du chercheur et cofondateur du Laboratoire d'intelligence artificielle du MIT, Marvin Minski, qui la définit comme "*la science des machines qui accomplissent des tâches qui requièrent normalement de l'intelligence humaine* ». Il explique ainsi que l'IA « *implique le développement de programmes informatiques pour effectuer des tâches telles que la reconnaissance de la parole, la prise de décision, la planification et la compréhension du langage naturel* » (Minsky 1986).

La notion d'algorithme, définie par l'informaticien Donald Knuth comme une suite finie et non ambiguë d'opérations ou d'instructions permettant de résoudre un problème où une tâche donnée, éclaire en outre, celle d'intelligence artificielle³.

Souvent classée dans le groupe des sciences cognitives, l'IA mêle en effet plusieurs disciplines dont les mathématiques, l'informatique... Elle recherche des méthodes de

¹<https://algorithmwatch.org/en/automated-discrimination-facebook-google/#:~:text=In%202019%2C%20they%20showed%20that,the%20lifetime%20of%20an%20ad.>

² C. Villani, « Donner un sens à l'intelligence artificielle, pour une stratégie nationale et européenne », Rapport au Premier ministre, mars 2018.

³ D. Knuth, Algorithmes, CSLI, 2011

résolution de problèmes à forte complexité algorithmique et désigne de fait des dispositifs imitant ou remplaçant l'Homme dans certaines mises en œuvre de ses fonctions cognitives (Barabel et al 2020)⁴.

En matière de recrutement, le terme d'IA recouvre ainsi différentes techniques pertinentes qu'il convient de détailler dans un souci d'intelligibilité du propos de ce papier :

- **Apprentissage automatique (Machine learning)⁵** : processus par lequel un algorithme évalue et améliore ses performances à partir d'un modèle de base sans l'intervention d'un programmeur, en répétant son exécution sur des jeux de données jusqu'à obtenir, de manière régulière, des résultats pertinents.
- **Apprentissage profond (Deep learning)⁶**: apprentissage automatique qui utilise un réseau de neurones artificiels composé d'un grand nombre de couches dont chacune correspond à un niveau croissant de complexité dans l'interprétation des données.
- **Neurone artificiel⁷**: Dispositif à plusieurs entrées et une sortie qui simule certaines propriétés du neurone biologique.

Dans la mesure où elle permet d'augmenter significativement les capacités humaines en matière de productivité, d'efficacité et de prise de décision, l'intelligence artificielle occupe une place de plus en plus prépondérante dans notre société et notre économie.

Dans le secteur du recrutement, l'IA a rapidement gagné une influence rapide et significative. Ainsi, en 2022, 22% des entreprises les plus performantes à l'échelle mondiale déclaraient avoir automatisé au moins une partie de leur processus de recrutement à l'aide d'outils basés sur l'intelligence artificielle⁸. Le domaine du recrutement est aujourd'hui considéré comme le domaine RH dans lequel le recours à l'IA est le plus répandu et cela tout au long du processus.

⁴ Barabel et al, L'IA au service des RH, Dunod, 2020

⁵ Barabel et al, L'IA au service des RH, Dunod, 2020, p.36.

⁶ Barabel et al, L'IA au service des RH, Dunod, 2020, p.37.

⁷ Barabel et al, L'IA au service des RH, Dunod, 2020, p.38

⁸<https://www2.deloitte.com/content/dam/Deloitte/us/Documents/human-capital/us-cons-hr-bots-new-super-power-for-workforce.pdf>

L'une des premières raisons de cette mutation digitale du processus de recrutement est l'affluence des candidatures. A titre d'illustration, le Groupe L'Oréal, présent dans 60 pays, comptabilise plus d'un million de candidatures par an⁹. Or dans un processus de recrutement classique, les études montrent qu'environ 30% des CV ne sont jamais lus et peuvent, quand ils le sont, faire l'objet de discriminations conscientes ou inconscientes. L'IA permet ainsi une forte accélération des processus de recrutement et une prise en compte accrue de tous les profils de candidats (Barabel et al, 2020)¹⁰

En outre, l'évolution de la culture des entreprises justifie également l'intérêt de ces dernières pour les outils de recrutement basés sur l'IA. A l'ère du digital, la génération Y, sensible au leadership émotionnel, est en effet demandeuse d'un encadrement de proximité interactif et informel. Ces aptitudes sociales sont donc structurantes de ce qui est recherché par les recruteurs. (Haag et Seguela, 2005). De plus en plus d'organisations dans le cadre de leur processus de sélection des carrières se mettent de ce fait à évaluer les compétences émotionnelles qui sont présentées comme un facteur de performance dans l'exécution d'un métier, dans la capacité à travailler en équipe et dans la mise en œuvre du leadership (Charrière et al, 2014).

Les promoteurs des outils de recrutement basés sur l'IA mettent notamment en avant l'accélération et l'affinement du processus de sélection des candidats. Les professionnels développant ces outils mettent également en lumière la promesse d'un recrutement exempt de discriminations en raison de l'objectivité de l'IA et de sa capacité à éliminer le biais de jugements humains interpersonnels liés par exemple à l'intériorisation de stéréotypes (Lacroux, Martin Lacroux, 2021).

Les algorithmes de recrutement seraient ainsi capables de proposer les profils les plus adaptés au recruteur, et de se détacher des « profils types » vers lesquels les employeurs se tournent généralement, réduisant les discriminations à l'embauche (Amabile et al, 2020)¹¹

⁹<https://www.avature.net/fr/video/60-countries-1-million-applications-per-year-how-loreal-manages-global-recruiting/>.

¹⁰ Barabel et al, L'IA au service des RH, Dunod, 2020, p.225.

¹¹ Amabile et al, Algorithmes: Contrôle des biais SVP, Rapport de l'Institut Montaigne, Mars 2020, P.17.

Cependant, en dépit de ces promesses, les outils de recrutement basés sur l'IA font l'objet de nombreuses controverses et le débat public témoigne d'une méfiance importante à leur égard. Les algorithmes sont principalement décriés pour sacrifier l'éthique au profit d'une plus grande efficacité. Dans la mesure où ils fonctionnent à partir de données collectées au sein de la population, ils sont accusés de reproduire, voire d'aggraver les discriminations qu'ils sont pourtant chargés d'éviter. La critique interroge ainsi leur pertinence dans le processus de recrutement et le prix auquel ces algorithmes fonctionnent en termes de déshumanisation de l'expérience du candidat, d'éthique et de collecte de données.

Le but de ce rapport est de réaliser un état de l'art des pratiques en matière de recrutement basé sur l'IA, de discuter les controverses au regard de l'utilisation faite des outils de recrutement basés sur l'IA et de faire des suggestions allant dans le sens d'une utilisation responsable de ces outils.

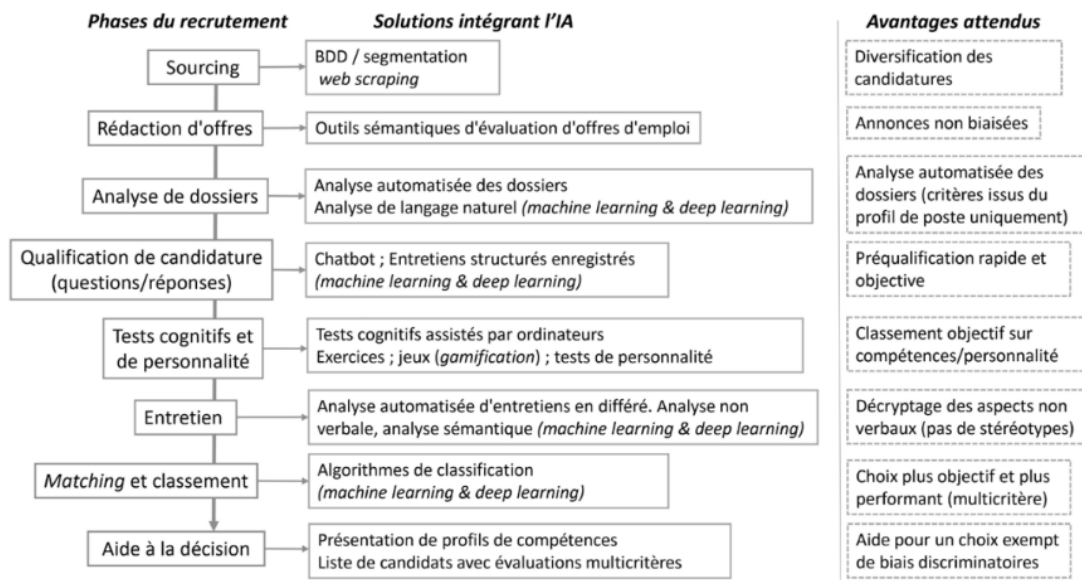
La première partie de ce papier sera ainsi consacrée à une typologie des outils de recrutement basés sur l'IA selon les phases du processus et à l'explication de leur fonctionnement. La seconde partie visera à discuter les controverses et les nouveaux risques que présente le recours à l'IA dans le recrutement. Enfin, la troisième partie de ce rapport détaillera les recommandations pertinentes permettant une utilisation responsable des outils digitaux de recrutement et les mesures à mettre en place afin de tendre vers une IA de confiance.

PARTIE 1 - TYPOLOGIE DES OUTILS DE RECRUTEMENT BASÉS SUR L'INTELLIGENCE ARTIFICIELLE SELON LES PHASES DU PROCESSUS DE SÉLECTION

Les entreprises déploient de plus en plus d'outils de recrutement basés sur l'intelligence artificielle et ce, à toutes les étapes du processus. Pour comprendre la nature de ces outils et l'utilisation spécifique qui en est faite, il convient donc de décomposer et d'étudier une à une les phases du recrutement.

Nous décomposerons cette étude des outils de recrutement en trois parties correspondant à ces différentes étapes. Nous étudierons d'abord la phase visant à constituer une base de données de possibles candidats et à les contacter (le sourcing), la phase visant à échanger et à évaluer les candidats préalablement à leur entretien (le

screening), la phase d'entretien et enfin l'étape de décision finale de l'employeur de sélection du candidat (le matching).



La place de l'IA dans le processus de recrutement (Lacroux, Martin-Lacroux, 2021).

I. Première phase du recrutement: Le sourcing, la diffusion de l'offre et la communication avec les candidats potentiels.

1. Le sourcing

1.1. Constitution d'une base de données par l'IA

L'un des avantages principaux de l'Intelligence artificielle réside dans sa capacité à générer un gain de temps et à automatiser des phases chronophages et répétitives du recrutement. Ainsi, l'assistance d'outils digitaux pour constituer des bases de données de candidats (sourcing) est fondamentale pour de nombreux recruteurs et entreprises de recrutement dans la mesure où il s'agit de la principale source de talents facilement disponibles.

Des outils basés sur l'IA utilisant des algorithmes d'apprentissage automatique sont donc utilisés avant même le lancement de tout processus de recrutement pour scanner et récupérer des informations de candidats accessibles au public, sur les réseaux sociaux et plus largement sur internet. Ces données sont ensuite analysées par l'IA qui utilise des algorithmes de traitement du langage naturel. L'IA trie ensuite les profils qu'elle classe et constitue ces bases de données de candidats (Aora, Kumari, 2021).

1.2. Mise à disposition des recruteurs des bases de données de candidats

Après avoir collecté de nombreuses données de candidats, les principaux sites mondiaux de recrutement et de recherche d'emploi tels que Naukri.com, Monster.com, Timesjobs.com ou encore la version payante de LinkedIn les proposent aux employeurs sur la base d'un modèle d'abonnement aux entreprises. Ces plateformes prennent souvent la forme d'un moteur de recherche permettant au recruteur de trouver les candidats potentiels en utilisant des mots clés.

Ces moteurs de recherche ou CVthèques filtrent les candidats sur la base d'un opérateur logique appelé opérateur booléen utilisé pour effectuer des associations logiques sur des valeurs booléennes (VRAI/FAUX). Les trois opérateurs booléens de base sont : ET, OU, SAUF. Ainsi, par exemple, en utilisant l'opérateur ET, les recruteurs peuvent rechercher des candidats qui répondent à plusieurs critères spécifiques (Marketing ET Logiciel pour un candidat disposant de compétences en marketing et une expérience de travail dans l'industrie du logiciel par exemple). En utilisant des expressions booléennes avancées, les recruteurs peuvent affiner leur recherche pour trouver des candidats qui répondent exactement à leurs critères spécifiques et réduire le temps nécessaire pour trier manuellement une grande quantité de candidatures (Casagrande, Gotti, Lapalme, 2017).

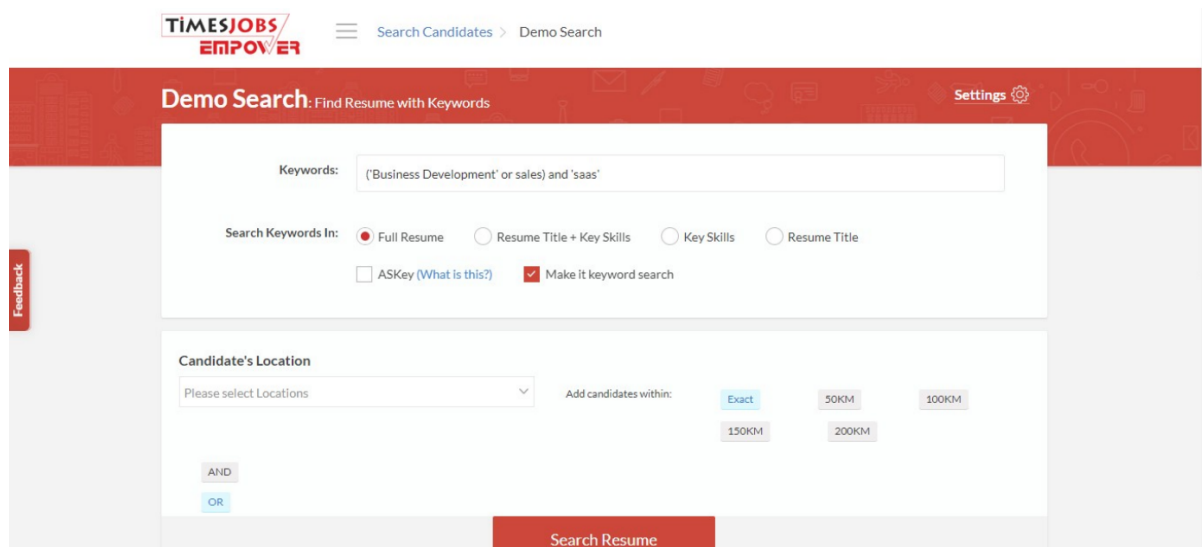


Image du moteur de recherche de la base de données Timesjob (Aora, Kumari, 2021).

2. La génération automatique d'offres de recrutement

2.1 Description des outils

Les algorithmes d'IA peuvent aider à la création de description de poste plus efficace en utilisant l'analyse de données pour identifier les mots clés et les phrases les plus attrayantes pour les candidats. En utilisant ces informations, les recruteurs peuvent créer des descriptions de poste claires et précises qui attirent l'attention des candidats qualifiés (Mihaljević et al., 2022),

Les générateurs automatiques de texte sont des outils qui utilisent l'intelligence artificielle et l'apprentissage automatique pour créer du texte de manière autonome. Ces outils sont de plus en plus utilisés dans le domaine du recrutement pour aider les recruteurs à atteindre des candidats potentiels de manière efficace et personnalisée. En analysant les données des candidats comme leur profil LinkedIn ou leur CV, les générateurs automatiques de texte peuvent créer des messages personnalisés qui correspondent à leurs compétences et à leurs expériences (Barabel et al, 2020).

Ces outils sont de plus en plus utilisés par les recruteurs et connaissent un important développement. Parmi les plus utilisés on peut citer:

- *Crystal*: qui utilise l'analyse de personnalité pour aider à créer des messages personnalisés en fonction de la personnalité du destinataire.¹²
- *Textio*: qui utilise l'analyse prédictive pour aider à créer des descriptions de poste et des annonces d'emploi attrayantes qui attirent l'attention des candidats.¹³
- *Jobscan*: qui analyse le profil LinkedIn du candidat et compare les mots clés utilisés dans son profil avec les exigences de l'emploi pour aider le recruteur à créer des messages personnalisés et ciblés.¹⁴
- *Linked Helper*: qui permet de créer des modèles de messages personnalisés pour différents types de candidats et d'automatiser l'envoi de messages.

2.1. Fonctionnement

La génération automatique de lettres de recrutements se base sur des techniques de génération automatique de texte (natural language processing), visant à extraire de l'information pertinente à partir de données textuelles non-structurées. Le Deep learning a fortement fait progresser la génération automatique de texte. L'utilisation des réseaux de neurones implique qu'il faut au préalable transformer le texte en une suite de vecteurs. Les tâches de traitement du langage naturel consistent à transformer un texte (une séquence) en une autre séquence. Pour ce faire, l'architecture la plus claire est l'encodeur/ décodeur. Grâce à la vectorisation du texte, un premier réseau de neurones encode le texte en un vecteur contexte qui est ensuite passé à un second réseau de neurones chargé de décoder ce contexte pour exécuter la tâche attendue transformant si besoin les différents vecteurs générés en mots (Barabel et al, 2020)

Plus explicitement, l'IA identifie les éléments se trouvant à la fois sur le profil du candidat et sur l'offre de recrutement, ce qui constitue la base des possibles candidats à contacter. Elle va ensuite hiérarchiser et organiser les éléments collectés pour obtenir le message personnalisé le plus clair et le plus efficace possible tout en imitant le style humain (Grand Maison, 2017).

¹² <https://www.crystalknows.com/>

¹³ <https://textio.com/blog/textio-for-job-seekers/13034792749>

¹⁴ <https://www.jobscan.co/>

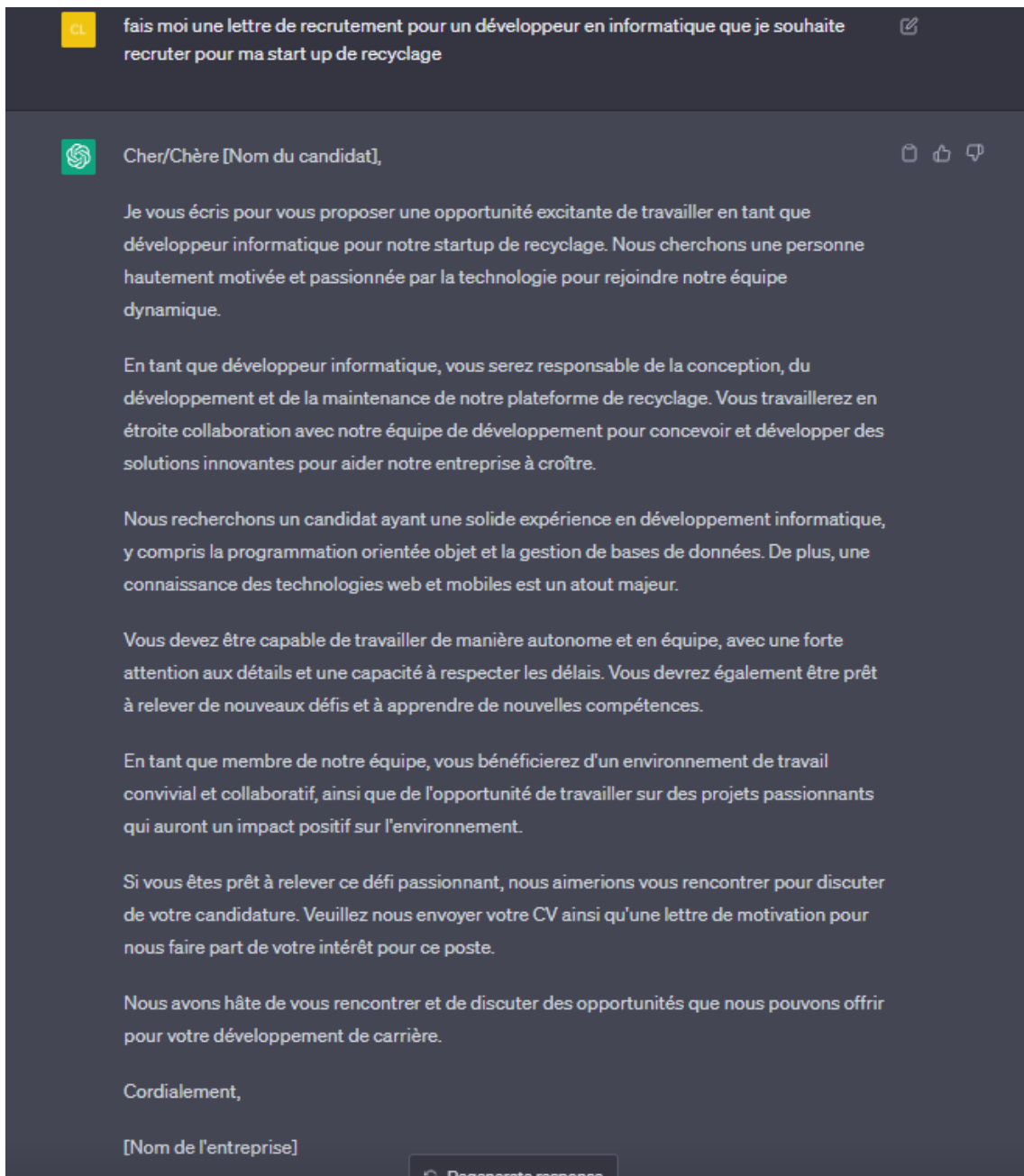
Prenons l'exemple de la formule d'appel utilisée pour contacter le candidat sur LinkedIn. Celle-ci répond à un patron (W/XY?Z), où X est le titre, Y est le prénom, Z est le nom et W est l'expression à utiliser en cas d'anonymat. Si le générateur de texte détecte que le candidat occupe une profession spécifique, le titre X pourra être Docteur s'il est médecin, Maître s'il est avocat, etc. Ce sera Monsieur ou Madame si aucune profession spéciale n'est détectée. Si le niveau de formalité paramétré est bas, le prénom peut également être ajouté. (Grand Maison, 2017).

En outre, l'apparition et la diffusion d'IA génératives (intelligences artificielles capables d'imiter l'humain de manière très rapide¹⁵) constitue une offre encore plus accessible de génération automatique de lettres de recrutement.

A titre d'exemple, en quelques secondes, l'IA ChatGPT¹⁶ prototype d'agent conversationnel utilisant l'intelligence artificielle, développé par OpenAI, peut générer en quelques secondes une lettre de recrutement à partir de critères extrêmement précis:

¹⁵ <https://intelligence-artificielle.com/generative-ai-tout-savoir/>

¹⁶ <https://openai.com/blog/chatgpt>



3. Les Chatbots, un outil digital de communication et d'évaluation

3.1. Définition et avantages attendus de l'outil

Les chatbots ou robots conversationnels sont des logiciels programmés pour simuler une conversation. Grâce à l'IA, ce programme informatique donne l'illusion de tenir une conversation humaine en répondant instantanément aux questions des internautes. Ces robots fonctionnent avec des algorithmes d'intelligence artificielle et font notamment appel à la technologie de Natural Language Processing mentionnée ci-avant. Cette technologie, capable d'analyser les entrées de l'utilisateur, de définir son intention et de

répondre correctement à ses messages, s'avère d'une grande utilité pour les responsables RH (Tawk, 2021), (Swapna et al, 2021). Il existe plusieurs types de chatbots:

- *Les chatbots informationnels*: programmés pour répondre aux questions fréquentes pour poster et obtenir des informations sur l'entreprise: ils correspondent plutôt à la phase de prescreening du processus de recrutement évoquée ci-avant.
- *Les chatbots de matching* : ils s'assurent de la correspondance des offres d'emploi et des candidats et offrent un accompagnement sur les démarches de candidature à suivre.
- *Les chatbot décisionnels*: qui utilisent deux technologies, l'IA et le Machine Learning Process, pour répondre aux questions des candidats et évaluer par la même occasion leur potentiel par rapport au poste à pourvoir. Il met ensuite les meilleurs profils en relation avec le recruteur.

En 2020, 79% des DRH s'accordaient à dire que les chatbots étaient une interface très avantageuse, permettant aux candidats d'obtenir des réponses en temps réel (Barabel et al 2020).

Les professionnels chargés du recrutement avancent de multiples arguments en faveur de ces technologies¹⁷. Ces robots conversationnels amélioreraient ainsi l'expérience du candidat qui obtient des réponses instantanées à ses questions sur le recrutement, diminuant ainsi le phénomène d'opacité après avoir postulé à une offre (Azoulay 2020). De même, les chatbots permettraient d'alléger le travail des recruteurs en leur économisant du temps de travail passé sur les premières phases de tri des candidatures et d'organisation des entretiens. L'exemple de L'Oréal et de Mya, son chatbot de recrutement, utilisé par le groupe jusqu'en 2019 illustre le gain de temps que peut représenter une telle technologie. Mya assiste en effet la société dans le "screening" des candidats aux 5000 postes à pourvoir au sein du groupe. Cet outil représente ainsi un gain de 40 minutes par entretien en moyenne et des économies conséquentes sur le salaire des recruteurs (Tawk, 2021).

¹⁷ <https://blog.hunteed.com/chatbots-recrutement>

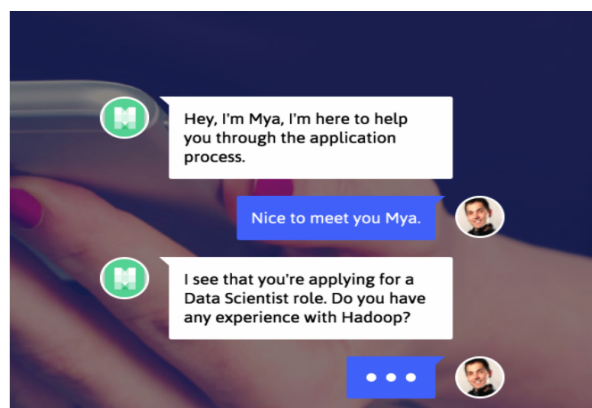
Enfin, ces outils permettent de recueillir plus de candidatures spontanées du fait des conseils donnés par le chatbot mais aussi de contacter les candidats passifs avec un fort potentiel.

Par exemple, le chatbot de XOR, l'un des célèbres chatbots de recrutement, peut prendre en charge de nombreux canaux de recrutement et peut communiquer avec les candidats par le biais de divers modes de communication notamment les SMS, WhatsApp, Facebook, le web, le courrier électronique, etc. Ce chatbot prend en charge la sélection des candidats, la planification des entretiens, l'intégration des employés, les questions fréquemment posées (FAQ) en matière de ressources humaines, et bien d'autres cas d'utilisation (Tawk, 2021).

En 2020 99,3 % des candidats et des employés engagés avec l'IA de XOR se disaient satisfaits de l'expérience. ("The Top 10 Best Recruiting and HR Chatbots - June 2020 | Select software Reviews").

3.2. Fonctionnement du Chatbot

Le chatbot s'active généralement par le clic d'un bouton démarrer. Dans de nombreux cas, il salue l'utilisateur et se présente par l'usage du pronom personnel "je" avant de décliner son prénom (Tibot, Issy, Co'Pilot, Mya), d'expliquer son activité (Pélissier, 2018).



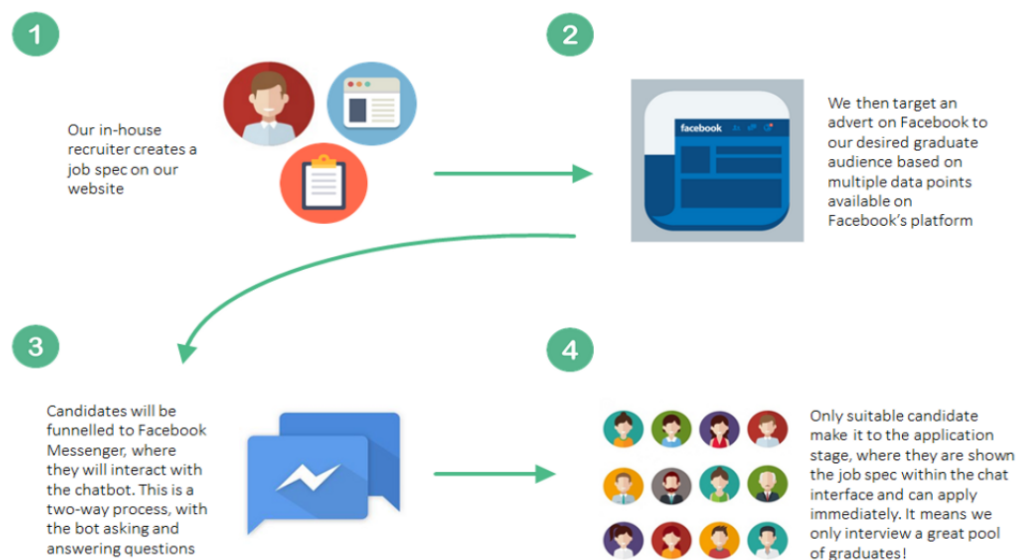
*Exemple de messages envoyés par le chatbot Mya*¹⁸

Généralement, afin d'éviter toute expérience déceptive ou frustrante et de favoriser l'efficacité du chatbot, le robot marquera nettement son caractère automatique et sa

¹⁸ <https://www.revechat.com/blog/chatbot-marketing/>

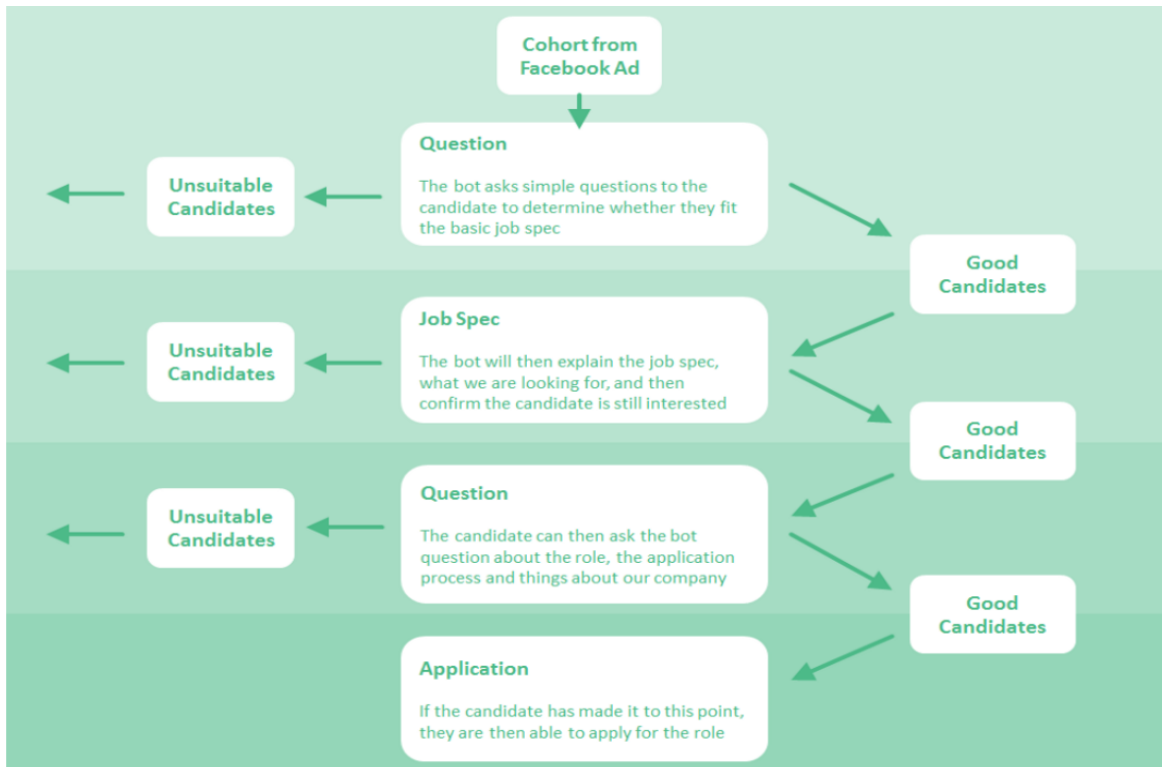
frontière avec l'humain « *Mais je [le chatbot] ne décide de rien ! Et il n'y a aucune mauvaise réponse à mes tests ! C'est toujours un conseiller humain qui étudiera votre candidature, échangera avec vous, et vous proposera des missions* » (chatbot Randstad).

Georges Perry, dans son article *How we changed our recruitment process using a Chatbot* (2017), a schématisé le gain de temps et l'optimisation du processus que représente l'introduction d'un chatbot de matching sur Messenger dans un processus de recrutement.¹⁹



Perry, 2017, How we changed our recruitment process using a Chatbot

¹⁹ <https://medium.com/p/b030b9820c1d>



Perry, 2017, How we changed our recruitment process using a Chatbot

Ces schémas illustrent clairement comment un chatbot peut constituer un premier tri de masse selon des critères extrêmement simples (niveau d'études, spécialisation, date du diplôme, années d'expérience), permettant ainsi un gain de temps aux responsables de recrutement.

Il convient de préciser que certains des chatbots mentionnés dans cette sous-partie correspondent à la fois à la première phase de communication avec les candidats et à la phase de screening que nous allons étudier ci-après.

II. Seconde phase du recrutement: L'évaluation des candidats ou screening.

1. Le screening du CV

Les recruteurs recevant une grande quantité de CV utilisent des outils de recrutement programmés pour effectuer un scan des CV et dégager une liste des candidats les plus prometteurs. Ces outils d'apprentissage automatique sont programmés pour faire des prédictions sur les performances futures d'un candidat sur les bases des termes utilisés dans son CV (Hunkenschroer et Luetge, 2022).

Les outils de screening de CV sont souvent basés sur des algorithmes créés à partir du LASSO, nom donné à une technique d'optimisation visant à trouver une représentation éparsée d'un signal par rapport à un dictionnaire prédéfini (Bonnetoy et al, 2014).

Dans un tel type de screening, il existe quatre types de variables prédictives: les études/éducation du candidat (Université, Master, Doctorat...), ses expériences professionnelles (anciens employeurs, postes occupés...), les compétences reportées par le candidat lui-même sur son CV, les autres mots clés utilisés dans le texte du CV qui sont classés par base linguistique par l'algorithme (nageur, natation, nage seront par exemple classés dans la même catégorie) qui déterminera ensuite combien de répétitions du même groupe lexical sont effectuées par le candidat sur son CV. Même si certaines de ces variables ne sont pas directement liées au poste à pourvoir, elles ont des facultés prédictives sur les performances d'un candidat. Par exemple : des verbes tels que "servir" et "diriger" peuvent indiquer des styles de leadership distincts qui conviennent mieux à certaines entreprises qu'à d'autres (Bo Cowgill, 2020).

Sur la base de ce "screening" initial, ces algorithmes peuvent également fournir une seconde analyse déterminant la meilleure association entre un candidat et un certain emploi lorsque plusieurs postes sont ouverts (Hunkenschroer et Luetge, 2022).

2. Le screening de l'activité du candidat sur les réseaux sociaux

2.1. Objectifs du screening

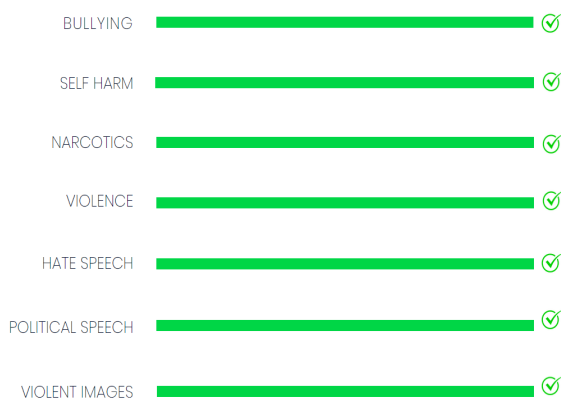
Afin d'évaluer l'adéquation du candidat avec la culture de l'entreprise, de vérifier certaines informations ou de compléter leur évaluation, nombreux sont les recruteurs qui utilisent les réseaux sociaux des candidats (Social media screening). De nombreux travaux ont en effet démontré qu'il existe une corrélation entre la présentation du profil LinkedIn d'un candidat et ses compétences réelles (Ruparel, Dhir, Tendon, Kaurf, Islam, 2020).

L'analyse de l'activité des candidats sur les réseaux sociaux permet également aux recruteurs d'évaluer certaines données telles que la personnalité des candidats, leur intelligence émotionnelle et leurs capacités sociales qui sont des données déterminantes dans la sélection des candidats pour la phase d'entretiens (Ruparel, et al, 2020); (Lacroux, Martin-Lacroux, 2021).

Ce screening vise également pour certains employeurs à vérifier la véracité des informations relatives à certains candidats et à révéler certains éléments dissimulés par ces derniers (Jeske, Shultz, 2016).

A l'appui de cette analyse de l'activité digitale des candidats, des fournisseurs de logiciels ont développé des technologies d'IA qui étudient l'activité numérique d'une personne (pages likées, posts sur les réseaux sociaux...) pour déterminer leur profil psychologique. Ces technologies réduisent ainsi le talent, la personnalité d'un candidat, ainsi que d'autres différences individuelles importantes afin de comparer ces données à la culture de l'entreprise recrutant (Hunkenschroer et Luetge, 2022), (Jeske, Shultz, 2016).

A titre d'illustration, la solution digitale Ferrelty utilise une solution de machine learning pour analyser l'activité des candidats sur les réseaux sociaux et prévenir d'éventuels comportements à risque pour l'entreprise et son image de marque.²⁰



2.2. Limites du screening des réseaux sociaux

Il convient de préciser que le déploiement de telles solutions de screening est plus répandu outre-atlantique dans la mesure où les outils permettant d'analyser les réseaux sociaux de candidats sont souvent difficilement compatibles avec le Règlement Européen sur la Protection des Données Personnelles (RDPD). Le recueil, par un recruteur européen, d'informations publiées par un candidat sur un réseau social ou plus généralement sur internet constitue en effet un traitement de données au regard de ce règlement auquel il est donc soumis.

²⁰ <https://www.ferretly.com/>

L'article 5 du RGPD prévoit en effet une collecte loyale et transparente des données des candidats et impose dans le cadre de toute collecte, une finalité déterminée, explicite et légitime. L'article 9 de ce même règlement interdit le traitement de données sensibles (opinions politiques, éthnie, appartenance syndicale), ce qui limite les possibilités d'un algorithme de screening des réseaux sociaux en Europe.²¹

Au regard de ces articles, les informations collectées par un recruteur doivent présenter un lien direct et nécessaire soit avec l'appréciation de la capacité du candidat à occuper le poste proposé, soit avec l'appréciation de ses aptitudes professionnelles (Art. L.1221-6 Code du travail du travail français). De même, les données considérées comme relatives à leur vie privée au regard de l'article 9 du RGPD ne peuvent en aucun cas être collectées sur les réseaux sociaux.²²

Ainsi, des plateformes de screening des réseaux sociaux conformes au RGPD telles que NEOTAS²³ se développent afin de toucher également des clients européens.

3. Les tests préalables à l'entretien

3.1. Les tests cognitifs

En amont des entretiens vidéo, le candidat peut faire l'objet de tests cognitifs opérés par des IA permettant une première évaluation de ses compétences, qualités et réflexes. Il s'agit généralement de questionnaires à choix multiples, de simulations virtuelles, de jeux vidéo permettant à l'IA de fournir une première appréciation de la motivation, des capacités organisationnelles ou du rapport au risque d'un candidat (Hunkenschroer et Luetge, 2022).

Par exemple, l'éditeur de logiciels Pymetrics utilise les données psychométriques des travailleurs les plus compétents pour établir un modèle permettant de prédire les aptitudes d'un candidat à un poste spécifique. Pour ce faire, les employés en poste jouent à une série de jeux en ligne d'une vingtaine de minutes constituant des évaluations ludiques permettant de mesurer de nombreux traits cognitifs et sociaux.

²¹ <https://www.cnil.fr/fr/reglement-europeen-protection-donnees>

²² https://www.cnil.fr/sites/default/files/atoms/files/guide_-_recrutement.pdf#page=30

²³ <https://www.neotas.com/social-media-screening-gdpr-truth/>

Les données recueillies lors de ces jeux sont ensuite utilisées pour établir un “profil de réussite” pour le poste en question. Les candidats sont à leur tour invités à jouer aux mêmes jeux et le modèle de Machine Learning détermine alors leur probabilité de réussite (Chiang, Berkoff, 2017).

La société française Goshaba s’est développée sur le même modèle. La personnalisation du processus de recrutement, le retour sur l’entretien fourni au candidat ainsi que l’évaluation des “soft skills” sont les avantages mis en avant par les créateurs de cette solution²⁴.

Du fait de la multiplication des tests cognitifs, des programmes d’entraînement des candidats sont même désormais commercialisés pour les préparer aux tests de logique. Il existe par exemple de nombreuses préparations aux tests psychotechniques intégrés dans le processus de recrutement de la société générale²⁵.

3.2. Les tests culturels

Des tests permettant d’évaluer l’adéquation des candidats à la culture de l’entreprise préalablement à tout entretien sont également proposés par certains outils algorithmiques. Par exemple, le groupe L’Oréal utilisait en 2019 la solution Seedlink (Azoulay, 2020).

Cet outil, paramétré et personnalisé selon les critères de L’Oréal permettait d’évaluer en amont certains traits de personnalité valorisés et recherchés dans la culture du groupe (volonté d’innover, capacités stratégiques, intelligence sociale, facultés d’intégration…) en posant des questions ouvertes aux candidats. Il pouvait par exemple être demandé aux candidats de raconter une difficulté rencontrée dans le cadre du travail de laquelle ils se seraient sortis grâce à une solution créative. L’outil de Seedlink était ainsi programmé pour repérer les “compétences” recherchées dans les réponses et donner un score permettant d’apprécier les concordances avec les habiletés recherchées (Azoulay, 2020).

²⁴ <https://goshaba.com/solution/>

²⁵ <https://www.psychotechniqua.com/test-psychotechnique-societe-generale/>; <https://www.decrochez-jo-b.fr/test-societe-generale/>; <https://www.glassdoor.fr/Entretien/Soci%C3%A9t%C3%A9-G%C3%A9n%C3%A9rale-questions-entretien-d-embauche-E10350.htm>

III. Troisième phase du recrutement: L'entretien vidéo

1. Définition et avantages attendus

Traditionnellement menés en face à face ou par téléphone, les entretiens d'embauche sont désormais souvent réalisés par des systèmes de visioconférence en ligne ou par des enregistrements vidéo asynchrones, notamment dans les grandes entreprises.

Le candidat se connecte ainsi à une plateforme web et répond à une séquence de questions prédéfinies par le recruteur en s'enregistrant en vidéo au moyen de sa webcam, son smartphone ou sa tablette. Cette vidéo sera ensuite traitée par l'algorithme qui générera un score d'employabilité du candidat (Hemamou et al, 2020).

Cette phase d'entretien vidéo vise à éviter les biais humains propres à l'entretien, les vidéos filmées par les candidats étant traitées automatiquement par des robots fournissant une analyse logistique du langage et de certains critères non verbaux tels que les expressions faciales, le débit de parole. Parmi les entreprises proposant de telles technologies, on peut citer HireVue ou encore Easyrecrue (Lacroux, Martin-Lacroux, 2021).

Ainsi, les entretiens vidéos permettent aux recruteurs d'accéder à une analyse plus rapide et plus approfondie des candidats dont certaines données n'auraient pas été perçues lors d'un entretien classique. Ces outils sont donc réputés pour optimiser les entretiens et économiser une part des efforts organisationnels et des ressources financières initialement allouées au processus de recrutement (Rupasinghe, Suntharalingam, Sanjeewa Atukorale, 2016).

A titre d'illustration, HireVue est l'un des outils d'entretien vidéo les plus répandus sur le marché américain. Il comptait ainsi parmi ses clients plus de 100 entreprises américaines en 2019, dont Hilton, Unilever ou encore Goldman Sachs pour qui HireVue a enregistré et traité plus de 50.000 vidéos d'entretien cette même année (Harwel, 2019).

Nathan Mondragon, ingénieur et psychologue chez HireVue expliquait, lors d'une interview accordée au Washington Post en 2019²⁶, qu'une vidéo d'entretien standard

²⁶<https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/>

durait 30 minutes, incluait une demie douzaine de questions (paramétrées au préalable par le recruteur) et permettait de collecter et de traiter plus de 500.000 “points de données” pertinents servant chacun à calculer le score du candidat (Harwel, 2019).

2. Fonctionnement

Dans les entretiens vidéo, le comportement non verbal des candidats fait l’objet d’une analyse approfondie de la part d’algorithmes de reconnaissance faciale et vocale. Dans le contexte du recrutement, plusieurs auteurs ont étudié les indices visuels captés chez les candidats qui seraient en lien avec la prédiction de l’anxiété, des performances, de la personnalité ou des tentatives de tromperie. (Hemamou et al 2020).

Selon leur comportement et leurs expressions, ces algorithmes peuvent diviser les candidats en cinq catégories comportementales: coopératif, nerveux, anxieux, extraverti, introverti. Les outils de recrutement vidéo déduisent d’un comportement extraverti des capacités de management et de performance en termes d’interaction sociale, d’intelligence émotionnelle et de communication, ce qui est beaucoup recherché dans les emplois du secteur tertiaire (Hemamou et al 2020).

L’algorithme d’analyse de la vidéo détecte les mouvements du candidat, ses expressions faciales, son timbre de voix, ses intonations et les mots et termes utilisés. Les études des algorithmes d’entretien vidéo ont démontré que les candidats adoptant un comportement non-verbal enthousiaste (regard direct, attitude souriante, hochements de tête) étaient évalués comme plus employables que les autres (Bailly, 2020).

Les outils de recrutement vidéo parviennent à extraire les expressions marquées des candidats (Facial action Units - AU) et les analysent à l’aune d’une méthode de description des mouvements du visage: le Facial Action Coding System (FACS), développée par les psychologues Paul Ekman et Wallace Friesen en 1978. Il s’agit encore de l’outil principal sur lequel se basent les systèmes et les études s’intéressant à la reconnaissance faciale. Sont notamment détectées les expressions universelles telles que la colère, le dégoût, la peur, la joie, la tristesse et la surprise.

Upper Face Action Units					
AU 1	AU 2	AU 4	AU 5	AU 6	AU 7
					
Inner Brow Raiser	Outer Brow Raiser	Brow Lowerer	Upper Lid Raiser	Cheek Raiser	Lid Tightener
*AU 41	*AU 42	*AU 43	AU 44	AU 45	AU 46
					
Lid Droop	Slit	Eyes Closed	Squint	Blink	Wink
Lower Face Action Units					
AU 9	AU 10	AU 11	AU 12	AU 13	AU 14
					
Nose Wrinkler	Upper Lip Raiser	Nasolabial Deepener	Lip Corner Puller	Cheek Puffer	Dimpler
AU 15	AU 16	AU 17	AU 18	AU 20	AU 22
					
Lip Corner Depressor	Lower Lip Depressor	Chin Raiser	Lip Puckerer	Lip Stretcher	Lip Funneler
AU 23	AU 24	*AU 25	*AU 26	*AU 27	AU 28
					
Lip Tightener	Lip Pressor	Lips Part	Jaw Drop	Mouth Stretch	Lip Suck

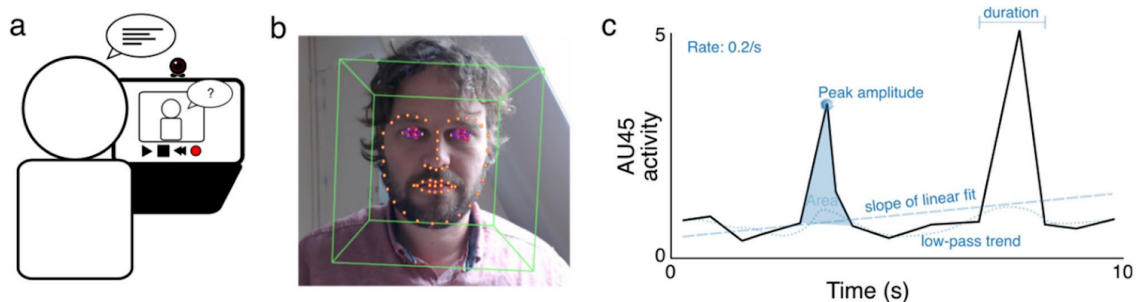
Facial action coding system, (Beh Mei Yin, Muklas Malaysian, Omar, Norain, Talip, Othman 2017).

Ces expressions faciales catégorisées sont ensuite étudiées à l'aide de logiciels de traitement des visages en temps réel développés pour l'analyse des expressions. A titre d'exemple, les chercheurs Rupasinghe, Suntharalingam, Sanjeewa et Atukorale expliquent le fonctionnement du *Computer Expression Recognition Toolbox (CERT)* (Rupasinghe et al, 2016).

Le CERT est un outil basé sur des algorithmes d'amplification (GentleBoost, WaldBoost), permettant de capter les micromouvements du visage en temps réel afin d'en déduire une expression ou une émotion connue du logiciel. Cet outil peut ainsi traiter des vidéos captées par une webcam standard, détecter les 19 unités d'action établies par le FACS et

en déduire les 6 émotions universelles (bonheur, tristesse, surprise, colère, dégoût et peur). Pour ce faire, le logiciel estime notamment l'intensité des sourires posés, l'orientation tridimensionnelle de la tête et les emplacements (x,y) de 10 points caractéristiques du visage (Whitehill et al, 2011).

Le processus d'analyse des expressions faciales de CERT peut être divisé en trois phases principales : la détection des visages, l'extraction des caractéristiques et l'enregistrement des visages (Rupasinghe et al, 2020).



Procédure et extraction des caractéristiques. (a) Enregistrement de la vidéo du candidat lors d'un entretien automatisé. (b) Exemple de marqueurs faciaux détectés (cercles orange) pour les mesures de l'activité de l'unité d'action. (c) Exemple schématique d'un enregistrement de l'unité d'action 45 (clignement des yeux). Les caractéristiques comportementales extraites du signal sont surlignées en bleu. (Kappen, Naber, 2021).

Selon Nathan Mondragon, psychologue industriel travaillant dans l'équipe de développement de l'outil HireVue, les expressions non-verbales (expressions du visage, gestuelle, variations de voix) sont clés dans la compréhension d'un candidat et de l'anticipation de ces réactions face à un consommateur pénible ou à une négociation difficile par exemple. Les unités d'expression faciale (AU) peuvent ainsi constituer jusqu'à 29% du score d'un candidat interviewé avec un outil HireVue, les paramètres audio, dont les intonations, les variations vocales et le contenu de ce qui est dit conditionnant le reste de la note de l'entretien (Harwel, 2019).

Une analyse similaire est également faite par l'algorithme sur la voix du candidat en particulier concernant l'impact de l'amplitude, du ton, des silences ou des disfluences verbales qui permettent de prévoir un comportement stable ou au contraire une anxiété marquée (Hemamou, et al 2020).

La recherche sur les algorithmes d'analyse vidéo, spécifiquement dans le secteur du recrutement, démontre que ce sont des instants de variation des expressions faciales ou vocales qui concentrent majoritairement l'attention de l'algorithme et conditionnent ainsi le score à l'issue de l'entretien qui représente lui-même en moyenne 50% de la note finale d'un candidat à l'issue d'un processus de sélection (Hemamou et al, 2020; Beraich et al, 2022).

IV. Dernières phases du recrutement : La notation finale des candidats et la prise de décision.

1. La classification des candidats par traitement de données

La classification des candidats par l'extraction et le traitement des nombreuses données collectées constitue l'une des phases finales du processus de sélection évoqué ci-avant. Les outils basés sur l'IA permettant de classer les données constituent un moyen efficace d'éviter la subjectivité des responsables RH qui n'interviennent ainsi qu'après une présélection algorithmique des candidats. (Amos Pah, Utama, 2020).

Les outils prédictifs de classement visent à sélectionner le meilleur candidat possible pour une entreprise particulière. Pour ce faire, ils se basent généralement sur une collecte de données des employés en poste (données relatives à la formation, expérience, source de recrutement, moyenne des résultats obtenus lors de la phase de présélection...) afin d'obtenir de nombreuses variables auxquelles se référer (Beraich, et al, 2022). Ces données seront ensuite préparées: une valeur prédictive leur sera attribuée afin de noter les candidats en lice pour un poste (Amos Pah , Utama, 2020).

Outre les données paramétriques initialement rentrées pour permettre de classer les candidats (algorithme d'apprentissage), les outils de matching sont généralement basés sur un réseau de neurones artificiels, capable d'apprendre et de mettre en œuvre le principe de l'induction, c'est-à-dire l'apprentissage par l'expérience. En se confrontant à des situations spécifiques, ils infèrent ainsi un système de décision intégré dont le caractère générique dépend du nombre de cas d'apprentissage rencontrés et de la complexité de la question à résoudre. Les algorithmes de classification des candidats

(C4.5, Support Vector Machine, Naive Bayes...) se réfèrent essentiellement au théorème de Bayes, principal théorème de probabilités²⁷, permettant de déterminer la survenance d'un événement à partir de la réalisation d'un autre événement interdépendant (Amos Pah, Utama, 2020).

Les logiciels de matching utilisés par les recruteurs sont généralement des logiciels tout-en-un proposant au préalable une prise en charge des étapes précédentes du recrutement (screening du CV, entretiens vidéos). Parmi les logiciels les plus populaires, on peut citer Softy qui attribue, à l'issue du matching, des étiquettes aux candidats selon leur pertinence (à rejeter, à contacter, à recruter...).

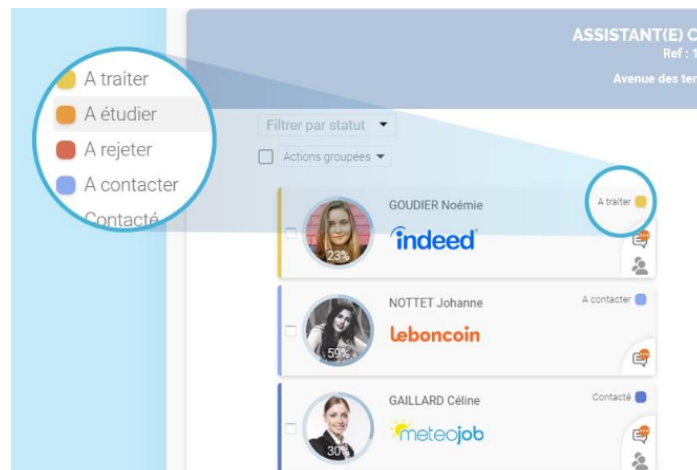


Image du site Softy ²⁸

2. L'étape finale du recrutement : le choix du candidat

²⁷ $P(A/B) P(B) = P(A \cap B) = P(B \cap A) = P(B/A) P(A)$ où

- $P(A)$ et $P(B)$ sont les probabilités à priori de A et B;
- $P(A/B)$ est la probabilité conditionnelle de A sachant B;
- $P(B/A)$ est la probabilité conditionnelle de B sachant A.

²⁸ <https://www.appvizer.fr/ressources-humaines/ats/softy>

Dans la mesure où l'IA renforce la compréhension du présent et permet une certaine prédiction de l'avenir, elle constitue, à l'issue d'un processus de recrutement, une source de fiabilité pour les responsables RH et les managers.

Si les choix multicritères des algorithmes dépassent les capacités humaines, les outils de recrutement prédictifs se présentent toujours comme des outils d'aide à la décision et jamais comme des solutions entièrement autonomes. De nombreux travaux de recherche insistent en effet sur l'incapacité de l'IA de se substituer à la faculté de jugement propre à l'humain (Abradi, et al, 2021).

Lorsque les outils de classement des candidats ont terminé de traiter les multiples données collectées lors du processus de recrutement, le recruteur se voit généralement proposer une liste de candidats présélectionnés avec évaluation du degré d'adéquation par rapport au poste offert. Cette liste peut présenter les candidats selon un score qui leur a été attribué, un niveau de recommandation ou encore un résumé des résultats. (Lacroux, Martin-Lacroux 2021).

Afin de conserver la neutralité et l'absence de biais recherchée dans l'utilisation d'un algorithme, certains professionnels RH préconisent même de demander à plusieurs managers de se prononcer indépendamment sur la décision finale et de délibérer selon la majorité des décisions prises (Kuncel, 2014).

En tout état de cause, la présence d'un humain dans la boucle de décision de recrutement est obligatoire au regard de l'article 22 du RGPD qui dispose du droit à toute personne de ne pas faire l'objet d'une prise de décision entièrement automatisée.²⁹

PARTIE 2. LES CONTROVERSES DES OUTILS DE RECRUTEMENT BASÉS SUR L'INTELLIGENCE ARTIFICIELLE

²⁹ <https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre3>

En dépit de leur efficacité et de leur faculté à traiter un nombre important de candidatures en prenant en compte le plus d'informations possible, les outils de recrutement font l'objet de questionnements techniques et éthiques et leurs avantages sont controversés par certaines critiques.

L'IA induit en effet une transformation et une digitalisation de l'expérience du candidat dont les données sont collectées et traitées massivement ce qui implique de nouveaux enjeux en termes de transparence des outils, de continuité du processus de recrutement et de données personnelles.

De même, les promoteurs des outils basés sur l'IA soulignent l'avantage de ce mode de recrutement dans la mesure où il élimine les biais discriminatoires propres à un recrutement humain puisqu'il ne se base que sur les performances et non sur des caractéristiques telles que le genre ou l'ethnie. Cependant, la recherche et l'expérience de candidats démontrent que ces promesses sont à nuancer dans la mesure où les outils de recrutement basés sur l'IA restent élaborés par l'homme et se basent sur des données passées pour prédire le futur. Ces outils présentent donc un risque de reproduction voire d'accentuation des discriminations dans la mesure où ils sont programmés pour reproduire une version optimisée de la population active en poste.

I. Les tensions entre une optimisation du processus de recrutement et les risques de complexification de ses modalités.

1. L'impact de la digitalisation des ressources humaines sur le ressenti des candidats

Les outils basés sur l'IA dans le recrutement, ont d'après leurs concepteurs et leurs utilisateurs l'avantage de permettre un retour plus rapide au candidat après les premières étapes de recrutement et de les tenir informés. Les Chatbots présentent en effet la capacité de prévenir les candidats que leur candidature a bien été reçue ou qu'elle est en cours d'examen (Hunkenschroer et Luetge, 2022). Cela permet donc de remédier efficacement au "trou noir" vécu par les postulants qui ne sont pas recontactés avant un long moment après avoir déposé leur candidature (Azoulay, 2018).

Au-delà des informations rapidement données aux candidats au cours du processus de recrutement, l'IA permet aux entreprises de fournir un retour détaillé à des centaines de postulants concernant leurs points forts et les aspects de leur profil qui pourraient être plus développés. Ce retour précieux permet ainsi aux candidats de retirer une utilité supplémentaire des processus de recrutement dans lesquels ils s'engagent (Hunkenschroer et Luetge, 2022).

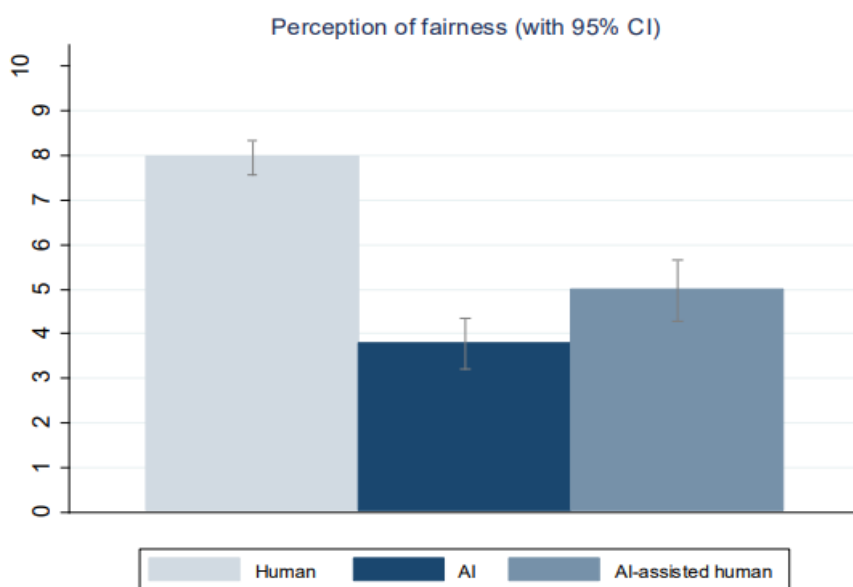
Cependant, l'amélioration de certains aspects de l'expérience du candidat par l'IA est fortement contrebalancée par certains défauts inhérents aux outils digitaux. Le secteur des ressources humaines étant par essence lié au relationnel et aux interactions avec des personnes réelles, la substitution des recruteurs par l'IA présente des limites importantes.

Tout d'abord malgré l'apparence humaine que peuvent prendre certains chatbots dans leur manière de communiquer, ces derniers ne sont capables d'aucune empathie dans leurs échanges avec les candidats. Ces échanges en demi-teinte, formulés d'une manière humaine mais quelque peu aseptisés peuvent donc s'avérer frustrant pour les postulants (Tawl, 2021). Dans une interview menée par le Washington Post en 2019, une douzaine de postulants passés par un entretien vidéo mené à l'aide de l'outil HireVue, ont souligné qu'il était aliénant et déshumanisant de devoir répondre à un ordinateur plutôt que de se voir consacrer du temps par un responsable RH à l'occasion d'un réel entretien. Emma Rasiel, professeur d'économie à l'université Duke, conseillant régulièrement des étudiants à la recherche d'emplois dans le conseil et la finance avait quant à elle déclaré au Post qu'elle constatait l'angoisse croissante des étudiants à l'idée de passer les entretiens digitaux HireVue (Harwell, 2019).

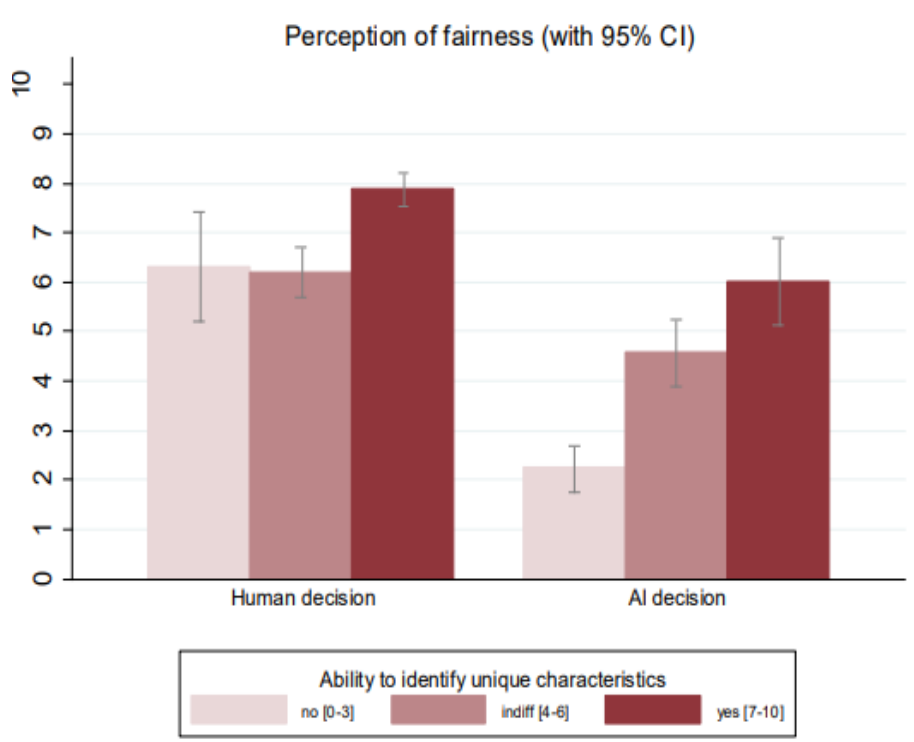
En outre, l'appréciation faite par l'IA des candidats est considérée par certains comme lacunaire. En effet, si certains algorithmes ont une forte capacité d'évaluation des compétences professionnelles "objectives" des candidats, de nombreux doutes sont émis quant à leur capacité d'évaluer leurs compétences émotionnelles. Les postulants déplorent ainsi le manque d'esprit humain dans le recrutement dans la mesure où les

algorithmes passeraient sous silence les spécificités faisant la force et le relief de certains candidats (Martin 2022).

De manière générale, la recherche montre que l'IA dans le recrutement est perçue comme moins équitable de la part des candidats. La littérature nuance cependant ce constat de base en démontrant que l'insatisfaction est plus propre aux candidats n'ayant pas été retenus tandis que les candidats retenus prêtent moins attention aux composantes du processus (Martin, 2022).



Perception de l'équité du processus de recrutement par les postulants, Applicants' Fairness Perceptions of Algorithm-Driven Hiring Procedures Lavanchy & al, 2022.



Perception de l'équité du processus de recrutement par les postulants répondant à la question "Selon vous, le responsable de recrutement/ l'algorithme seraient capables d'identifier vos caractéristiques singulières ?". Applicant Fairness perceptions of AI-Enabled Hiring Procedures, K. Martin, 2022.

Enfin, si l'IA peut se substituer efficacement à l'humain dans certains maillons du processus de recrutement, il convient de souligner que ce dernier reste cependant indispensable, ne serait-ce que dans les dernières phases. En effet, si la recherche démontre que l'IA pourrait techniquement prendre en charge du début à la fin la sélection des candidats (Hunkenschroer et Luetge, 2022). L'article 22 du RGPD relatif à la décision individuelle automatisée, dispose en effet du droit de la personne à ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé. Aux termes de cet article, les outils de recrutement basés sur l'IA doivent rester cantonnés à leur rôle d'aide à la prise de décision et ne peuvent fonctionner indépendamment de l'humain.³⁰

³⁰ https://www.cnil.fr/sites/default/files/atoms/files/guide_-_recrutement.pdf

2. La tension entre efficacité des outils de recrutement et transparence du processus

2.1. L'opacité concernant la présence d'IA

L'introduction d'un élément relativement méconnu du candidat dans la prise de décision, à savoir les algorithmes et l'intelligence artificielle, induit une opacification certaine du processus. D'une part, les candidats ne sont pas toujours clairement informés de la présence d'une IA, de ce qu'implique cette présence dans le processus et le fonctionnement des outils utilisés leur est rarement clairement expliqué.

Cette opacité représente une difficulté éthique dans le contexte de la sélection du personnel en raison de son influence sur la vie professionnelle des candidats et parce que, du fait de leur complexité, ces systèmes sont difficilement contestables. Afin de remédier à cette problématique, le RGPD (art. 22) justifie d'un droit à l'explication en vertu duquel les personnes peuvent demander une justification sur les décisions (algorithmiques) prises à leur sujet (Pena et al., 2020).

2.2. Les difficultés d'explicabilité posés par les outils basés sur l'IA

Comme le souligne le rapport Villani, « *L'explicabilité des systèmes à base d'apprentissage constitue un véritable défi scientifique, qui met en tension notre besoin d'explication et notre souci d'efficacité* » (Villani et al. 2018).

En effet, malgré les contraintes légales (notamment les articles 13 à 15 du RGDP concernant le droit d'explication), certains obstacles techniques s'opposent souvent à une explication claire du fonctionnement des outils algorithmiques aux candidats. Les processus prédictifs et décisionnels des algorithmes étant souvent opaques, pour les programmeurs eux-mêmes. Lorsque les algorithmes utilisent des millions de points de données pour évaluer un candidat, il devient en effet difficile de fournir une explication qualitative des attributs qui motivent les décisions (Hunkenschroer et Luetge, 2022).

Le progrès des outils algorithmiques les a fait évoluer du Machine learning³¹, qui se fonde sur des approches mathématiques et statistiques pour donner aux ordinateurs la capacité d'apprendre à partir de données et de résoudre des tâches, au Deep Learning³², sous catégorie du Machine learning, qui est une méthode d'apprentissage automatique s'inspirant du fonctionnement du système nerveux des êtres vivants. Les algorithmes du Deep Learning traitent l'information d'une façon analogue à ce que feraient des réseaux de neurones humains, répondant aux signaux nerveux qui leur sont destinés. En fonction du type et de la fréquence des messages reçus, certains réseaux de neurones vont se développer quantitativement et qualitativement tandis que d'autres régressent. Cependant le Deep learning nécessite la plupart du temps un grand volume de données et une forte puissance de traitement pour constituer et exploiter un réseau de neurones (Janiesch, Zschech, Heinrich, 2021).

Contrairement aux modèles dits explicites utilisés dans le Machine learning qui conduisent à des décisions explicables et reproductibles, il existe donc un effet "boîte noire" inhérent au deep learning. Il n'est donc pas possible de connaître et de reproduire le mode de décision de l'algorithme de deep learning en raison de la complexité du processus de classification. On parle dans ce cas de modèle de boîte noire car la seule information qu'il est possible de donner est la mesure de l'importance de certains prédicteurs (autrement dit le résultat du calcul) (Lacroux, Martin-Lacroux 2021).

Ainsi, la littérature existante s'accorde à dire que les entreprises et les recruteurs ne devraient pas se fier aux informations produites par un algorithme à boîte noire qu'ils ne comprennent pas entièrement. Il s'agit là d'un défi technique ouvert à relever : construire des algorithmes et des applications d'IA qui conduisent à des résultats explicables (Schumann et al., 2020).

Les entreprises créant et promouvant les outils de recrutement sont, en outre, elles-mêmes réticentes à donner des explications claires sur le fonctionnement de ces IA. Si l'on prend l'exemple de Hirevue, l'entreprise ne donne que de vagues descriptifs des mots ou des comportements offrant les meilleurs résultats (motivé, enthousiaste).

³¹ Cf introduction pour la définition de cette notion.

³² Idem

Hirevue a par ailleurs affirmé que ses experts révisent et perfectionnent régulièrement leur algorithme, mais a refusé un audit extérieur (Harwell, 2019). Il ne s'agit pas de la seule entreprise qui refuse de dévoiler les rouages de son fonctionnement, les développeurs d'outils de recrutement basés sur l'IA étant nombreux à vouloir préserver leur avantage concurrentiel (Lacroux, Martin-Lacroux, 2021).

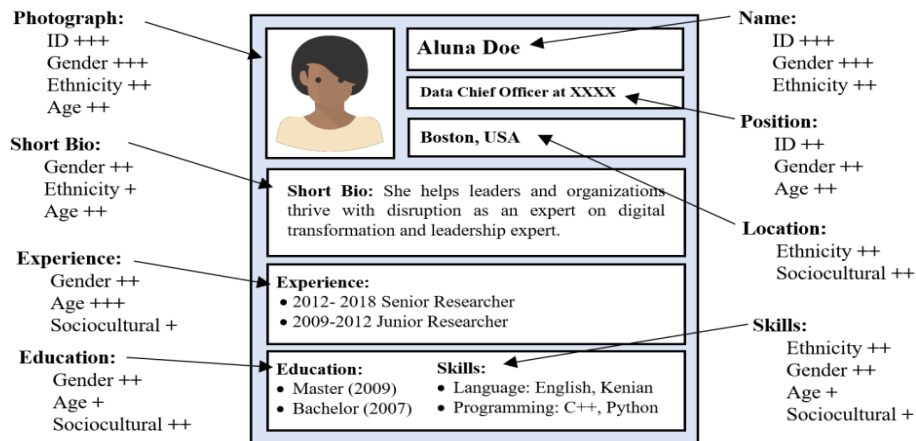
I. La tension entre la précision des données collectées durant le processus et les limites légales à la collecte et au traitement

Les méthodes de recrutement basées sur l'IA se basent très souvent sur le Big Data (données massives), avant tout pour des raisons techniques. En effet, les outils de recrutement basés sur le Machine Learning ou encore le Deep learning, dépendent de la quantité de données collectées dans la base d'apprentissage. Pour garantir la fiabilité des outils de recrutement basés sur l'IA et dans une logique technique portée par les data scientists en entreprise, il y a donc une incitation à collecter et accumuler les données qui conduit presque inmanquablement à des problématiques de respect de la vie privée. De ce fait, il existe sur le terrain des tensions entre data scientists et juristes ou responsables RH autour des solutions fondées sur les big data RH concernant le sujet des données personnelles (Lacroux, Martin-Lacroux 2021).

Le terme de données personnelles est défini à l'article 4 du RGPD comme *“toute information se rapportant à une personne physique identifiée ou identifiable (...) un nom, un numéro d'identification, des données de localisation, un identifiant en ligne, ou à un ou plusieurs éléments spécifiques propres à son identité physique, physiologique, génétique, psychique, économique, culturelle ou sociale”*³³

Le potentiel de récolte de données personnelles lors d'un processus de recrutement réalisé avec l'assistance d'IA serait extrêmement important en l'absence de régulation sur les données personnelles. A titre d'illustration, un simple screening de CV permet de prendre connaissance du genre, de présomptions concernant l'origine du candidat, du nom du candidat, de ses capacités, de son âge ... (Pena et al., 2020).

³³ <https://www.cnil.fr/fr/reglement-europeen-protection-donnees/chapitre1#Article4>



Blocs d'informations disponibles dans un CV et attributs personnels qui peuvent en être déduits. Le nombre de croix représente le degré de sensibilité de l'information (+++ haut; ++ moyen, + faible), (Pena et al. 2020).

Ainsi, toutes ces données ne sont pas collectées notamment au sein de l'UE et les candidats à un processus de recrutement ont en théorie un certain nombre de droits garantis par le RGPD en ce qui concerne le traitement des données à caractère personnel (transparence, identité du responsable de traitement, droit d'accès aux données collectées, droit de s'opposer au traitement de données, consentement libre et éclairé, interdiction d'une décision entièrement automatisée).

Cependant, certains de ces droits sont remis en question par la pratique dans le cadre d'un processus de recrutement par l'IA. Tout d'abord on peut questionner le droit de s'opposer et la faculté de consentir au traitement des données personnelles (art. 13 et 14 RGPD) dans le cadre d'un processus de recrutement. En effet, si les candidats sont effectivement informés de la collecte de leurs données, il est difficile pour eux de s'opposer à leur traitement au risque de se trouver défavorisés (Sánchez-Monedero et al 2020). On peut ainsi questionner la validité du consentement du candidat dont le RGPD exige qu'il soit libre et éclairé (art. 4 et 7)³⁴ étant donnée l'asymétrie du rapport de force entre le recruteur et le postulant (Hunkenschroer et Luetge, 2022).

En outre, si ce processus de screening est limité en Europe du fait du RGPD (interdiction de traiter les données sensibles), il existe un débat sur la pertinence à traiter

³⁴ <https://www.cnil.fr/fr/les-bases-legales/consentement>

automatiquement les données des candidats disponibles sur les réseaux sociaux. S'il s'agit légalement d'un contenu accessible au public, l'éthique d'une telle pratique est à questionner. De même, la recherche interroge la pertinence du screening des réseaux sociaux des candidats dont il est montré que les informations sur les réseaux sociaux n'avaient pas de lien causal direct avec les performances au travail (Hunkenschroer et Luetge, 2022).

Enfin, en ce qui concerne l'interdiction d'une décision purement automatisée, on peut questionner la pleine effectivité de l'article 22 du RGPD donnant droit à un réexamen de la décision par un humain alors que le recrutement peut être basé sur l'IA dès les premières phases. La décision ne peut ainsi jamais être réexaminée dans son intégralité (Sánchez-Monedero et al 2020).

II. La tension entre optimisation du processus de recrutement et risque d'endogénéité

Les promoteurs d'outils de recrutement basés sur l'IA mettent en avant le fait qu'ils permettent un processus de sélection plus précis dans la mesure où les algorithmes d'embauche prédisent le comportement lié au travail et le potentiel de performance d'un candidat en se basant sur les données des personnes actuellement en poste. L'IA pourrait ainsi surpasser en précision les déductions humaines de la personnalité et des aptitudes car elle peut traiter un éventail beaucoup plus large de signaux comportementaux et cognitifs (Chamorro-Premuzic et al. 2016).

Cependant, cette vision est grandement nuancée par de nombreux travaux de recherche qui avertissent sur l'endogénéité des décisions prises à l'aide de l'intelligence artificielle. Pour de nombreux observateurs, les systèmes algorithmiques contribuent en effet à l'émergence d'un nouveau paradigme scientifique en vertu duquel la corrélation l'emporte sur la causalité, sans modèles cohérents et sans théories unifiées (Anderson, 2008). Le recours au Big data et l'analyse en masse de données marqueraient la venue d'une nouvelle forme d'empirisme, nourri par les corrélations inédites que les algorithmes parviennent à faire émerger des masses de données, sans qu'il ne soit plus nécessaire d'en comprendre les causes (Lazaro, 2018).

En effet, la capacité de réflexion et d'apprentissage d'un algorithme de machine learning ou de deep learning est basée sur des données humaines récoltées auprès des travailleurs les plus performants (Abradi, Cambini, Rondi, 2021). Il existe ainsi un problème notable dans la conception et dans la mise en marche de ces outils dont les données d'entraînement sont à "bias in, bias out" c'est à dire qu'elles incorporent de la partialité dans une prétendue impartialité (Lacroux, Martin-Lacroux 2021).

Les données d'entraînement sont des bases de données sur lesquelles les tests de validation et d'ajustement des modèles prédictifs sont menés. Il s'agit principalement des données disponibles sur les candidats ayant postulé dont les caractéristiques sont croisées avec celles des salariés effectivement recrutés. Ces données basant le fonctionnement d'un algorithme de recrutement comportent donc plusieurs biais notables. Tout d'abord, puisque ces données représentent les travailleurs en poste dans une société donnée, elles incorporent inévitablement les biais institutionnels et structurels inhérents à cette société. Ces données se basant sur un échantillon préalablement biaisé, les outils de recrutement se basant dessus produisent inévitablement les biais humains et désavantagent systématiquement les candidats sous-représentés dans les données d'apprentissage. Par exemple, reflétant les actuels chefs d'entreprises ou de service, le leadership ou le comportement assertif seront donc potentiellement par la machine comme des traits masculins (Lacroux, Martin-Lacroux 2021).

Ce fonctionnement basé sur des données empiriques a donc plus pour effet de cloner la population travaillante déjà existante plutôt que de supprimer les biais de recrutement humains (Lacroux, Martin-Lacroux 2021). Les algorithmes supposés émettre des prédictions favorisent majoritairement des situations connues que des évènements rarement ou jamais observés. C'est ce qui s'appelle en statistique le problème de l'antécédent rare. Il s'agit d'un problème commun à tous les modèles de statistiques qui conditionnent sur des évènements rares. Le modèle court le risque de décrire comme impossible ou peu probable une situation qui n'a simplement pas été observée ou observée trop peu de fois (Grand-Maison, 2017).

Cette logique est en outre accentuée lorsque les outils algorithmiques répondent à des critères pré-remplis par le recruteur qui peuvent générer des biais dès la diffusion des offres. A titre d'exemple, la plateforme LinkedIn propose plus de 15 critères de diffusion d'une offre. L'algorithme LinkedIn Talent Match utilise ces critères pour inférer les besoins des employeurs à partir de leurs embauches précédentes ; il y donc de forts risques que les employeurs qui ont une tendance à la discrimination se voient proposer des candidats « adéquats », reflétant leurs choix précédents (Barocas, 2014; Lacroux et Martin-Lacroux , 2021).

III. La tension entre l'efficacité prédictive de l'algorithme et les risques de discrimination induits par l'utilisation de l'IA dans le processus de recrutement

1. L'impossibilité de contrôler les discriminations reproduites ou créées par l'algorithme

L'un des postulats principaux justifiant l'utilisation de l'IA dans le recrutement est basé sur le fait que les algorithmes, au raisonnement purement logique, élimineraient les biais humains aux différentes étapes du recrutement (Hunkenschroer et Luetge, 2022). Il est en effet par exemple prouvé que l'IA peut éliminer les biais de genre dans une annonce d'emploi en créant des descriptions inclusives (Mihaljević et al., 2022) ou que les outils algorithmiques tels que Pymetrics peuvent amenuiser les biais humains en faisant passer au candidat des tests cognitifs uniquement basés sur son intelligence et non sur d'autres caractéristiques discriminatoires ³⁵.

Cependant, le recours massif aux algorithmes d'évaluation et de prise de décisions concernant les candidats présente le risque (comme démontré ci-avant) d'entraîner une plus grande homogénéité des recrutements et donc de fait de reproduire les discriminations existantes. Plus encore, ces outils pourraient même accentuer les biais et la discrimination humaine, voire en créer de nouveaux.

³⁵ <https://www.pymetrics.ai/>

Les algorithmes, essentiellement créés et paramétrés par des hommes de catégories socio-économiques supérieures (Lacroux et Martin-Lacroux, 2021) n'obéissent en effet qu'à un schéma prédéterminé, là où dans les modèles de recrutement classiques plusieurs décideurs humains aux points de vue possiblement différents pouvaient donner leur avis, ce qui implique une perte possible de diversité (Hunkenschroer et Luetge, 2022).

En outre, au-delà des critères apparents recherchés chez un candidat (intelligence, motivation, compétences) les algorithmes, pour effectuer les prédictions les plus pertinentes, vont prendre en compte de nombreuses données qui n'apparaissent pas explicitement mais qui auront un impact majeur dans la notation du candidat : il s'agit des variables latentes qui se réfèrent aux caractéristiques optimales sur lesquelles la décision finale de recruter ou non sera basée (le sexe, les qualifications, les expériences professionnelles passées). Ces variables se réfèrent à un espace observé (les candidats les plus performants en poste dans le cas du recrutement) pour créer l'espace d'hypothèse le plus fiable possible. Dans les utilisations traditionnelles des méthodes de machine learning, les data scientists peuvent vérifier la qualité de l'articulation entre ces deux espaces et analyser la pertinence et l'éthique des données prises en compte par la machine.

Cette surveillance du datascientist des variables utilisées a longtemps été la base de la vigilance sur la qualité et l'équité des modèles algorithmiques. Par exemple, le data scientist peut vérifier si, à niveau de compétence égal, l'algorithme crée des décisions différentes entre les personnes noires et les personnes blanches. Cependant, avec l'apparition du deep learning et la massification des données récoltées et utilisées pour faire les prédictions, cette relation entre l'espace observé et l'espace latent des hypothèses est de plus en plus inaccessible à la vérification et à la critique (John-Matthews, Cardon, Balagué, 2022).

On assiste donc à une opacification du processus de recrutement basé sur l'IA dont on ne peut plus contrôler l'éthique du fait d'une augmentation considérable du nombre de données générée par une granularisation des entités statistiques qui perdent de leur interprétabilité (John-Matthews, Cardon, Balagué, 2022).

2. Le risque d'accentuation et de créations de discriminations par l'algorithme

Cette complexification générant une perte de contrôle sur les décisions de l'IA, supposée améliorer et accélérer le raisonnement de l'humain, certains travaux de recherche avertissent sur le potentiel discriminatoire des outils algorithmiques qui risqueraient de reproduire et d'amplifier les préjugés (Hunkenschroer et Luetge, 2022).

A titre d'exemple, la recherche interroge la propension des algorithmes de traitement du langage naturel à reproduire les stéréotypes de genre qui apparaissent dans les documents écrits, notamment lorsqu'ils s'appuient sur des techniques de "prolongements de mots" qui cherchent à repérer et à créer des associations de mots. L'association de la base d'articles Google news a ainsi permis de mettre en évidence une forte association entre le sexe et les co-occurrences de mots neutres en anglais (qui peuvent être en théorie associées aux deux genres) : le mot « femme » est fortement associé aux mots « foyer, bibliothécaire » ; le mot « homme » est fortement associé aux mots « maestro, skipper, philosophe » (Lacroux et Martin-Lacroux, 2021).

De même, des chercheurs avertissent sur la propension des logiciels de reconnaissance des émotions utilisés lors des entretiens vidéo à accentuer les préjugés raciaux dans la mesure où ils ne fournissent qu'une analyse ethnocentrée des expressions, intonations qui peuvent pourtant avoir différentes significations selon les cultures. Les algorithmes peuvent ainsi systématiquement désavantager certains groupes spécifiques (Fernandez-Martinez et Fernandez, 2020). Plusieurs recherches ont par exemple démontré une mauvaise détection des expressions faciales des candidats au physique atypique (par rapport aux données d'entraînement) : par exemple, le taux d'erreur pour la reconnaissance d'une expression peut varier de 1 % pour un homme blanc à 35 % pour une femme noire (Lacroux et Martin-Lacroux, 2021).

L'analyse des expressions faciales supposées révéler certains traits de personnalité est ainsi grandement controversée au regard des mécanismes qu'elle met en œuvre (ethnocentrisme, stéréotypes raciaux ou de genre). Les travaux récents menés sur la

reconnaissance faciale des émotions fondamentales (colère, angoisse...) viennent renforcer les doutes sur la possibilité d'une reconnaissance automatique en montrant que le décodage de ces émotions est culturellement dépendant : une même mimique faciale est interprétée différemment selon les cultures (Lacroux et Martin-Lacroux, 2021). De même, ces algorithmes s'appliquant de manière uniforme aux personnes sans handicap et aux personnes handicapées, ils peuvent s'avérer pernicieux pour ces dernières. En effet, dans la mesure où ils sont basés sur un référentiel de personnes sans troubles spécifiques, l'algorithme court donc le risque de moins bien classer les personnes atteintes d'un handicap (Packin, 2021).

Concernant les outils d'entretien vidéo tels que Hirevue, Meredith Whittaker, cofondatrice de l'institut AI Now, centre de recherche à New York se montrait radicale, affirmant qu'ils relèvent de la "pseudoscience" et constituent un "permis de discriminer" plutôt que des outils réellement fiables. Selon elle, de tels outils ne peuvent effectivement pas prendre sérieusement en compte les particularités de chacun qui constituent pourtant des éléments qu'un recruteur est supposé examiner. De même, Lisa Feldman Barrett, neuroscientifique spécialiste des émotions et professeur de psychologie à Northeastern University (Boston) exprime également son scepticisme quant à la capacité du système à comprendre ce qu'il regarde. Ses travaux de recherche ont notamment démontré que si les instruments basés sur l'IA pouvaient effectivement détecter les émotions faciales, ils restaient très limités quant à l'interprétation de leur signification et n'étaient pas préparés aux grandes différences culturelles et sociales dans la manière dont les gens manifestaient leurs émotions ou leur personnalité. Utilisant l'exemple du froncement de sourcils qui peuvent impliquer de la concentration ou du stress, mais souvent interprétés par la machine comme de la colère, elle mettait en avant cette faiblesse d'analyse de l'IA pouvant s'avérer dangereuse (Harwell, 2019).

Ces mêmes outils d'entretien vidéo sont également critiqués pour leur faculté à analyser le ton et la voix des candidats, soulevant de nouvelles questions éthiques sur le fait de jauger les compétences des individus sur des caractéristiques physiques inchangeables. De même, la capacité d'analyser la voix et l'apparence pour une IA implique également celle d'en tirer des conclusions sur des données sensibles telles que l'origine du

candidat, son état de santé, son âge, son orientation sexuelle, son attractivité physique (Chamorro-Premuzic 2016).

Cette analyse accrue des caractéristiques personnelles des candidats peut non seulement accroître le risque d'abus et de discrimination intentionnelle (Fernandez-Martinez et Fernandez, 2020) mais également renforcer l'asymétrie d'information et de pouvoir entre les candidats et les employeurs potentiels (Sanchez-Monedero, 2020).

De nombreux travaux de recherche démontrent également que même les outils de recrutement supposés être les plus neutres et ne s'intéresser qu'aux facultés cognitives présentent des risques discriminatoires majeurs.

Des travaux de recherche empiriques ont ainsi mis en évidence la discrimination des groupes initialement sous-représentés par les tests cognitifs traditionnels, tels que les candidats ayant un statut socio-économique inférieur (QCM, problèmes logiques), (Chamorro-Premuzic 2019). De même, les tests et questionnaires à choix multiples génèrent généralement une autocensure des femmes qui ne répondent pas plutôt que de donner une réponse fautive, générant ainsi une surperformance des hommes, plus nombreux à répondre au hasard. Ces tests présentent en effet une distorsion liée à la menace du stéréotype qui représente l'effet psychologique qu'un stéréotype peut avoir sur un sujet qui en est habituellement victime. Face à certaines situations de test, le sujet peut en effet avoir la sensation d'être jugé à travers un stéréotype négatif visant son groupe, ou alors l'avoir intériorisé, ce qui peut provoquer un stress et une diminution des performances sur certains tests cognitifs utilisés en recrutement (Lacroux, Martin-Lacroux 2021).

Enfin, la digitalisation des tests soumis aux candidats implique un désavantage évident des plus vieilles générations face aux jeunes (Lacroux, Martin-Lacroux, 2021).

Au-delà des propensions discriminatoires de ces tests, la recherche interroge en outre leur plus-value dans la mesure où l'introduction de tests digitaux, jeux-vidéos, QCM en ligne, dans le processus de recrutement, peut brouiller la lisibilité de la performance

des candidats et donc nuire à la mesure de leurs compétences (Hunkenschroer et Luetge, 2022).

Cette propension des modèles de recrutement basés sur l'IA à reproduire des discriminations s'est illustrée à travers certains cas d'espèce dont le plus connu est sans doute celui d'Amazon qui a mis au point à partir de 2014 un modèle prédictif intégrant un algorithme de machine learning entraîné sur une base comprenant les recrutements des dix années précédentes. Suite à la mise en place de cet outil, il a été démontré qu'il conduisait au déclassement des candidates féminines (Lacroux et Martin-Lacroux, 2021; John-Matthews, Cardon, Balagué, 2022).

3. L'accumulation des outils de recrutement : source de multiplication des biais

Les risques de biais énoncés ci-avant sont accentués dans la mesure où les processus de recrutement cumulent généralement plusieurs IA. Ainsi, il existe une possibilité d'accumulation des erreurs et des biais non négligeable au cours du processus de recrutement.

Les logiciels de matching utilisés par les recruteurs sont d'ailleurs généralement des logiciels tout-en-un proposant au préalable une prise en charge des étapes précédentes du recrutement (screening du CV, entretiens vidéos). Parmi les logiciels les plus populaires, on peut citer Softy³⁶ qui attribue, à l'issue du matching, des étiquettes aux candidats selon leur pertinence (à rejeter, à contacter, à recruter...).

IV. Les questions juridiques inédites posées par les algorithmes de recrutement

1. Le glissement de la responsabilité : la déresponsabilisation de l'humain par la machine ?

³⁶ <https://recrutement.softy.pro/>

Au-delà des avantages techniques relatifs au gain de temps et éthiques concernant la diversification de l'embauche, les instruments de recrutement basés sur l'IA présentent également un intérêt légal dans la mesure où ils promettent à l'entreprise d'éviter de faire l'objet de poursuites pour pratique de recrutement discriminatoire (Lacroux, Martin-Lacroux, 2021).

Cependant, cet aspect sécurisant pour les entreprises risque de compromettre les possibilités de recours pour un candidat qui serait victime d'une discrimination à l'embauche. L'utilisation d'algorithmes, souvent produits par une entreprise tierce, complique la désignation d'un responsable et donc la formation d'un recours. Se pose en effet la question de l'imputabilité des décisions de recrutement prises étant donné que plusieurs groupes de personnes (les scientifiques des données, les responsables de l'embauche ou l'entreprise dans son ensemble) sont impliqués dans ces décisions (Hunkenschroer et Luetge, 2022). Cette difficulté à établir une chaîne claire de responsabilité est amplifiée par la volonté des fournisseurs de logiciels de protéger leur propriété intellectuelle en ne dévoilant pas le fonctionnement de leurs algorithmes (Sanchez-Monedero et al. 2020).

Les algorithmes peuvent ainsi constituer une vitrine de protection pour les employeurs chargés de respecter certaines lois ou certains quotas en matière de discrimination. Un des exemples les plus connus est celui du respect aux Etats-Unis de l'American with Disabilities Act (ADA), loi votée en 1990, dont le but est de protéger la population des Etats-Unis contre les discriminations basées sur le handicap.³⁷

L'ADA vise à garantir que les handicaps qui n'entravent à aucune fonction essentielle du poste de travail ne soient pas une source d'exclusion injustifiée des candidats. Cependant, les algorithmes de reconnaissance faciale fondés à partir de données d'entraînements de personnes valides peuvent être source de discrimination pour les personnes atteintes d'un handicap (physique ou mental) dont l'apparence, le rapport aux expressions faciale ou vocales est différent³⁸. Dans la mesure où il n'existe pas d'algorithmes spécifiques aux personnes handicapées ou que les données d'entraînement

³⁷[ADA.gov](https://www.ada.gov)

³⁸ <https://www.justice.gov/opa/pr/justice-department-and-eeoc-warn-against-disability-discrimination>

ne prennent pas en compte le handicap, les outils basés sur l'Intelligence artificielle constituent une défense utile pour les employeurs qui peuvent justifier le non-respect des lois anti-discriminatoires par le fonctionnement des algorithmes (Packin, 2021).

Dans ce contexte de déresponsabilisation de l'employeur induit par l'utilisation d'algorithmes d'assistance au recrutement, plusieurs travaux de recherche insistent sur le caractère uniquement complémentaire de ces outils qui sont uniquement censés atténuer les préjugés des responsables de recrutement et non totalement les relayer dans la prise de décision. Les chercheurs alertent ainsi sur le principe de "confiance automatisée", phénomène conduisant les employeurs à se fier entièrement au jugement de l'outil digital. De même, ils soulignent que l'utilisation des outils de recrutement basés sur l'IA doit faire l'objet d'une coordination au sein de l'entreprise et doivent être ancrés dans une gestion fonctionnelle des ressources humaines, sinon quoi, leurs bienfaits seraient dévoyés (Mihaljević et al., 2022).

2. La complexification de la charge de la preuve et la question du droit de la non discrimination

Pour prendre un exemple à l'échelle nationale des répercussions juridiques de l'introduction d'outils basés sur l'IA, il semble pertinent de s'intéresser au droit de la non-discrimination dans l'emploi.

La présence d'algorithmes vient en effet complexifier cette branche récente du droit du travail qui permet en théorie de lutter contre les discriminations volontaires ou systémiques perpétuées par la politique d'emploi d'une entreprise.

Dans la mesure où certaines discriminations peuvent être subtiles (notamment en matière d'embauche), il existe un aménagement de la charge de la preuve en droit de la non-discrimination. Si le demandeur peut identifier des effets discriminatoires qui touchent régulièrement un ou plusieurs groupes (par exemple si les femmes sont systématiquement moins embauchées que les hommes), il y a ainsi présomption de discrimination. L'employeur aura alors à sa charge de prouver qu'il ne perpétue pas une telle discrimination et de renverser la présomption en démontrant l'exécution d'une

obligation de moyen dans l'élaboration de plans d'action visant à l'inclusion des groupes visés (Mercat-Bruns, 2015).

Le droit de la discrimination se focalise donc sur l'intention de l'employeur ou recruteur et tient compte du fait que l'entreprise ne peut pas forcément justifier certains effets discriminatoires de son fonctionnement intrinsèque ou historique (Mercat-Bruns, 2015). On comprend donc comment la présence d'algorithmes dans le processus de recrutement peut compromettre la démonstration d'une discrimination systémique par un groupe de candidats dans la mesure où l'usage d'un outil basé sur l'IA témoigne d'emblée de l'envie d'objectivité de l'entreprise et des recruteurs.

En outre, le manque de transparence des systèmes de recrutement basés sur l'IA et les corrélations de données permises par les algorithmes, souvent de manière invisible, rendent les protections offertes par le droit incertaines voire inefficaces³⁹ (Pena et al., 2020). Ce manque d'intelligibilité du processus de recrutement compromet d'éventuelles contestations des décisions qui sont prises et donc la naissance de tout contentieux (Sánchez-Monedero et al 2020). Il est en effet impossible d'exercer son droit au recours lorsqu'on ignore que l'on a été victime de discrimination, que l'organisation utilisatrice de l'algorithme n'en a elle-même pas conscience et que le concepteur de l'outil refuse d'expliquer en détail son fonctionnement⁴⁰.

En ce sens, le Défenseur des droits et la CNIL, dans leur rapport publié en 2020, *Prévenir l'automatisation des discriminations*, appellent à une mobilisation des pouvoirs publics afin de s'assurer que le droit de la non-discrimination remplisse son office.

Une solution adaptée à l'IA est cependant complexe à trouver dans la mesure où le droit civil et le droit pénal qui régissent la non-discrimination ont été pensés pour des personnes humaines, impliquant souvent une faute et/ou un élément intentionnel et un élément matériel dans la désignation d'un responsable. Si certains pensent qu'un statut juridique particulier devrait être attribué aux IA, de nombreux rapports soulignent le

³⁹ Rapport du défenseur des droits en partenariat avec la CNIL, Algorithmes : prévenir l'automatisation des discriminations, 2020, P.6.

⁴⁰ Idem

risque de déresponsabilisation totale des recruteurs et des concepteurs d'outils⁴¹ (CF Partie III, 2. Infra).

3. Confrontation des cadres juridiques

L'une des plus grandes complexités juridiques relatives aux algorithmes de recrutement réside dans la confrontation de plusieurs cadres juridiques qu'implique leur utilisation.

Si la plupart des systèmes d'embauche automatisés sont développés aux Etats-Unis, ils sont utilisés dans le monde entier. Or si les objectifs d'embauche (profit, compétences, correspondance avec l'entreprise) sont relativement les mêmes à travers les cultures, la réglementation et la régulation des systèmes automatisés et basés sur des données est très propre au système juridique dans lequel elles s'inscrivent (Sánchez-Monedero et al 2020).

Au sein de l'UE, le Règlement Général sur la Protection des Données personnelles fournit notamment une série de droits très protecteurs des ressortissants européens et de leurs données qui n'ont pas d'équivalent aux Etats-Unis. L'article 48 du RGPD prévoit notamment que le transfert de données personnelles par une entreprise à une autorité administrative d'un pays tiers ne peut avoir lieu que s'il est fondé sur un accord international tel qu'un traité d'entraide judiciaire.

Or, aux Etats-Unis, le Clarifying Lawful Overseas Use of Data Act (CLOUD Act) permet aux autorités fédérales étatiques américaines d'exiger d'un fournisseur de services américain qu'il leur communique des données hébergées en dehors des États-Unis. De fait, il contredit l'article 48 du RGPD et nuit aux droits des citoyens européens.

C'est du fait de ces incompatibilités juridiques que le Privacy Shield (accord d'adéquation validant le transfert des données européennes vers les Etats-Unis), avait été invalidé par la Cour de Justice de l'Union Européenne en juillet 2020, obligeant les entreprises traitant les données de ressortissants européens à relocaliser le stockage de leurs données en Europe, ou à mettre en place des clauses contractuelles conformes au

⁴¹ S. MERABET, Vers un droit de l'Intelligence artificielle, Dalloz, Juillet 2020, P.489

droit européen⁴². Cependant, ces mesures de transition n'étant pas complètement fonctionnelles, certaines questions restent en suspens. On peut par exemple se demander quelles sont les garanties en matière de traitement des données d'un candidat européen postulant aux États-Unis via des outils basés sur l'IA.

De même, en matière d'égalité et de norme anti discrimination, les pays, qu'il s'agisse du Royaume-Uni, des États-Unis ou des pays de l'UE, n'ont pas les mêmes standards, les mêmes variables protégées et les mêmes normes (Sanchez-Montero et al, 2020).

Au sein de l'UE, la Charte des droits fondamentaux de l'UE et plusieurs directives telles que la directive sur l'égalité entre les hommes et les femmes de 2004⁴³ ou encore la directive-cadre relative à l'égalité de traitement en matière d'emploi et de travail de 2000⁴⁴ régissent le principe d'égalité et posent les variables protégées. Au Royaume-Uni et aux États-Unis, ce sont respectivement l'Equality Act de 2010 et le Titre VII du Civil Rights Act de 1964 qui posent de telles normes.

Ainsi, il est légitime de se questionner sur la force normative de ces textes quand les principaux outils de recrutement basés sur l'IA disponibles émanent de concepteurs américains, risquant ainsi de gommer les spécificités et les exigences nationales (Sanchez-Montero et al, 2020).

⁴² <https://www.goodintech.org/preview/blog-article.html?blogID=338>

⁴³ <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:32004L0113>

⁴⁴ <https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX%3A32000L0078>



Figure 1: Schématisation des controverses afférentes à l'utilisation d'outils basés sur l'Intelligence artificielle dans le recrutement

PARTIE 3 - RECOMMANDATIONS POUR UN USAGE RESPONSABLE DES OUTILS DE RECRUTEMENT ET UNE IA DE CONFIANCE

A l'ère de la généralisation de l'Intelligence artificielle dans le monde du recrutement et plus généralement dans la société, les travaux de recherche, les rapports d'organisations gouvernementales ou non gouvernementales visant à élaborer des recommandations et des suggestions de bonnes pratiques se sont multipliés.

La dernière partie de ce rapport vise à synthétiser les recommandations les plus pertinentes afin de tendre à un usage responsable des algorithmes de recrutement et à conserver une certaine éthique et équité dans la sélection des candidats, même si une partie du processus est automatisée.

I. Prévenir l'introduction de biais algorithmiques dès la conception des outils de recrutement.

Avant même de s'intéresser au contenu des algorithmes de recrutement et aux biais qu'ils peuvent présenter a posteriori, il convient de s'intéresser à leurs concepteurs et aux informations qu'ils mobilisent dans la fabrication de ces outils fonctionnant à partir de bases de données.

1. La diversification et la sensibilisation des concepteurs des algorithmes.

L'apparition des biais des algorithmes a pour première racine l'identité de leurs concepteurs. La recherche met en effet en évidence l'influence des caractéristiques sociodémographiques des programmeurs sur le fonctionnement des algorithmes, leur milieu étant extrêmement homogène. (Lacroux, Martin-Lacroux, 2021).

Or, comme le rappelle un rapport de l'Institut Montaigne, publié en mars 2020, les cadres et les techniciens des entreprises centrées sur les données sont hautement qualifiés pour comprendre le fonctionnement et l'utilisation faite des algorithmes qu'ils développent mais peuvent présenter des lacunes dans leur vision de l'équité (Bouverot et al, 2020)⁴⁵.

⁴⁵ Rapport de l'Institut Montaigne, Algorithmes Contrôle des biais SVP, Mars 2020

La recherche démontre en effet que les résultats des algorithmes dépendent de la manière dont les programmeurs les ont “composés”. Une homogénéité dans la population des concepteurs d’algorithmes générera donc inéluctablement des biais cognitifs, c’est-à-dire une distorsion de l’information par rapport à une seule réalité, à savoir la vision du monde des concepteurs (Bertail et al, 2019).⁴⁶

Les rapporteurs soulignent ainsi la nécessité d’insister sur l’importance de mobiliser une diversité de profils tant sociale que professionnelle dans les projets de développement d’algorithmes. L’hétérogénéité des profils des concepteurs représente en effet la garantie d’une compréhension plus globale des enjeux relatifs aux algorithmes. A titre d’exemple, les femmes sont plus susceptibles que les hommes de détecter le potentiel de harcèlement que présente une technologie.

Cependant, la recherche de diversité de genre, d’ethnie ou de milieu social, au sein d’une équipe de développeurs n’est pas chose aisée dans la mesure où les formations d’informatique, de statistiques ou de data science sont extrêmement composées d’hommes de classes sociales élevées. De même, les experts des données sont très peu formés à la réalité sociale des discriminations⁴⁷. Dans un premier temps, pour pallier ce manque de diversité, il pourrait donc être bénéfique de former les experts des données afin de renforcer leur compréhension de ce que sont les biais de société et l’équité réelle. Cela permettrait à ces experts de comprendre l’importance de recueillir un échantillon de données d’apprentissage reflétant une réalité objective et universelle et non seulement celle qui leur est familière (Bouverot et al, 2020).

2. L’intégration d’une plus grande variété de données d’entraînement

2.1. La création d’une base de données plus représentative

L’une des principales sources de biais des algorithmes de recrutement et des algorithmes en général provient de la représentativité extrêmement restreinte des données d’apprentissage. Les algorithmes, apprenant sur des exemples limités, ne peuvent alors pas produire des décisions appropriées pour les populations qui sont exclues de l’échantillon d’apprentissage. À titre d’exemple, les premiers algorithmes de

⁴⁶ Rapport de Télécom Paris Tech & Fondation Abeona, Algorithmes: Biais, discrimination et équité, février 2019.

⁴⁷ Rapport de la CNIL, Algorithmes : prévenir l’automatisation des discriminations, 2020

reconnaissance faciale de Google, uniquement entraînés sur des hommes blancs, classifiaient certaines personnes noires comme des gorilles (Bouverot et al, 2020).

Ainsi, le caractère biaisé des algorithmes se retrouve dans toutes les franges possibles de discrimination. Comme l'explique G. Packin, dans son papier spécialement axé sur les algorithmes de scoring et le handicap, la qualité des données disponibles sur les personnes handicapées est loin d'être représentative de la population active. Etant donnée la dépendance des systèmes d'IA à l'égard des données empiriques, cette lacune est particulièrement problématique et se doit d'être comblée. Ainsi, la création, le développement et la conception de nouvelles technologies devraient impliquer davantage de personnes handicapées, notamment dans la formation des systèmes basés sur l'IA (Packin 2021).

Ce raisonnement peut donc s'étendre à toutes les catégories de population active actuellement défavorisées par les algorithmes d'apprentissage. C'est ce que les chercheurs appellent le redressement des données de l'apprentissage, méthode qui consiste à pondérer les bases d'apprentissage avec pour but de corriger les biais subis par les catégories discriminées (Lacroux, Martin-Lacroux, 2021) (Besse 2020).

Exclure les variables sensibles d'un algorithme est en effet insuffisant. A l'ère du deep learning, les algorithmes peuvent en effet discriminer sur des critères auxquels ils n'ont pas accès de manière directe. Ainsi, selon plusieurs chercheurs, une approche par calibration, c'est-à-dire une approche par laquelle on s'assure que les résultats obtenus sont indépendants des variables protégées est préférable. Il s'agit ainsi d'accepter de mesurer les discriminations et de tester la présence de biais sur des critères protégés, établis en avance. Plus que l'exclusion de variables sensibles dans un algorithme, c'est l'indépendance et la neutralité du résultat par rapport à ces variables spécifiquement protégées qui garantira donc un algorithme et un processus d'embauche équitable (Bouverot et al, 2020; John-Matthews, Cardon, Balague, 2022). Les auteurs qualifient ce redressement des données de représentation équitable de la société (John-Matthews, Cardon, Balague, 2022).

Le processus de redressement des données consiste à contrôler les outils de recrutement basés sur l'IA, ex ante, c'est -à -dire avant leur mise en fonctionnement sur le marché afin de garantir leur inclusivité. Une telle méthode nécessite des compétences

à la fois techniques (statistiques, développement, apprentissage automatique) et juridiques pour auditer d'une manière conforme au droit les résultats de tels algorithmes, les statistiques ethniques étant interdites en France⁴⁸ et l'article 9 du RGPD interdisant à l'échelle européenne le traitement des données à caractère personnel révélant l'origine raciale ou ethnique d'une personne, ses données génétiques ou de santé⁴⁹ (Besse, 2020).

Cependant, comme l'expliquent plusieurs travaux de recherche, cette méthode s'avère plus compliquée pour les outils d'algorithme basés sur le deep learning en raison de l'opacité de ces systèmes due à l'infinité de données utilisées pour les faire fonctionner (Villani 2018) (Lacroux, Martin-Lacroux 2021).

2.2. Processus envisagés pour élargir les bases de données des algorithmes de recrutement

Dans la mesure où la collecte de données pouvant faire l'objet d'une discrimination est extrêmement sensible en Europe, une approche des algorithmes visant à mesurer et exclure certaines variables doit être strictement encadrée. Dans leur rapport: *Algorithmes: Biais, discrimination et équité*, publié en 2019, Patrice Bertail, David Bounie, Stephan Clémenton et Patrick Waelbroec abordent les possibles manières de corriger les biais inhérents aux données d'entraînement des algorithmes.

Afin de compléter l'information et d'enrichir les données d'entraînement, la première piste statistique envisagée consiste à reconstituer les données manquantes par l'intermédiaire d'un modèle statistique approprié. Les statistiques ethniques étant complexes à réaliser dans plusieurs pays, il s'agit donc de comprendre pourquoi certains individus sont peu représentés dans la base de données de l'algorithme et de modéliser la probabilité qu'ils y figurent effectivement (score de propension à être observé/sélectionné). Si la faible représentation des candidats n'a aucun rapport avec le mécanisme de sélection, les biais peuvent être ignorés. Dans le cas contraire, une correction du biais de sélection est indispensable (Bertail et al, 2019).

⁴⁸ https://www.legifrance.gouv.fr/loda/article_lc/LEGIARTI000031932112/2016-06-14

⁴⁹ https://www.legifrance.gouv.fr/loda/article_lc/LEGIARTI000031932112/2016-06-14

Le redressement des données par ré-échantillonnage est également une option face aux complexités statistiques. Il s'agit de créer une population fictive sur la base de la population cible, en supprimant et/ajoutant des individus artificiellement (Bertail et al, 2019).

Enfin, le rapport *Algorithmes: Biais, discrimination et équité*, suggère d'adopter une analyse moins séquencée des données. Les capteurs fonctionnant en continu, les auteurs expliquent que les données récoltées sont analysées sur des fenêtres temporelles extrêmement restreintes, excluant ainsi des tendances de long terme, des effets périodiques ou certains événements pouvant avoir une influence particulière sur les données. Ainsi, incorporer dans les méthodes d'apprentissage des modèles temporels décrivant les éventuels mécanismes d'évolution des phénomènes analysés est nécessaire à l'élimination des biais et au redressement des données d'entraînement (Bertail et al, 2019).

II. L'instauration d'une vigilance constante dans l'usage des algorithmes de recrutement.

1. La sensibilisation du personnel de recrutement

Afin que les algorithmes de recrutement puissent trouver une place optimale dans un processus de RH dont le fonctionnement était initialement basé sur l'interpersonnel, il est primordial d'intégrer ces outils dans des processus fonctionnels conservant des responsables physiques, formés et préparés à la digitalisation des processus.

Il ne s'agit donc pas simplement d'ajouter des outils permettant aux RH de gagner en efficacité, mais d'adapter leur formation et leurs pratiques pour prévenir et alerter la diffusion des biais algorithmiques. Il est donc primordial que les professionnels du recrutement soient préparés à "garder la main" et à développer un regard critique sur les algorithmes de recrutement.⁵⁰

La recherche insiste ainsi sur l'importance de conditionner les recruteurs et plus généralement les utilisateurs d'algorithmes à interroger la validité prédictive des

⁵⁰ Rapport du défenseur des droits en partenariat avec la CNIL, *Algorithmes : prévenir l'automatisation des discriminations*, 2020

solutions d'IA et à démystifier ces solutions en intégrant les larges possibilités d'erreurs qu'elles présentent et en nuancant de fait les promesses anti-discriminatoires formulées par les promoteurs de ces outils (Lacroux, Martin-Lacroux, 2021; Lazaro, 2023).

Pour inciter les recruteurs à faire ce travail, certains travaux préconisent ainsi d'intégrer une part d'aléatoire dans la présentation des résultats finaux au recruteur afin d'éviter les biais de présentation susceptibles d'influencer son choix au profit des profils apparaissant en première position (Lacroux, Martin-Lacroux, 2021).

La mise en place de chartes de bonne pratiques, ou de lignes directrices éthiques propres aux algorithmes de recrutement incitant les responsables RH à maintenir un degré de vigilance accru face à ces outils pourrait également contribuer à former et sensibiliser le personnel du recrutement aux risques de biais induits par les outils basés sur l'IA (Bouverot et al, 2020; John-Matthews, Cardon, Balagué, 2022).

2. La sensibilisation et l'information des cibles des algorithmes de recrutement

Pour garantir l'équité des outils de recrutement basés sur l'intelligence artificielle, il est primordial d'assurer l'information des candidats quant au processus dans lequel ils s'engagent, la nature de leurs interlocuteurs (humains ou digitaux) ainsi que la manière dont leurs données sont traitées et évaluées.

De nombreux chercheurs insistent ainsi sur l'usage éthique de l'IA et le devoir de transparence et d'explication du recruteur vis-à-vis des candidats dont le consentement libre et éclairé doit impérativement être obtenu afin de poursuivre le processus (Chamorro-Premuzic 2019; Hunkenschroer et Luetge, 2022).

L'appréhension par le candidat de son expérience de recrutement dépendra de sa compréhension plus globale de ce qu'est une intelligence artificielle et de son fonctionnement. Il est donc nécessaire de donner au grand public les moyens de cette compréhension, indispensable à une IA de confiance. Les connaissances de base sur les algorithmes devraient ainsi faire partie de l'éducation dans les écoles et être vulgarisées et accessibles aux adultes. En Finlande, de telles politiques d'éducation sont progressivement mises en place. Ainsi, en 2020, 3,5% de la population avait déjà reçu

une formation à l'IA, un cours de formation ayant été conçu localement pour la population (Bouverot et al, 2020).

En outre, dans son avis sur le livre Blanc sur l'Intelligence artificielle, publié par la Commission Européenne en 2020, le Comité Économique et Social Européen (CESE), insiste sur le rôle des partenaires sociaux et plaide pour qu'ils soient étroitement associés à la mise en place des systèmes d'IA dans le monde professionnel afin de garantir que les systèmes respectent les droits des candidats⁵¹.

3. Le renforcement des obligations de transparence

Afin de garantir un réel équilibre des forces entre employeur et candidat, il est primordial de renforcer les obligations en matière d'information, de transparence et d'explicabilité des algorithmes. L'opacité des systèmes et leur caractère secret constituent en effet un frein à la découverte des biais éventuels.

Le RGPD apporte certaines réponses légales à ces questions, ses articles 13, 14 et 15 imposant notamment l'obligation de fournir "*des informations utiles concernant la logique sous-jacente*" de toute décision automatique entraînant un effet important sur la personne.

Cependant, et comme le précisent le Défenseur des droits et la CNIL dans leur rapport publié en 2020⁵² ces exigences légales de transparence et d'explicabilité doivent aller plus loin afin de lutter efficacement contre les biais discriminatoires. Concrètement, dans le cadre du recrutement, il serait par exemple bénéfique de prévoir que les tiers, et non les seuls destinataires des décisions individuelles puissent accéder aux critères utilisés et à la manière dont ces critères sont agrégés par l'algorithme afin de le comprendre préalablement au processus de recrutement (Sanchez-Mondero et al, 2020).

Cependant, les possibilités de transparence et d'explicabilité des algorithmes de recrutement rencontrent de nombreuses limites. Tout d'abord, la multiplicité des

⁵¹<https://www.eesc.europa.eu/fr/our-work/opinions-information-reports/opinions/livre-blanc-sur-lintelligence-artificielle>

⁵² Rapport du Défenseur des droits en partenariat avec la CNIL, Algorithmes : prévenir l'automatisation des discriminations, 2020, P.11

données utilisées par les algorithmes de deep learning les rend difficilement explicables, y compris pour leurs développeurs (Lacroux, Martin-Lacroux, 2021; Bouverot et al, 2020). En outre, nombreuses sont les entreprises qui refusent de divulguer les détails des algorithmes qu'elles utilisent ou qu'elles ont développés dans un souci concurrentiel (Bertail et al, 2019).

Enfin, selon certains travaux, parvenir à la transparence des algorithmes pourrait générer des effets pervers inattendus dans la mesure où un système où les exigences de l'algorithme sont trop transparentes peut être détourné par les candidats, pouvant adapter leur réponse et "tromper l'algorithme" (Lacroux, Martin-Lacroux, 2021).

4. L'audit des algorithmes de recrutement

Si l'explicabilité des algorithmes présente de nombreux obstacles techniques et commerciaux, l'audit peut constituer une alternative plus facilement réalisable permettant de détecter la présence de biais.

4.1. Les audits par les concepteurs et/ou les utilisateurs des algorithmes de recrutement

Afin qu'ils soient réalisés le plus précisément possible, ces tests devraient idéalement être réalisés par les entreprises déployant elles-mêmes les algorithmes dans la mesure où elles sont les plus aptes techniquement à les effectuer. Elles y ont en outre un intérêt pour limiter les risques juridiques et réputationnels même si de tels tests représentent un certain coût. Le rapport de l'institut Montaigne préconise en outre de recourir à l'accompagnement de la CNIL afin d'aider les entreprises à déterminer les seuils d'équité de leurs algorithmes et les méthodologies de test adéquates (Bouverot et al, 2020).

Dans le cadre du recrutement, il pourrait être bénéfique d'utiliser les données d'échantillons de candidats pour réaliser de tels audits. Il existe en effet aujourd'hui des cas d'exception pour la collecte des données sensibles, notamment dans le cadre d'usages statistiques. Certains chercheurs proposent d'étendre ces exceptions au cas très spécifique de tests d'identification de discriminations éventuelles, sur une collecte de données d'un nombre réduit mais significatif de candidats (20% environ). Un

candidat engagé dans un processus d'embauche pourrait par exemple accepter de partager certaines informations pour contribuer à constituer une base de test, permettant de vérifier l'absence de biais. Cependant, il pourrait être difficile d'acquiescer la confiance de candidats et les convaincre que ces données ne seront utilisées que dans le cadre d'un test et non dans le but de leur sélection (Bouverot et al, 2020).

Une autre option, moins précise, mais permettant de surmonter cette difficulté pourrait consister à adopter une identification de l'origine des candidats par le patronyme (Besse, 2020).

Enfin, en cas de difficultés légales et techniques pour une entreprise déployant ou utilisant un algorithme de recrutement à tester son équité, il pourrait être envisageable de mettre à disposition des bases de données de test publiques pour permettre aux entreprises d'évaluer les biais de leur méthodologie en leur fournissant des bases de données plus représentatives que celles dont disposaient les concepteurs d'algorithmes. A l'échelle française par exemple, ces bases de données pourraient être élaborées en prenant en compte les 25 critères de discrimination prohibés par la loi (Age, genre, ethnique...) ⁵³ afin de s'assurer que certains outils, tels que ceux basés sur la reconnaissance faciale, ne soient pas discriminatoires et qu'une variable protégée n'a pas été "redécouverte" par l'algorithme (Bertail et al, 2019).

4.2. Les audits publics des algorithmes de recrutement

Si un audit public des algorithmes pose des questions évidentes en termes de secret commercial, certaines solutions peuvent néanmoins être envisagées.

Des auteurs préconisent ainsi la nomination d'un tiers de confiance, permettant un examen de l'algorithme par ce dernier, sans permettre la publication de l'algorithme. La CNIL, appuie ce point de vue et recommande même l'introduction d'une plateforme nationale d'audit des algorithmes qui pourraient être réalisés par un organisme public d'experts en algorithmes qui surveilleraient et testeraient les algorithmes (en vérifiant par exemple leur propension discriminatoire) ⁵⁴ (Bertail et al, 2019).

⁵³ https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000045391831;

<https://www.defenseurdesdroits.fr/fr/institution/competences/lutte-contre-discrimination>

⁵⁴ Rapport de la CNIL, Comment permettre à l'Homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle, Décembre 2017, P.59

Qu'ils soient publics ou privés, ces audits permettraient donc aux développeurs des outils de recrutement basés sur l'IA, aux entreprises ainsi qu'aux pouvoirs publics de prendre conscience des biais de ces algorithmes et de ce fait, de tendre à leur perfectionnement ou à une plus grande régulation de ces solutions digitales.

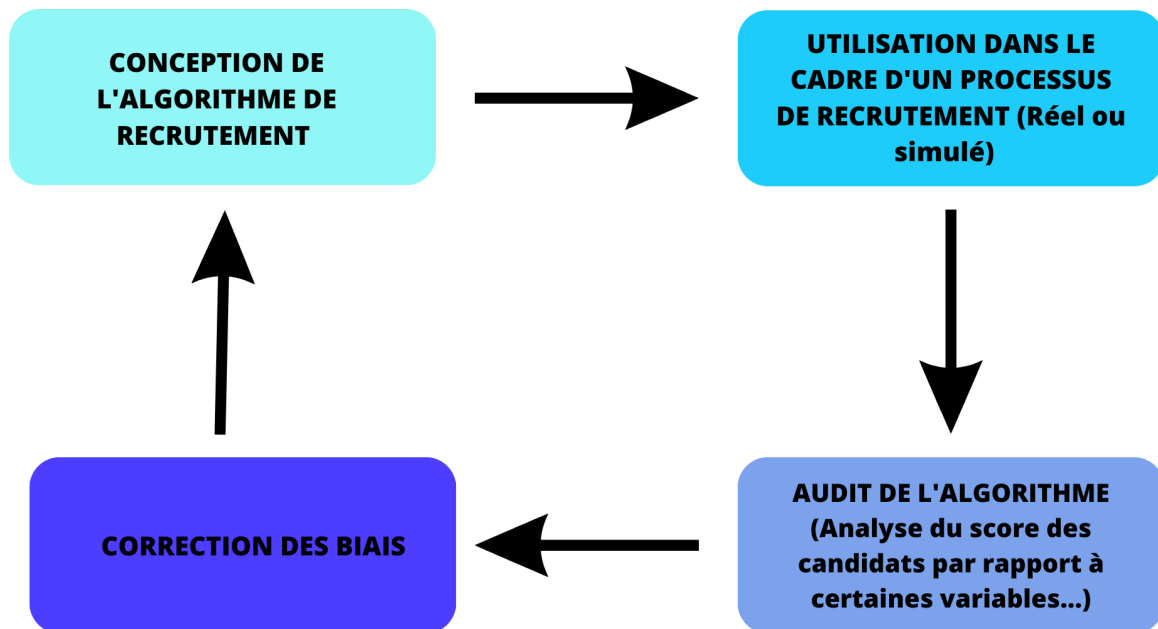


Figure 2: L'audit des algorithmes de recrutement pour la correction des biais incorporés lors de leur conception

4.3. L'élaboration d'une labellisation des algorithmes

Le développement de labels, garantissant la conformité de l'algorithme à certains principes éthiques et non-discriminatoires pourrait s'inscrire dans une logique d'audit des algorithmes. Cela permettrait d'informer les candidats quant à la nature et à la fiabilité ou la légalité du processus utilisé pour les évaluer. L'élaboration de certifications et de labels spécifiques aux algorithmes de recrutement étant un long processus, des premiers labels pourraient se concentrer sur l'auditabilité des algorithmes, sur la qualité des données d'entraînement, sur la sensibilisation du personnel RH ou encore sur la présence d'un processus d'évaluation des risques de biais au sein de la structure (Bouverot et al, 2020).

Une telle démarche irait en outre dans le sens de l'opinion des possibles candidats européens prenant part à un processus de recrutement. D'après un sondage mené en 2019 par le fonds public d'investissement finlandais SITRA, 66% des européens seraient en effet rassurés par l'introduction d'un label dans le cadre d'un processus impliquant la collecte de leurs données personnelles⁵⁵.

III. La spécification des règles juridiques encadrant les algorithmes de recrutement

1. La nécessité de réguler les algorithmes de recrutement

1.1. La mise en place d'un cadre de régulation clair

De nombreux travaux s'accordent pour constater que le cadre juridique en matière de discrimination et de numérique offre déjà de nombreuses solutions. Les Etats et l'UE ayant déjà de nombreuses difficultés à garantir le respect de ce cadre légal, il serait superflu de créer de nouvelles prérogatives plutôt que de garantir l'implémentation, l'explicitation et la pleine effectivité des normes existantes (Bouverot et al, 2020). Il est ainsi primordial de trouver des moyens efficaces de contrôler les algorithmes de recrutement selon ces normes existantes. (Packin 2021).

La complexité des questions numériques et plus particulièrement celles afférentes à l'IA, nécessitent des acteurs spécialisés afin de veiller au respect du cadre normatif par les acteurs publics. En ce sens, le décret n°2020-1102 du 31 août 2020 a créé en France le pôle d'expertise de la régulation numérique (PENEN) dont le rôle est de contribuer à la régulation des plateformes en ligne notamment en matière de traitement des données.

⁵⁶

En sus de la création d'entités spécialisées, il serait en outre nécessaire de mettre en place une coopération efficace entre tous les acteurs axés sur le digital (CNIL, HADOPI...) permettant de renforcer le contrôle du respect des normes par les entreprises en matière d'algorithmes de recrutement (protection des données sensibles, présence d'humains dans la boucle...).

⁵⁵ <https://www.sitra.fi/en/news/the-public-demands-companies-be-transparent-in-their-use-of-data/>

⁵⁶ <https://www.peren.gouv.fr/>

Cependant, en dépit des moyens considérables qui pourraient être mis en place à l'échelle étatique, le monde numérique est bien trop vaste pour être contrôlé ex ante par les pouvoirs publics. Ainsi, certains travaux s'accordent pour affirmer que l'Etat ne peut être seul à contrôler l'absence de biais dans les algorithmes de recrutement. Il est ainsi nécessaire d'envisager de transférer cette mission de contrôle aux entreprises, aux laboratoires de recherche et de certification ou à des tierces parties telles que des sociétés d'audit. Ce contrôle décentralisé pourrait ainsi s'accompagner en parallèle d'une responsabilisation des acteurs face aux conséquences des biais de leurs algorithmes (Bouverot et al, 2020).

Le Comité Économique et Social Européen (CESE) partage en outre cette vision, insistant sur le fait qu'une IA de confiance "*nécessite une approche socio-technique continue et systématique, constituant à examiner la technologie sous tous les angles et à travers différents prismes*"⁵⁷, impliquant ainsi une mise à contribution de tous les acteurs, publics comme privés, dans le contrôle de l'IA.

Enfin, il serait bénéfique que l'initiative législative et réglementaire sur les algorithmes et l'intelligence artificielle actuellement examinée par le Parlement Européen et qui devrait être adoptée d'ici 2024 (IA Act), détermine précisément les limites de l'usage de l'IA dans un contexte décisionnel et ouvre la possibilité de refuser explicitement le recours à certains types d'outils.⁵⁸

1.2. Régulation "molle" : La mise en place de lignes directrices "éthiques"

Plusieurs travaux de recherche ou encore législatifs tels que le Livre Blanc sur l'Intelligence Artificielle publié par la Commission européenne en 2020 ou encore les lignes directrices en matière d'éthique publiées en 2019 par le Groupe d'experts de haut niveau sur l'IA, missionné par la Commission européenne⁵⁹ préconisent l'adoption de lignes directrices fixant un code d'éthique en matière d'IA.

⁵⁷ CESE, avis sur le livre Blanc sur l'Intelligence artificielle, publié par la Commission Européenne en 2020, le Comité Économique et Social Européen (CESE), <https://www.eesc.europa.eu/fr/our-work/opinions-information-reports/opinions/livre-blanc-sur-lintelligence-artificielle>, P.4.

⁵⁸ <https://www.eesc.europa.eu/fr/our-work/opinions-information-reports/opinions/livre-blanc-sur-lintelligence-artificielle>

⁵⁹ Lignes directrices en matière d'éthique pour une IA digne de confiance groupe d'experts de haut niveau sur l'intelligence artificielle constitué par la Commission européenne en juin 2018, 2019

Cependant, ces recommandations restent très générales et peu sectorielles tandis que le domaine du recrutement dispose de spécificités qu'il convient de considérer.

Il pourrait en effet être bénéfique de s'appuyer sur ces cadres généraux de l'éthique de l'IA pour en tirer des lignes directrices détaillées, opérationnelles dans le contexte précis de la sélection de candidats. La contribution des experts du domaine des RH compléterait ainsi celle des développeurs d'IA, impliquant une collaboration étroite entre les disciplines. De telles lignes directrices permettraient de se poser des questions très concrètes et plaquées sur la réalité du recrutement : le candidat est-il informé de la nature du processus ? Comprend-t-il le fonctionnement et les implications des outils utilisés ? Est-il apte à consentir ou à refuser certaines pratiques ? In fine, la sélection fait-elle l'objet d'un jugement humain ? (John-Matthews, Cardon, Balagué, 2022).

2. Sur l'élaboration d'une chaîne de responsabilité claire en présence d'un algorithme de recrutement.

L'utilisation d'algorithmes dans le cadre du recrutement soulève plusieurs questions du fait de la prise de décision par un outil digital. En cas de discrimination, il est en effet sensible de déterminer un régime de responsabilité approprié dans la mesure où elle n'est pas commise directement par un humain et où les discriminations peuvent être involontaires et dissimulées. Se pose alors la question du régime de responsabilité applicable aux algorithmes, à leurs concepteurs et aux entreprises les utilisant.

Si certains décideurs politiques et professionnels de la justice militent en faveur de la création d'un nouveau cadre juridique spécifique visant à attribuer une personnalité aux algorithmes et à les traiter comme des "personnes électroniques" (Bouverot et al, 2020), tel n'est pas l'avis des institutions européennes.

Le livre blanc sur l'IA publié par la Commission européenne en 2020 se positionne en effet, en faveur de l'applicabilité à l'intelligence artificielle des régimes de responsabilité existants⁶⁰. Le CESE partage également cet avis et s'oppose à l'introduction de toute forme de personnalité juridique spécifique pour l'intelligence artificielle, avançant

⁶⁰ Livre Blanc sur l'Intelligence artificielle, publié par la Commission Européenne en 2020, P.16.

notamment qu'un tel régime pourrait permettre aux entreprises de s'exonérer de leur responsabilité en cas de discrimination⁶¹.

En ce même sens, la norme ISO/IEC RR 2408⁶² Overview of trustworthiness in artificial intelligence dispose qu'il est nécessaire de spécifier sur qui la confiance relative au fonctionnement de l'IA repose. Les individus impliqués doivent ainsi pouvoir avoir confiance dans l'IA et dans les autres acteurs de l'écosystème. L'identification et la définition de tous les acteurs humains est une première étape en vue de la constitution de l'écosystème. A cet effet, plusieurs acteurs (le concepteur, l'individu, le certificateur, les prestataires...) peuvent dès lors être identifiés (Chiaroni & Pons, 2022).

Le CESE et la Commission européenne s'accordent donc à se positionner en faveur de la conservation des régimes de responsabilité classiques qu'il s'agirait possiblement d'adapter à l'évolution des fonctionnalités des systèmes d'IA.

En effet, une fois la question du régime juridique de responsabilité écartée, la difficulté principale réside dans la constitution d'une chaîne de responsabilité claire, permettant à une personne discriminée ou dont les données ont été abusivement collectées ou traitées, d'exercer un recours juridique.

Écarter la responsabilité de l'IA en tant que telle revient à attribuer cette responsabilité soit à l'entreprise créatrice de la technologie et/ou à l'entreprise qui l'achète et l'utilise.

Le droit européen (Directive de 1985⁶³) et le droit interne transposé (Article 1245 du Code civil) ont établi des normes spécifiques concernant la responsabilité des produits défectueux. Ces textes prévoient notamment la responsabilité d'un producteur lorsqu'un produit présente un défaut. Cependant, ce régime spécifique suppose la démonstration dudit défaut, or il peut être complexe de prouver avec précision la défectuosité d'une IA lorsque celle-ci génère des biais lors du processus de recrutement. Le régime spécial des produits défectueux ne semble donc pas approprié au cas de l'IA⁶⁴.

⁶¹ CESE, avis sur le livre Blanc sur l'Intelligence artificielle, publié par la Commission Européenne en 2020, le Comité Économique et Social Européen (CESE), <https://www.eesc.europa.eu/fr/our-work/opinions-information-reports/opinions/livre-blanc-sur-lintelligence-artificielle>, P.7.

⁶² Norme internationale créée par des experts de l'organisation internationale de normalisation

⁶³ <https://eur-lex.europa.eu/FR/legal-content/summary/defective-products-liability.html>

⁶⁴ Livre Blanc sur l'Intelligence artificielle, publié par la Commission Européenne en 2020, P.14.

Une analogie avec le droit civil général existant pourrait alors être imaginée afin de déterminer qui est responsable pour les fautes commises par l'IA. Par exemple en droit civil français, la règle posée par l'arrêt Bertrand (Cass. civ. 2ème, 19 février 1997, n°94-21-111) dispose que les parents sont responsables du fait de leur enfant dès lors que ce dernier cause un dommage, même sans faute. On peut facilement imaginer un parallèle dans un contexte où une personne morale (l'entreprise) conçoit un outil générant des dommages sociétaux (un algorithme discriminatoire).⁶⁵

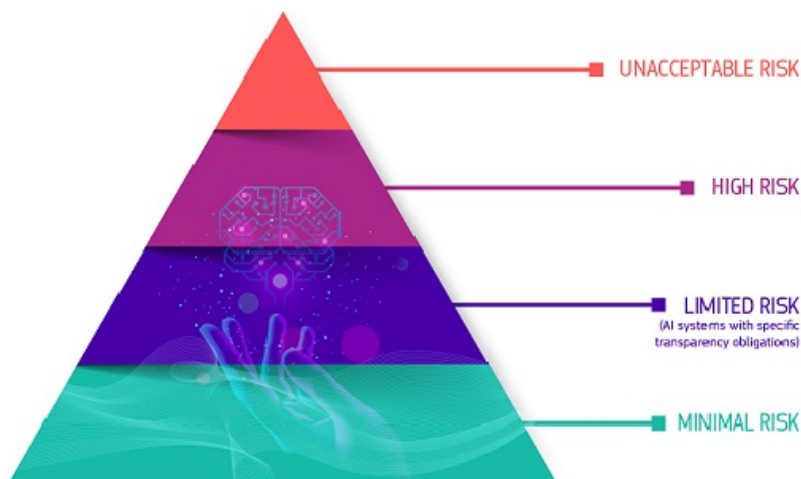
Une faute du fait d'un manquement à leurs obligations de diligence et de vigilance pourrait également être imaginée pour les entreprises concevant et déployant des outils de recrutement basés sur l'IA sans prendre les précautions appropriées (Audits, information du candidat...). Pour rappel, le devoir de diligence est un processus que les entreprises devraient mettre en œuvre pour identifier, prévenir, et atténuer les conséquences négatives réelles et potentielles de leurs activités.⁶⁶

L'établissement d'une telle chaîne de responsabilité nécessiterait donc de clarifier les obligations d'audit et de transparence des algorithmes mentionnées ci-avant. Une carence dans ces obligations pourrait alors constituer une preuve démontrant que les entreprises ont failli à leurs obligations en matière de non-discrimination, la faute étant constituée par la carence d'actions positives par l'entreprise (Mercat-Bruns, 2015).

Le droit communautaire est également en recherche de solutions à cet égard. Ainsi, l'Artificial Intelligence Act, premier cadre réglementaire de l'UE, visant à encadrer l'IA, proposé en 2021 par la commission européenne a été adopté le 14 juin 2023 par le Parlement européen. Ce texte pourrait clarifier plusieurs points quant à la chaîne de responsabilité en matière d'IA. L'IA Act classe en effet les IA selon la gravité des risques qu'elles posent avec des garde-fous adaptés à chaque degré de risque (Chiaroni et Pons, 2022).

⁶⁵ S. MERABET, Vers un droit de l'Intelligence artificielle, Dalloz, Juillet 2020, P.489

⁶⁶<https://economie.fgov.be/fr/themes/entreprises/developper-et-gerer-une/conduite-responsable-des/quest-ce-que-le-devoir-de>



Catégorisation des systèmes d'IA prévue par l'IA Act, Commission européenne⁶⁷

Ce texte considère ainsi les systèmes destinés à servir pour le recrutement (ou encore pour la sélection de personnes physiques) comme “à haut risque”. Ce marquage implique des normes plus strictes visant à limiter les risques identifiés: données d'entraînement de qualité, enregistrement des activités de l'algorithme visant à retracer la genèse d'une éventuelle discrimination, auditabilité de l'algorithme, transparence vis-à-vis des utilisateurs...⁶⁸

Concernant ces IA au degré de risque élevé, l'AI Act propose en outre que soit suivie la procédure réglementaire de marquage CE (Conformité Européenne). Ce marquage, créé dans le cadre de l'harmonisation des législations techniques européennes, est un marquage réglementaire indiquant que le fabricant engage sa responsabilité sur la conformité du produit à l'ensemble des exigences fixées par la législation de l'Union européenne applicable à ce produit. Il inclut en outre la mise à disposition d'une notice d'utilisation (Chiaroni & Pons, 2022).

⁶⁷ <https://digital-strategy.ec.europa.eu/fr/policies/regulatory-framework-ai>

⁶⁸ Proposition de RÈGLEMENT DU PARLEMENT EUROPÉEN ET DU CONSEIL ÉTABLISSANT DES RÈGLES HARMONISÉES CONCERNANT L'INTELLIGENCE ARTIFICIELLE (LÉGISLATION SUR L'INTELLIGENCE ARTIFICIELLE) ET MODIFIANT CERTAINS ACTES LÉGISLATIFS DE L'UNION
<https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A52021PC0206>

IV. Sur le besoin de recherche et d'investissement spécifiques en matière d'algorithmes de recrutement.

A l'heure où les algorithmes sont à la fois présentés comme des outils infaillibles permettant de pallier la subjectivité et la partialité des humains et comme des entités extrêmement opaques et difficilement compréhensibles par le grand public, il est urgent de renforcer les moyens intellectuels et financiers investis dans la recherche concernant l'IA et plus spécifiquement les algorithmes de recrutement (Lazaro, 2023).

La recherche académique sur la question de la validité prédictive des outils de sélection digitaux est en effet encore trop insuffisante pour guider efficacement les professionnels qui utilisent pourtant de plus en plus ce type de technologies dans les processus de recrutement (Lacroux, Martin-Lacroux, 2021).

1. La nécessité d'une recherche transversale

De nombreux travaux de recherche soulignent l'importance de donner un aspect transversal et interdisciplinaire à la recherche en matière d'algorithmes (John-Matthews, Cardon, Balagué, 2022). De tels travaux pourraient ainsi impliquer toutes les parties prenantes au domaine du recrutement algorithmique, à savoir des spécialistes en IA, en réseaux neuronaux (pour notamment comprendre et maîtriser les enjeux), en économie (pour mesurer d'éventuelles discriminations) mais aussi en ressources humaines, en sociologie et en éthique afin d'approfondir les perspectives de "fair learning" c'est-à-dire la conception d'algorithmes répondant à des objectifs d'égalité et d'explicabilité, et pas seulement de performance (Bouverot et al, 2020).

C'est d'ailleurs ce que préconise le Comité Economique et Social Européen qui invite, dans son avis sur le livre blanc sur l'IA, la Commission européenne à favoriser la pluridisciplinarité dans la recherche en mêlant des disciplines différentes (droit, éthique, philosophie) et les parties prenantes intéressées (syndicats, organisations professionnelles d'entreprises)⁶⁹.

⁶⁹ CESE, avis sur le livre Blanc sur l'Intelligence artificielle, publié par la Commission Européenne en 2020, le Comité Économique et Social Eurpéen (CESE), <https://www.eesc.europa.eu/fr/our-work/opinions-information-reports/opinions/livre-blanc-sur-lintelligence-artificielle>, P.3.

Il est en effet nécessaire que la recherche se base sur plusieurs disciplines et plusieurs perspectives afin de développer des études et des méthodologies de prévention de biais. La mise en place de travaux de recherche sur le terrain spécifique qu'est l'Europe est en outre indispensable dans la mesure où la plupart des analyses démontrant les biais discriminatoires des algorithmes de recrutement sont encore basées sur le cas des Etats-Unis. Il est donc primordial de réaliser des travaux similaires, prenant toutefois en compte nos spécificités démographiques et sociales (Bouverot et al, 2020).

2. La nécessité d'une recherche spécifique au secteur du recrutement

Les programmes de recherche relatifs à l'IA sont assez généralistes. Cependant, une recherche spécifique au secteur du recrutement, croisant la réalité de la sélection des candidats et celle du développement d'algorithmes, est nécessaire afin de garantir l'élimination des biais de recrutement.

Il pourrait ainsi être utile de mener des recherches sur la manière dont les départements RH travaillent avec ces nouveaux outils basés sur l'IA afin de capter la réalité de ce secteur spécifique ainsi que l'impact de ces technologies sur le métier de RH. De tels travaux permettraient également d'étudier les effets cognitifs des algorithmes sur les responsables de recrutement, mesurant ainsi le degré de confiance accordé par les professionnels à la machine et à ses décisions (Hunkenschroer et Luetge, 2022).

Afin d'établir un audit et des lignes directrices adéquates dans le secteur du recrutement, l'enjeu de la recherche sur l'IA sera également de déterminer comment des injustices qui ne sont plus basées sur des catégories sensibles (sexe, l'âge, le diplôme, l'éthnie) mais sur des entités plus composites, telles que par exemple les expressions faciales, la voix, peuvent être détectées et corrigées (John-Matthews, Cardon, Balague, 2022).

A l'échelle française, le Gouvernement a lancé en 2018, suite au rapport Villani, une stratégie nationale pour l'intelligence artificielle, visant à accroître les capacités de recherche, former et attirer de nouveaux talents en IA⁷⁰. Afin de traiter au mieux les spécificités de l'IA appliquée au recrutement, il serait primordial que cette stratégie s'étende aux IA utilisées en matière de RH.

⁷⁰ <https://www.economie.gouv.fr/strategie-nationale-intelligence-artificielle#>

3. La constitution de bases de données

Afin de contribuer efficacement à l'élimination des biais, la recherche en matière d'algorithmes de recrutement nécessite de se baser sur des sources de données représentatives de la société.

Il est ainsi nécessaire de développer des infrastructures de recherche à la fois respectueuses des données personnelles et fournissant des informations significatives sur la composition de la population active, cible du recrutement algorithmique. Ces espaces de données sécurisés pourraient constituer une base de données d'entraînement et pourraient servir de cadre dans la recherche.

La CNIL suggère ainsi le lancement d'une grande cause nationale participative pour dynamiser la recherche en IA, la capacité à disposer de très vastes quantités de données constituant en effet l'un des fondements du développement de la recherche. Selon la CNIL, la puissance publique pourrait ainsi jouer un rôle moteur dans la concrétisation d'un tel projet, se portant garante du respect des libertés⁷¹.

4. La nécessité d'investir dans la recherche à l'échelle européenne pour une IA éthique

Pour développer une recherche d'envergure, qui donne les moyens aux concepteurs d'algorithmes ainsi qu'aux entreprises en charge du recrutement d'élaborer des outils efficaces et non-discriminatoires, une augmentation des investissements nationaux et européens dans la recherche et le développement s'impose.

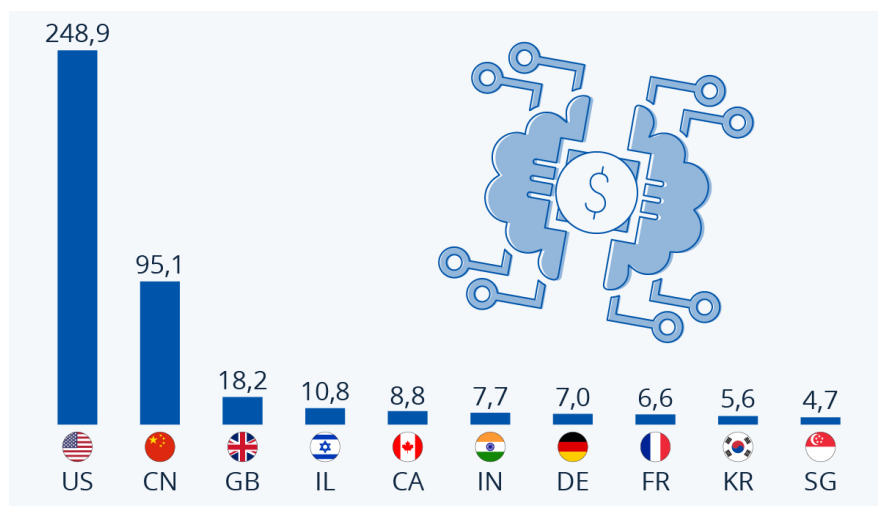
Il est nécessaire que l'Europe renforce son indépendance en matière d'IA par rapport à d'autres régions du monde plus avancées en la matière. Afin d'imposer ses propres standards éthiques et juridiques, l'UE doit en effet être financièrement capable d'investir et de développer ses propres outils.

Face à la concurrence des acteurs américains et chinois, souvent mieux armés financièrement, l'Europe doit en effet imposer ses valeurs non seulement par la régulation mais aussi par une stratégie industrielle par le marché, par les usages et les

⁷¹ Rapport du Défenseur des droits en partenariat avec la CNIL, Algorithmes : prévenir l'automatisation des discriminations, 2020, P.59.

technologies. L'Europe doit ainsi analyser les marchés et les chaînes de valeur de l'IA, dans tous les secteurs (notamment celui du recrutement) et y introduire la notion de confiance (Chiaroni et Pons, 2022).

Cette approche est indispensable dans la mesure où actuellement, les financements consacrés à l'IA au sein de l'UE restent inférieurs à ceux réalisés dans d'autres régions du monde. A titre d'illustration, d'après les données du rapport "2023 AI Index Report" de l'université de Stanford, en 2022, les Etats-Unis arrivaient en tête des investissements mondiaux avec un montant de 248,9 milliards de dollars, suivis par la Chine (95,1 milliards de dollars). L'Allemagne et la France ne se classaient qu'aux 7ème et 8ème rangs avec des investissements de 7 et 6,6 milliards de dollars.⁷²



Quels sont les pays où l'on investit le plus dans l'IA, Pays avec les plus grands investissements privés dans l'IA (2013-2022) en milliards de dollars, Artificial intelligence Index report, Statista

L'UE se mobilise néanmoins de plus en plus afin d'accroître ses investissements dans l'IA et de devenir un leader mondial en la matière. Le plan coordonné sur l'intelligence artificielle, adopté en 2021 par la Commission européenne, vise notamment à accélérer

⁷²<https://fr.statista.com/infographie/29666/pays-avec-les-plus-grands-investissements-privés-dans-ia-entre-2013-et-2022/>

les investissements dans l'IA par les pays membres, à agir sur les stratégies et programmes en matière d'IA et à aligner et unifier la politique portant sur l'IA⁷³.

Dans le cadre du programme Horizon Europe, la Commission a également proposé de consacrer un montant ambitieux au soutien des centres d'essai et d'expérimentation numériques, faisant appel à des investissements européens, nationaux et privés⁷⁴.

Enfin, toujours dans le cadre du programme Horizon Europe, la Commission a mis en place un nouveau partenariat public-privé dans le domaine de l'IA (The IA Data Robotics Association - ADRA) afin de conjuguer les efforts, d'assurer la coordination de la recherche et de l'innovation dans le domaine de l'IA⁷⁵.

CONCLUSION :

L'intelligence Artificielle étant souvent évoquée de manière globale et parfois abstraite pour les non-connaisseurs du secteur du digital, ce papier de recherche visait à aborder le sujet des algorithmes de recrutement in concreto, en décrivant précisément les outils, leur contexte d'utilisation et leur fonctionnement afin d'expliquer leurs controverses de manière illustrée et pratique.

Nous avons ainsi démontré que de nombreux outils basés sur l'IA se superposent au cours des phases du processus de recrutement, présentant ainsi un risque d'accumulation de plusieurs biais. Basés sur des données d'entraînement traitées par des réseaux de neurones (deep learning), les algorithmes de recrutement sont notamment critiqués pour reproduire, voire amplifier les discriminations à l'embauche en dehors de tout contrôle humain. Cette digitalisation du processus de sélection des candidats pose par la même occasion un problème d'opacification du recrutement et d'asymétrie d'information entre les postulants et les recruteurs/développeurs.

De tels outils posent ainsi de nouvelles questions juridiques en termes de données personnelles, de responsabilité en cas de discrimination qui sont encore loin d'être résolues.

⁷³ <https://digital-strategy.ec.europa.eu/fr/policies/plan-ai>

⁷⁴ Livre Blanc sur l'Intelligence artificielle, publié par la Commission Européenne en 2020, P.7.

⁷⁵ <https://adr-association.eu/>

Afin de contrer ces effets négatifs, plusieurs types de bonnes pratiques sont ainsi préconisés. Une sensibilisation de tous les individus (candidats, développeurs, RH, partenaires sociaux) en contact avec ces algorithmes semble d'une part s'imposer. Une correction des données d'entraînement sur lesquelles les algorithmes basent leur fonctionnement est également évoquée dans de nombreux travaux ou rapports. Un audit (public et/ou privé) des algorithmes, une labellisation et la mise en place d'un cadre juridique clair semblent également s'imposer.

Enfin le développement de la recherche en matière d'algorithmes de recrutement et l'allocation de moyens est indispensable afin de garantir notre indépendance en tant qu'Européen et de développer nos propres outils, conformes à nos cadres juridiques tout en mesurant les effets.



Figure 3: Typologie des recommandations pour un usage responsable et éthique des algorithmes de recrutement

Travaux cités:

Abrardi, Laura, et al. "Artificial Intelligence, Firms and Consumer Behavior: A Survey." *Journal of Economic Surveys*, 26 July 2021.

Amos Pah, Clarissa Elfira. "Decision Support Model for Employee Recruitment Using Data Mining Classification." *International Journal of Emerging Trends in Engineering Research*, vol. 8, no. 5, May 2020.

Amzil Karim & Al, *towards a Digital Enterprise: The Impact of Artificial Intelligence on the Hiring Process*, *Journal of Intelligence Studies in Business* 12(3):18-26 . March 2023.

Arora, Sunaina, and Neeraj Kumari. "Recruitment Search Engines for Screening Resumes through AI by Using Boolean Search Functions." *Journal of Asian Development*, vol. 7, no. 2, Oct. 2021.

Azoulay E. *Transformations Numériques L'intelligence Artificielle Pour Diversifier Les Recrutements Chez L'Oréal*, *Journal de l'École de Paris Du Management* 2020/2 (N° 142), Pages 16 à 22.

Bailly, Kevin. "Apprentissage Automatique Pour l'Analyse Des Expressions Faciales." *Theses.hal.science*, 12 Feb. 2019, theses.hal.science/tel-02489704v1. Accessed 19 June 2023.

Barabel et al, *L'IA au service des RH*, Dunod, 2020

Barocas, Solon. *DATA MINING and the DISCOURSE on DISCRIMINATION*. 2014.

Bergman, Peter, et al. "Hiring as Exploration." *SSRN Electronic Journal*, 2020.

Bertail, Patrice, et al. *Algorithmes : Biais, Discrimination et Équité*. 2019.

Besse, Philippe. "Détecer,Évaluer Les Risques Des Impacts Discriminatoires Des Algorithmes D'IA.", May 2020.

Bonnefoy, Antoine, et al. "A Dynamic Screening Principle for the Lasso." *IEEE Xplore*, Sept. 2014.

Casagrande, Annette, et al. "Cerebra, Un Système de Recommandation de Candidats Pour l'E-Recrutement.", juin 2023

Chamorro-Premuzic, Tomas, et al. "New Talent Signals: Shiny New Objects or a Brave New World?" *Industrial and Organizational Psychology*, vol. 9, no. 3, 11 May 2016, pp. 621–640,

Chamorro-Premuzic, "Will AI Reduce Gender Bias in Hiring?" *Harvard Business Review*, 10 June 2019.

Charrière Valérie et Al, *L'impact Des Réseaux Sociaux et Des Compétences Émotionnelles Dans La Recherche d'Emploi : Étude Exploratoire*, *Management & Avenir* (N° 68), Pages 137 à 163, 2014.

Chiaroni et Pons, *IA de Confiance, opportunité stratégique pour une souveraineté industrielle et numérique*, Think Tank Digital New Deal, Collection digital new deal "numérique de confiance", juin 2022.

Chris Anderson. "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete." *WIRED*, *WIRED*, June 2008

Rapport CNIL "Comment Permettre à l'Homme de Garder La Main ? Rapport Sur Les Enjeux Éthiques Des Algorithmes et de l'Intelligence Artificielle", 2017.

Commission Européenne, *Livre Blanc Sur l'Intelligence Artificielle – Une Approche Européenne Axée Sur l'Excellence et La Confiance*, 2020.

Cowgill, Bo, et al. *Bias and Productivity in Humans and Algorithms: Theory and Evidence from Résumé Screening*, 2020.

Crump, Matthew J. C., et al. "Evaluating Amazon's Mechanical Turk as a Tool for Experimental Behavioral Research." *PLoS ONE*, vol. 8, no. 3, 13 Mar. 2013.

Dattner, Ben, et al. "The Legal and Ethical Implications of Using AI in Hiring." *Harvard Business Review*, 25 Apr. 2019

Fernandez Martinez Carmen et Al, *AI and Recruiting Software: Ethical and Legal Implications*, Journal Paladyn, Journal of Behavioral Robotics, 2020 .

Grand'Maison, Philippe. "Génération Automatique de Lettres de Recrutement." Université de Montréal, Mar. 2017.

"Groupe d'Experts de Haut Niveau Sur l'Intelligence Artificielle | Bâtir l'Avenir Numérique de L'Europe." *Digital-Strategy.ec.europa.eu*, 2023.

Harwell, Drew. "A Face-Scanning Algorithm Increasingly Decides Whether You Deserve the Job." *The Washington Post*, 22 Oct. 2019.

Hemamou, Léo, et al. "Attention Slices Dans Les Entretiens d'Embauche Vidéo Différés." *Inria.hal.science*, 3 June 2020.

Janiesch, Christian, et al. "Machine Learning and Deep Learning." *Electronic Markets*, vol. 31, 8 Apr. 2021.

Jeske, Debora, and Kenneth S Shultz. "Using Social Media Content for Screening in Recruitment and Selection: Pros and Cons." *Work, Employment and Society*, vol. 30, no. 3, 20 Nov. 2015.

John-Mathews, Jean-Marie, et al. "From Reality to World. A Critical Perspective on AI Fairness." *Journal of Business Ethics*, 25 Feb. 2022.

Kappen Mitchel & Al, *Objective and Bias-Free Measures of Candidate Motivation during Job Applications* Springer Nature, *Scientific Reports*, November 2021.

Kleinberg, Jon, et al. "Inherent Trade-Offs in the Fair Determination of Risk Scores." Nov. 2016.

Koivunen, Sami, et al. "The March of Chatbots into Recruitment: Recruiters' Experiences, Expectations, and Design Opportunities." *Computer Supported Cooperative Work (CSCW)*, 9 June 2022.

Lacroux, Alain, and 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.

Lavanchy, Maude, et al. "Applicants' Fairness Perceptions of Algorithm-Driven Hiring Procedures." *Journal of Business Ethics*, 12 Jan. 2023.

Lazaro, Christophe. "Le Pouvoir "Divinatoire" Des Algorithmes de La Prédiction à La Prémption Du Futur." *Anthropologie et Sociétés*, vol. 42, no. 2-3, Mar. 2023.

Littlewort, Gwen, et al. *The Computer Expression Recognition Toolbox (CERT)*. 2011.

Comité économique et social européen "Livre Blanc Sur l'Intelligence Artificielle.", 21 Feb. 2020.

Lopez Marcelo, "Institutional Impact of Emerging Artificial Intelligence (AI) Technologies in the Human Resources (HR) Recruitment Sector" Interviews January 2023 Conference.

Mercat-Bruns, Marie. *L'identification de La Discrimination Systémique*, *Revue de Droit Du Travail*. Nov. 2015.

Mihaljević, Helena, et al. "Towards Gender-Inclusive Job Postings: A Data-Driven Comparison of Augmented Writing Technologies." *PLOS ONE*, vol. 17, no. 9, 9 Sept. 2022.

Naulet Antoine, *IA et Décision*, *Revue Défense Nationale* 2019 (N° 820), 2019.

Packin, Nizan Geslevich. "Disability Discrimination Using AI Systems, Social Media and Digital Platforms: Can We Disable Digital Bias?" *SSRN Electronic Journal*, 2020.

Défenseur des droits et CNIL, *Algorithmes : Prévenir l'Automatisation Des Discriminations Face Au Droit, Nous Sommes Tous Égaux*. 2020.

Peña, Alejandro, et al. "Bias in Multimodal AI: Testbed for Fair Automatic Recruitment." Apr. 2020.

Pomerol, Jean-Charles & Frédéric Adam. "Understanding Human Decision Making - a Fundamental Step towards Effective Intelligent Decision Support." *Hal.science*, Springer-Verlag, 2008.

Ruparel, Namita, et al. "The Influence of Online Professional Social Media in Human Resource Management: A Systematic Literature Review." *Technology in Society*, vol. 63, Aug. 2020.

Sánchez-Monedero, Javier, et al. "What Does It Mean to "Solve" the Problem of Discrimination in Hiring?" Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 22 Jan. 2020.

Schumann, Candice, et al. "We Need Fairness and Explainability in Algorithmic Hiring." Semantic Scholar, 2020.

Swapna Jain, Chatbots as a Game Changer in E-Recruitment: An Analysis of Adaptation of Chatbot, in Book: Next Generation of Internet of Things (Pp.61-69), June 2021 .

Tamzid, Mohammad. WILL ARTIFICIAL INTELLIGENCE (AI) REPLACE HUMAN RECRUITERS? Journal of Green Business School, Volume 04, Issue 01, Jan. 2021.

Tanya Perelmuter. "Algorithmes : Contrôle Des Biais S.V.P." Institut Montaigne, Mar. 2020,

Tawk, Charles Joseph, Chatbots: Recruitment Game Changers? Source: International Journal of Business Management & Economic Research . , Vol. 12 Issue 2, P1909-1910. 2p., 2021

Villani , Cedric et al, "Donner Un Sens à l'Intelligence Artificielle" (236p.) [Rapport de La Mission Villani Sur l'Intelligence Artificielle], 2018.

Will, Paris, et al. "People versus Machines: Introducing the HIRE Framework." Artificial Intelligence Review, May 2022.