

Titre : Étendue et enjeux de l'intelligence artificielle dans les emplois professionnels : une perspective pluridisciplinaire

Rubrique : Article de recherche

Auteur(s)

- 1 : Mélanie Trottier, professeure
- 2 : Ewan Oiry, professeur
- 3 : Dominic Martin, professeur
- 4 : Sébastien Gambs, professeur
- 5 : Anne Thibault-Bellerose, conseillère syndicale

Citation : Trottier, M., Oiry, E., Martin, D., Gambs, S. et Thibault-Bellerose, A. (2024). Étendue et enjeux de l'intelligence artificielle dans les emplois professionnels : une perspective pluridisciplinaire. *Ad Machina*, 8(1), 177-199, <https://doi.org/10.1522/radm.no8.1844>

Affiliation des auteurs

- | | |
|---|---|
| 1 : Université du Québec à Montréal | Courriel : trottier.melanie@uqam.ca |
| 2 : Université du Québec à Montréal | Courriel : oiry.ewan@uqam.ca |
| 3 : Université du Québec à Montréal | Courriel : martin.dominic@uqam.ca |
| 4 : Université du Québec à Montréal | Courriel : gambs.sebastien@uqam.ca |
| 5 : Confédération des syndicats nationaux | Courriel : anne.thibault-bellerose@csn.qc.ca |

Remerciements

Déclaration des conflits d'intérêts

- Aucun conflit d'intérêts à déclarer
 Conflit d'intérêts à déclarer (veuillez détailler)

Détails :

Résumé (250 mots)

Les applications de l'intelligence artificielle (IA) sont susceptibles de transformer le travail de professionnels(les) (soignants, juristes, enseignants, travailleurs sociaux, etc.). Cette étude de portée (*scoping review*) s'inscrit dans une collaboration entre les milieux syndicaux et de la recherche afin d'identifier les applications, usages et enjeux de l'IA qui sont documentés en lien avec le travail de professionnels(les). Les résultats montrent que l'IA est très présente dans des secteurs comme la santé, l'administration, le droit et l'enseignement. Ses finalités sont multiples : depuis l'archivage de données jusqu'à la prise de décision en passant par le traitement de textes, les interactions, la reconnaissance ou la simulation. Elle pourrait se développer dans de nombreux métiers et créer des enjeux transversaux majeurs, en particulier autour des compétences, des emplois, de l'éthique et du fonctionnement des organisations. Des enjeux plus spécifiques à chaque métier sont aussi identifiables. Ces résultats permettent de proposer une discussion pluridisciplinaire de ces enjeux en traitant de l'éthique dans la problématique du consentement à l'IA, de la dualité technologique de l'IA, du rôle d'un syndicat par rapport à l'IA et du défi informatique de l'explicabilité de l'IA.

Abstract

Applications of artificial intelligence (AI) are likely to transform the work of professionals (i.e. caregivers, lawyers, teachers, social workers). This scoping review is part of a collaboration between union and research circles to identify the applications, uses and issues of AI that are documented in connection with the work of professionals. The results show that AI is very present in fields such as health, administration, law and education. Its purposes are multiple: from data archiving to decision-making, including word processing, interactions, recognition or simulation. AI could develop in many professions and create major cross-disciplinary issues, in particular around areas of competency, jobs, ethics and the functioning of organizations. Issues more specific to each profession are also identifiable. These results enable to propose a multidisciplinary discussion of these issues by dealing with ethics in the problems of consent to AI, the technological duality of AI, the role of a union in relation to AI and the computational challenge of AI explainability.

Mots clés

Applications IA, éthique, ressources humaines, syndicat, informatique, étude de portée

Droits d'auteur

Ce document est en libre accès, ce qui signifie que le lectorat a accès gratuitement à son contenu. Toutefois, cette œuvre est mise à disposition selon les termes de la licence [Creative Commons Attribution \(CC BY NC\)](https://creativecommons.org/licenses/by-nc/4.0/).

Étendue et enjeux de l'intelligence artificielle dans les emplois professionnels : une perspective pluridisciplinaire

Mélanie Trottier
Ewan Oiry
Dominic Martin
Sébastien Gambs
Anne Thibault-Bellerose

Introduction

L'intelligence artificielle (IA) est un sujet omniprésent dans la sphère organisationnelle et qui anime tous les acteurs de la société (Agrawal, Gans et Goldfarb, 2018; Cardon et al., 2018; Musk, 2014). Les développements rapides en matière d'intelligence artificielle constituent une préoccupation pour tous : travailleurs(euses), gestionnaires, syndicats ou employeurs. Les applications les plus fréquemment discutées sont celles des milieux industriels (Fréour et al., 2021). Or, les employés(es) et leurs organisations syndicales observent maintenant une pénétration de l'IA dans les métiers professionnels (p. ex., outils de prise de décision juridique et d'analyse de risque de délinquance), ce qui pose de nombreuses questions relevant non seulement du droit du travail, mais également de l'éthique, de la gestion des ressources humaines, de l'autonomie et de la responsabilité professionnelles.

La présente recherche s'inscrit dans une collaboration entre les milieux syndicaux et de la recherche afin de documenter ces enjeux, dont il importe de se saisir afin d'anticiper les développements à venir dans les milieux de travail et s'y préparer. C'est dans cette optique que la présente étude pose la question suivante : **quels sont les applications, usages et enjeux de l'IA documentés en lien avec le travail des professionnels(les)?** Plus spécifiquement, elle propose une large recension des applications de l'IA existantes dans les emplois professionnels et des enjeux qui y sont associés.

Ces résultats sont ensuite analysés et problématisés en regard de plusieurs disciplines. Tout d'abord, les recherches en éthique des affaires interrogent ces résultats en identifiant la question suivante : pour quel groupe au sein d'une organisation l'IA est-elle développée? L'activité des employés(es) ne fait partie de la réflexion que si ceux-ci sont considérés comme partie prenante de l'organisation. Dans les théories de la primauté actionnariale, l'activité des employés(es) n'est ainsi pas présente. Cette problématisation préalable est donc cruciale et permet également de poser la question des conditions du consentement volontaire chez les employés(es) (Olsaretti, 2004). Ces questionnements éthiques initiaux permettent ensuite, du point de vue de la recherche en GRH, de problématiser les résultats de l'étude de portée sous l'angle de la dualité qui apparaît comme centrale dans les effets de l'IA sur le travail des professionnels(les) dans les organisations (Meijerink et Bondarouk, 2022). Par la suite, cette problématisation est enrichie par le point de vue d'un syndicat de professionnels(les) sur cette dualité des effets de l'IA. Enfin, cette problématisation aboutit, du point de vue des recherches techniques en informatique, à un questionnement sur l'explicabilité des résultats produits par les algorithmes d'IA utilisés dans les organisations (Lipton, 2018).



Cette étude revêt un caractère exploratoire en visant à recenser les applications de l'IA utilisées ou projetées dans certains milieux de travail ainsi que les enjeux et impacts documentés. Elle contribue aux connaissances actuelles de trois manières complémentaires. Sa première contribution est de présenter un panorama détaillé des applications d'IA qui existent dans les métiers professionnels. Sa deuxième contribution est d'identifier les impacts (positifs et négatifs) que ces applications ont sur le travail de ces professionnels(les). Ce faisant, elle permet de dépasser les enjeux simplement pressentis ou projetés. Sa troisième contribution est de proposer une problématisation pluridisciplinaire de l'IA dans les métiers professionnels. Tous ces éléments sont autant de contributions à l'orientation de l'action des praticiens(nes) qui devront gérer l'arrivée des applications d'IA dans leurs milieux de travail.

1. L'intelligence artificielle, le travail et les emplois professionnels

L'IA est habituellement définie comme « *a broad class of technologies that allow a computer to perform tasks that normally require human cognition, including adaptive decision making* » (Tambe, Cappelli et Yakubovitch, 2019, p. 16). L'automatisation du travail consiste à confier à une machine une tâche jusqu'alors réalisée par un humain (Freysenet, 1992). L'objectif est que la machine soit plus fiable que l'humain et ne connaisse aucune variation dans l'exécution de cette tâche. L'émergence de données massives (*big data*) au début des années 2000 a permis de transformer fondamentalement cette situation. Avec l'IA, l'objectif est recadré afin que la machine soit capable d'adapter l'automatisation des tâches à des situations nouvelles (comportement des individus, variations de l'environnement, etc.). Grâce à une base d'entraînement composée de très nombreuses données, l'IA parvient à développer des algorithmes qui sont progressivement capables de proposer des réponses et réaliser des actions adaptées au profil d'un individu spécifique, à une situation nouvelle, etc. (Tambe, Cappelli et Yakubovitch, 2019). Par exemple, en disposant d'une base d'entraînement qui contient des milliers de décisions de recrutement prises par des recruteurs à partir de curriculum vitae (CV), l'IA permet d'identifier, parmi de nouveaux CV soumis, ceux qui correspondent le mieux à une offre d'emploi donnée (Burell, 2016).

Dans cette recherche, nous nous intéressons à la fois aux applications basées sur l'IA « symbolique », où la connaissance est encodée par un humain sous la forme de règles que le système utilise pour raisonner, qu'aux applications basées sur l'IA « connexionniste », où le modèle qui sert à faire la prédiction n'est pas codé directement par un humain, mais plutôt appris automatiquement à partir des données (Mira, 2008). Étant en mesure de prendre des décisions complexes dans des situations nouvelles, les applications issues de l'IA ont la capacité de prendre une place conséquente dans le travail des professionnels(les).

Au Québec, le Code des professions (2021) définit les professionnels(les) comme « toute personne qui est titulaire d'un permis délivré par un ordre et qui est inscrite au tableau de ce dernier », Le Québec compte plus de 400 000 professionnels(les) regroupés(es) dans 46 ordres professionnels et 55 professions très variées telles qu'ingénieurs(es), médecins, avocats(es), enseignant(e)s, travailleurs(euses) sociaux(ales) et agronomes. Le Code des professions précise que les professionnels(les) possèdent des savoirs très spécialisés et d'un niveau de complexité élevé (Office des professions, 2021). Étant donné cette complexité, les actes d'un(e) professionnel(le) ne peuvent généralement être évalués que par des pairs. Il est donc aujourd'hui tout particulièrement pertinent de s'intéresser aux impacts de l'IA sur le travail des professionnels(les). En effet, l'automatisation puis l'IA ont principalement touché le secteur industriel (Roll et Ifenthaler, 2020; Schaeffer, 2017). L'industrie 4.0 est généralement le terme utilisé pour désigner le fait que les processus industriels tentent de profiter du traitement des données massives pour produire des algorithmes qui permettent d'améliorer leurs performances (p. ex., en matière de maintenance préventive ou d'optimisation de la production) (Kohler et Weisz, 2016). Dans ce contexte, c'est l'impact de l'IA sur les métiers faiblement qualifiés qui est analysé (Brynjolfsson et al. 2018). Or, on constate aujourd'hui que l'IA peut être appliquée dans des métiers qualifiés. Contrairement à l'automatisation, l'IA permet d'adapter

des actions et des réponses à des situations nouvelles et à la spécificité des cas analysés. Elle a donc la capacité de soutenir, et peut-être de remplacer, des actes d'analyse et diagnostic qui sont au cœur du travail des professionnels(les) (Lacroux et Martin-Lacroux, 2021). C'est dans ce contexte que cette étude pose la question suivante : **quels sont les applications, usages et enjeux de l'IA documentés en lien avec le travail des professionnels(les)?**

2. Méthodologie

La méthode de l'étude de portée est mobilisée pour rendre compte de l'étendue, de l'ampleur et de la nature des applications et enjeux de l'IA actuels et projetés dans les métiers professionnels. Pour ce faire, la méthode d'Arksey et O'Malley (2005) est utilisée. Cette méthode offre un cadre structuré autour de cinq étapes : 1) identification de la question de recherche; 2) identification de la documentation visée; 3) sélection de la documentation à analyser; 4) cartographie des données; 5) synthèse des résultats.

Tel qu'introduit plus tôt, cette étude s'attarde à identifier les applications de l'IA qui sont ou qui vont être utilisées dans les métiers des professionnels. Afin d'identifier la documentation susceptible de répondre à cette question, un vaste appel de textes a été conduit sur des bases de données de plusieurs disciplines afin de couvrir divers milieux dans lesquels œuvrent des professionnels(les) (tableau 1). Outre ces bases de données, des recherches ont été effectuées sur les sites d'associations et ordres professionnels (p. ex., Association du personnel de soutien administratif du Québec, Association des archivistes du Québec).

Tableau 1
Bases de données sondées pour l'appel de textes scientifiques par secteur d'activité

Secteur	Bases de données
Administration et employabilité	ABI/INFORM, PsycINFO, Business Source Complete, Emerald, JSTOR, ERIC, SCOPUSWeb.
Communication sociale et publique	Ad Age Data Center, Communication and Mass Media Complete, Emerald Insight, IBSS (ProQuest), Sage Journals.
Éducation	PsycINFO, Repère, Cairn, Érudit, Education Source, ERIC, LearnTechLib.
Psychologie	psycARTICLES, Psychology + Behavioral Sciences Collection.
Santé	Cochrane Library, MEDLINE, PubMed, SCOPUS, SciFinder.
Sciences politiques	Academic Search Complete, Sociological Abstracts.
Science de l'activité physique	SPORTDiscuss.
Science juridique	Doctrinal Plus, HeinOnline, Jurisquare, Kluwer Law Online Journal
Travail social	Banque sur la recherche sociale et en santé, Social Services Abstract

Une combinaison de mots-clés recherchés dans le résumé (*abstract*) a été utilisée pour identifier la documentation s'intéressant à une diversité de professions (tableau 2). Sur la base de l'appel de textes, 867 documents ont été identifiés (incluant les doublons).



Tableau 2
Les mots-clés utilisés pour l'appel de textes scientifiques

Mots en lien avec l'IA		Mots en lien avec les professionnels(les)
Artificial intelligence/AI/Intelligence artificielle Deep learning Machine learning/apprentissage machine Neural network/réseau de neurones Natural langage understanding Cognitive computing	ET	Administration/Employabilité Admin*; Direction; Emploi; Comptabilité; Organisation; Travail administratif; Business developer; aide entreprises; développement économique; startup; Employabilité; ressources humaines; Human resources ; conseiller en orientation; emploi; carrière; Coordo*; Host, accueil; réception; Archives; Archivist; Archiviste; bibliothèque; Librarian. Droit Lawyer; Droit; Stagiaire; Stage pratique. Éducation Education; psychoéducation; Professeur; Université. Loisirs Safeguard; Accompagnement; garde; sauvetage. Politique Attaché de presse; Politique. Psychologie/Travail social Social worker; Psychosocial worker; Travail social; Psychologist; Social services; Care; caregiver. Santé Ergotherapy; Kinesiology; Art therapy; Nutrition; Biomedical engineering ; Orthopedagogy; Homoeopathy; Obstetrician; Midwife; Biomed*; Diet; Radiology; radiologie; Medical research ; Cytogenic; Genes; Transfusional security; Spiritual; Religion; Archivist; Archiviste; bibliothèque; Librarian.

Afin de sélectionner la documentation à analyser, les critères d'inclusion suivants ont été appliqués : 1) document publié (article de revue, de journal, éditorial ou statistiques) en français ou en anglais; 2) document traitant d'une application, d'un enjeu ou d'un impact de l'IA; 3) document portant sur un ou plusieurs milieux ou métiers professionnels. Suivant cette étape, 158 documents ont été retenus et analysés à l'aide d'une grille de codage visant à mettre en lumière la nature des applications de l'IA recensées, mais également les impacts et enjeux qui y sont associés. Plus spécifiquement, chaque document a été codé en regard des critères suivants : année, pays, discipline, nature du document, professions/milieux visés, application/usage de l'IA, utilisation actuelle/projetée et enjeux identifiés. Le tableau 3 présente un portrait global de la documentation retenue.

Tableau 3
Caractéristiques générales des documents retenus pour analyse

Caractéristiques	Littérature scientifique nb (%)	Littérature grise nb (%)	Total nb (%)
Années de publication	<i>N=50</i>	<i>N=108</i>	<i>N=158</i>
1986-1999	3 (6 %)	–	3 (2 %)
2000-2009	7 (14 %)	–	7 (4,5 %)
2010-2015	6 (12 %)	1 (1 %)	7 (4,5 %)
2016-2020	34 (68 %)	104 (96 %)	138 (87 %)
Non spécifié	-	3 (3 %)	3 (2 %)
24 pays différents dont le plus souvent	<i>N=52</i>	<i>N=110</i>	<i>N=162</i>
Canada	9 (17,3)	33 (30)	42 (26)
France	1 (1,2)	34 (30,9)	35 (21,7)
États-Unis	13 (25)	19 (17,3)	32 (19,8)
Grande-Bretagne	5 (9,6)	3 (2,7)	8 (5)
Finlande	6 (11,5)	-	6 (3,8)
Non spécifié	1 (1,9)	5(4,5)	6 (3,8)
Type de documents			
Littérature grise		<i>N=108</i>	
Articles de magazines/revues en ligne		15 (13,9)	
Articles/communiqués/éditoriaux de presse		42 (38,9)	
Articles de sites web (p. ex., ordres professionnels, ministères, entreprises, spécialisés dans un domaine)		38 (35,2)	
Livres		1 (0,9)	
Rapports/dossiers		12 (11,1)	
Littérature scientifique	<i>N=50</i>		
Revue de la littérature – Théorique	19 (38)		
Revue de la littérature – Empirique	29 (58)		
Méta-analyse	0 (0)		
Thèse	1 (2)		



3. Résultats

L'analyse a permis de caractériser les documents retenus de trois façons : en identifiant les applications et usages de l'IA utilisés par les professionnels(les); en repérant des enjeux associés à des types d'applications de l'IA; en relevant des enjeux d'applications de l'IA en lien avec des professions spécifiques.

3.1 Application et usages de l'IA chez divers professionnels(les)

Les résultats de la présente étude permettent d'abord de rendre compte de la vaste étendue des applications de l'IA. Le tableau 4 présente le nombre d'articles traitant d'applications issues de l'IA dans un secteur d'activité spécifique. On constate que l'IA est présente dans de très nombreux secteurs d'activité des professionnels(les). Elle doit donc être considérée comme une technologie qui, potentiellement, pourrait toucher tous les secteurs d'activité d'une économie. Ce tableau permet aussi d'identifier que certains secteurs d'activité semblent plus actifs quant à l'intégration d'applications issues de l'IA. Si la documentation suit la réalité sur le terrain, l'analyse des effets de l'IA sur le travail des professionnels(les) devrait donc prioriser ces secteurs d'activité pour en comprendre les effets. La santé est le secteur dans lequel les applications d'IA semblent être plus nombreuses avec près de 50 % des applications discutées dans la littérature. L'éducation et l'administration sont également très bien représentées.

Tableau 4
Applications de l'IA selon les secteurs d'activité

Secteur	Littérature scientifique (%) nb N=50	Littérature grise (%) nb (%) N=108	Total nb (%) N=158
	N=54*	N=143*	N=197*
Santé	31 (57,4)	57 (39,8)	88 (44,7)
Administration (incluant recrutement)	10 (26,5)	30 (21)	40 (20,3)
Éducation	5 (9,3)	29 (20,3)	34 (17,3)
Politique	2 (3,7)	11 (7,7)	13 (6,6)
Droit	5 (9,3)	6 (4,2)	11 (5,6)
Services sociaux	1 (1,9)	7 (4,9)	8 (4)
Loisir	0 (0)	2 (1,4)	2 (1)
Services auxiliaires	0 (0)	1 (0,7)	1 (0,5)

* Certains articles traitent de plus d'un secteur à la fois.

L'analyse de la documentation permet également de distinguer les applications de l'IA selon l'objectif qu'elles poursuivent. Le tableau 5 reprend la classification de Leicht-Deobald et al. (2019) qui propose une forme de caractérisation des applications de l'IA selon leur objectif. Ce tableau présente également quelques exemples d'applications issues de la présente analyse de littérature qui y sont associées. Cette catégorisation permet de constater un niveau d'implication de l'IA dans le travail des professionnels(les) qui diffère selon les applications recensées et suscite des questionnements quant aux impacts sur le travail même des professionnels(les). Ces enjeux seront discutés à la section 5.

Tableau 5
Caractérisation des applications de l'IA selon l'objectif (Leicht-Deobald et al. 2019)

Objectifs de l'IA	Exemples d'applications
L'analyse descriptive. De nature exploratoire, cette analyse vise à résumer et visualiser un ensemble de données.	Classement de dossiers médicaux.
L'analyse prédictive. Elle synthétise les données collectées dans le passé ou en temps réel pour créer des modèles capables de prévoir ou deviner l'état futur des variables.	Évaluation du risque de récurrence d'un(e) client(e).
L'analyse prescriptive. Elle analyse les données pour prévoir l'état futur des variables, mais peut également commander ce qu'il faut faire à l'avenir et comment le faire.	Soutien à la décision par la proposition de traitement/thérapie.

Alors que la catégorisation précédente (tableau 5) permet de comprendre comment l'IA peut utiliser les données de différentes façons, le codage de la documentation effectué dans la présente étude permet également de dégager des catégories d'applications au sein des métiers/milieus professionnels selon la finalité visée par l'application. Le tableau 6 présente les applications recensées selon qu'elles ont pour but de : 1) traiter du texte; 2) créer une interaction avec un(e) utilisateur(trice); 3) reconnaître l'expression faciale, la voix ou des images; 4) simuler une situation spécifique; 5) stocker ou transmettre de l'information; 6) soutenir la prise de décision que ce soit en prédisant la probabilité de survenue d'un événement, proposer une option ou réaliser des tâches. Ces résultats permettent de constater l'étendue des applications quant aux types de tâches qu'elles soutiennent ainsi qu'à l'égard de l'ampleur de l'implication qu'elle suppose avec les clients(es) ou usagers(ères). On y voit notamment de nombreux exemples de fonctions que les applications de l'IA recensées dans la littérature peuvent avoir dans les milieux organisationnels.



Tableau 6
Exemples d'applications de l'IA classées selon leur finalité

Fonction	Nb articles issus de la littérature scientifique	Nb articles issus de la littérature grise	Exemples
Traitement de textes	15	16	<ul style="list-style-type: none"> • Exploration, analyse et classification de textes (analyse de CV, rédaction de descriptions de postes, triage de demandes de visa ou d'admission, repérage d'enjeux médiatiques dans les documents, classement de dossiers médicaux, analyse sémantique de comptes rendus médicaux, analyse de la jurisprudence) • Traduction (traduction simultanée pour soutenir la relation avec un(e) client(e)). • Rédaction intelligente (rédaction automatique de contrats ou autres documents, transcription automatique de comptes rendus de réunions, d'archives).
Interaction	8	26	<ul style="list-style-type: none"> • Plateforme interactive (évaluation psychologique, synchronisation de dossiers de patients, suivi des patients, travail collaboratif sur des documents). • Agent conversationnel/virtuel (entretiens avec des candidats(es), patients(es) en psychothérapie, suivi de femmes enceintes, triage de patients(es) en milieux médicaux, réponses à des demandes téléphoniques (clients(es), étudiants(es), etc.), robot compagnon pour les personnes âgées). • Tutoriel intelligent (formation d'étudiants(es) à distance). • Assistant numérique (pilulier intelligent, diagnostic à partir d'images envoyées par un(e) patient(e)). • Interfaces pour susciter l'engagement des patients(es)/clients(es) (envoi de messages de motivation au cours de la journée lors de l'arrêt de fumer, interventions virtuelles d'infirmiers(ères) pour encourager la prise de médicaments ou l'activité physique).
Reconnaissance	7	9	<ul style="list-style-type: none"> • Reconnaissance faciale (soutien d'un diagnostic psychologique, évaluation de candidats(es), évaluation de la crédibilité d'un témoin, assister l'interaction avec des clientèles spécifiques (p. ex., enfants autistes). • Reconnaissance vocale (soutien d'un diagnostic, transcription, facilitation de la communication avec des personnes présentant des problèmes d'élocution, identification de la répartition du travail dans des activités d'équipe). • Reconnaissance d'images (association de documents, identification de caractéristiques spécifiques dans les édifices pour faciliter l'entretien, cellules cancéreuses et autres maladies, besoins/faiblesses/forces de patients(es) à domicile, difficultés familiales, difficultés langagières, erreurs des étudiants(es), compétences d'un(e) candidat(e)).
Simulation	1	10	<ul style="list-style-type: none"> • Avatar/réalité virtuelle (traitement du stress post-traumatique, formation clinique du personnel hospitalier, rééducation du trouble de l'équilibre). • Jeux éducatifs (développement de compétences [p. ex., immobilières, mathématiques]).

Fonction	Nb articles issus de la littérature scientifique	Nb articles issus de la littérature grise	Exemples
Stockage et transmission d'information	2	0	<ul style="list-style-type: none"> • Suivi d'échanges entre spécialistes du droit. • Archiver du contenu clinique.
Analyse de données pour soutenir la prise de décision (prédire une situation ou un état)	18	38	<ul style="list-style-type: none"> • L'état d'un patient (glycémie, potentiel de réadmission, risque de récurrence, succès d'un traitement, progression d'une maladie). • Les besoins de ressources (hôpitaux). • La durée (temps d'attente dans une clinique ou dans le traitement d'un dossier juridique). • L'employabilité d'un(e) candidat(e) (traits psychologiques). • Le comportement d'un(e) étudiant(e) (chute de l'attention, décrochage). • Les conditions d'apprentissage adéquates (ressources fonctionnant le mieux, choix de cours/programme). • Une issue juridique (décision de la cour, résultats d'une réclamation). • Le taux de succès (étudiant(e), avocat(e)). • La survenue d'un événement (précipitations mettant en péril un événement récréotouristique).
Analyse de données pour soutenir la prise de décision (proposer une option, réaliser des tâches)	5	8	<ul style="list-style-type: none"> • Soutenir le suivi d'un patient/sélection d'un traitement (choix des tests complémentaires, choix du traitement selon l'expression faciale en cas de dépression). • Offrir une rétroaction au patient (conseils diététiques). • Automatiser des tâches (tâches du radiologue, tâches chirurgicales, remplacer l'opérateur(trice) d'échographie, gestes de suture, gestion de la ventilation mécanique, planification/reprogrammation de réunions, entretiens d'embauche, prise de présence en classe, rétroaction/correction de travaux scolaires). • Assister la prise de décision (évaluation du type d'accouchement adéquat, score de risque à l'intention des juges, évaluation de la crédibilité de faits, choix des clauses d'un contrat, système expert de gestion).

3.2 Enjeux de l'IA chez les professionnels(les)

Le codage de la documentation a permis de mettre en exergue une variété d'enjeux relatifs aux applications de l'IA, enjeux actuels ou potentiels en lien avec le travail ou les conditions de travail des professionnels(les). Les résultats (tableau 7) montrent que de nombreux enjeux sont transversaux à de multiples professions et milieux. On y constate que ces enjeux sont principalement liés aux compétences requises, aux risques pour les professionnels, aux questionnements éthiques et aux enjeux pour les organisations.



Tableau 7
Enjeux transversaux aux professions/milieus

Cible de l'enjeu	Enjeux transversaux
Compétences	<ul style="list-style-type: none"> • L'IA pourrait provoquer un phénomène de surspécialisation des métiers poussant certains(es) professionnels(les) à développer des compétences très pointues. • Certains parcours professionnels dans les secteurs ou sous-secteurs qui seront fortement impactés par l'IA pourraient être fragilisés. • Les programmes de formation et de développement professionnel devraient inclure des bases de données scientifiques pour préparer les cliniciens à la pratique assistée par l'IA. • Besoin de collaboration entre les disciplines ainsi qu'entre les développeurs d'applications et les professionnels(les) afin d'assurer une pertinence de ces applications pour les usagers(ères) et clients(es) /patients(es).
Professionnel(le)s	<ul style="list-style-type: none"> • Résistance potentielle de la part des utilisateurs(trices). • Potentiel de stress lié au risque d'obsolescence des compétences par rapport à l'automatisation ainsi que la potentielle augmentation de la charge de travail. • Plusieurs risques potentiels pour les travailleurs(euses), notamment la surcharge informationnelle, la généralisation du travail à distance et la surveillance accrue.
Éthique/confidentialité	<ul style="list-style-type: none"> • Divers enjeux quant aux données collectées : éthique, confidentialité, légalité. • Les algorithmes de l'IA peuvent être contaminés par les préjugés humains. • La responsabilité en cas d'erreur issue de l'utilisation d'une application de l'IA est difficile à déterminer. • Plusieurs applications sont rendues disponibles malgré l'absence d'évidences scientifiques à leur endroit. • Les algorithmes doivent être transparents pour pouvoir être évalués et critiqués en regard de la qualité du résultat. Il y a une incapacité à expliquer la logique et l'argumentation se cachant derrière des réponses ou des informations offertes par des applications.
Organisations	<ul style="list-style-type: none"> • Coûts élevés de développement et d'intégration des applications de l'IA; l'IA reste inaccessible pour plusieurs milieux. • La rareté de la main-d'œuvre pourrait pousser les organisations souhaitant mettre en œuvre des solutions fondées sur l'IA à se tourner vers des organismes extérieurs.

Outre ces enjeux qui sont présents pour l'ensemble des milieux organisationnels où œuvrent les professionnels(les), certains enjeux sont quant à eux plutôt spécifiques à certains métiers/emplois. Le tableau 8 présente ces enjeux complémentaires qui sont identifiés dans la documentation comme étant particuliers à des professions ou des milieux. On y constate la diversité des questionnements que soulèvent ces applications pour les acteurs(trices) de ces milieux.

Tableau 8
Enjeux spécifiques à des professions ou des milieux

Professions/ milieux	Enjeux spécifiques
Psychologie/ Travail social	<ul style="list-style-type: none"> • Évaluation nécessaire du degré de confort des patients(es) avec la technologie. • Les professionnels(les) pourraient voir leur jugement professionnel moins sollicité; risque de susciter une éventuelle déqualification de leurs compétences. • Les systèmes d'information pourraient faire évoluer le travail social généraliste vers d'autres fonctions telles que spécialiste du traitement algorithmique.
Droit	<ul style="list-style-type: none"> • Besoin de collaboration entre cabinets et chercheurs pour l'ingénierie des processus juridiques. • Plusieurs tâches sont susceptibles d'être bouleversées par l'IA, notamment la recherche documentaire, le traitement de textes et la production de statistiques juridiques. • L'utilisation de la justice analytique semble mener vers une diminution du coût moyen d'une consultation et vers une escalade de la « finesse » des arguments juridiques. • La démocratisation des applications de justice prédictive pourrait jouer un rôle potentiellement significatif dans l'abandon d'un désir d'action en justice. • Certains outils permettent aux avocats(es) de se concentrer sur les tâches les plus compliquées, en automatisant les tâches plus faciles et répétitives. • Besoin de clarifier ce qui constitue une responsabilité civile et contractuelle lorsqu'un professionnel suit les recommandations de l'IA.
Santé	<ul style="list-style-type: none"> • Nécessité de valider l'IA en milieu clinique et d'augmenter la collaboration entre scientifiques et cliniciens. • Les différences quant aux normes de confidentialité selon les pays freinent le partage de données. • Des fournisseurs d'applications pourraient favoriser certains traitements. • L'exploitation des données médicales est complexe puisque 80 % des informations sur les patients sont textuelles. • Les algorithmes pourraient renforcer des préjugés (les principaux biais sont dus à la surreprésentation de certaines catégories de personnes dans les données, comme les personnes âgées ou des patients d'origine géographique particulière). • Les nouvelles technologies peuvent décentraliser une partie des tâches réservées aux hôpitaux, comme la détection précoce de maladies graves, et peuvent aider à réduire les coûts du système. • Les professionnels(les) de l'imagerie médicale sont particulièrement touchés par la généralisation prévisible de la lecture d'images automatisée. • L'évaluation des soins à domicile à l'aide l'IA peut déshumaniser l'évaluation des soins (engendre plus de travail administratif et réduit le temps passé auprès des patients(es)).
Éducation	<ul style="list-style-type: none"> • Nécessité de développer de nouvelles collaborations interuniversitaires et créer de nouvelles spécialités dans le milieu de l'éducation. • Trouver un équilibre entre le maintien d'aspects traditionnels qui ont fait la richesse de l'enseignement (interactions sociales) et l'IA. • La technologie de l'IA demeure dispendieuse et reste inaccessible pour l'ensemble des établissements d'enseignement supérieur. • La technologie risque de progresser plus rapidement que les systèmes éducatifs.
Employabilité/ Administration	<ul style="list-style-type: none"> • Risque d'accentuation de la complexité des tâches à effectuer, donc un risque d'augmentation de l'intensité du travail, avec la disparition des tâches plus simples. • Risque d'isolement du travailleur, car l'IA permet d'accéder aux informations requises pour son activité au moyen d'une application plutôt que par un contact humain.
Politique	<ul style="list-style-type: none"> • Les applications de l'IA dans le secteur public sont encore peu connues • Les écarts entre les attentes des citoyens et les capacités gouvernementales en matière d'IA sont grands.



Les résultats de l'étude de portée que nous venons de présenter démontrent que l'IA est susceptible d'être présente dans l'ensemble de l'activité des professionnels(les). Ses finalités très variées (traitement de textes, soutien des interactions et différentes formes d'analyse des données pour prédire des évolutions, proposer des options ou réaliser des tâches) font que des applications de l'IA existent déjà pour la plupart des secteurs d'activité des professionnels(les) et qu'elles pourraient modifier profondément certaines pratiques de travail. Une perspective pluridisciplinaire nous permet maintenant de proposer une problématisation de ces effets.

4. Discussion – Une analyse pluridisciplinaire pour problématiser les effets de l'IA dans l'activité des professionnels(les)

Les résultats de l'étude de portée que nous venons de présenter démontrent que l'IA peut avoir des effets profonds (p. ex., disparition de certaines tâches de diagnostic), très variés (dans la relation comme dans le traitement des données ou la prise de décision) et positifs comme négatifs (l'IA pourrait à la fois améliorer la recherche de jurisprudence et renforcer des biais dans les décisions de justice). Une analyse pluridisciplinaire de ces effets multiples permet d'en proposer une problématisation et, de ce fait, des pistes d'actions pour les différents acteurs(trices) dans les organisations.

Tout d'abord, en opposant la théorie de la « primauté actionnariale » (qui considère que l'intérêt des actionnaires est le seul objectif légitime pour l'action d'une entreprise) (Stout, 2012) et la « théorie des parties prenantes » (Freeman, 2010) (qui affirme plutôt qu'une entreprise doit répondre aux besoins de l'ensemble des groupes qui y sont impliqués), les recherches en éthique permettent de poser la question : pour qui l'IA doit être développée? Qui sont le ou les groupes qui peuvent juger de la pertinence des objectifs qui sont donnés à l'IA? Avec la théorie des parties prenantes, les employés(es) font partie des groupes légitimes pour décider des effets de l'IA. Leur activité est donc beaucoup plus souvent inscrite dans la réflexion. Les recherches en éthique permettent également d'identifier des thématiques de la confidentialité et la protection de la vie privée qui, bien que seulement indirectement liées à l'activité de travail, sont centrales dans l'analyse des effets de l'IA.

La réponse à ces questionnements éthiques permet alors de mieux problématiser les effets de l'IA sur le travail, son organisation et la gestion des ressources humaines. Ceux-ci sont caractérisés par la dualité (Meijerink et Bondarouk, 2022; Leonardi, Nardi et Kallinikos, 2012; Orlikowski, 1992). Renouant avec une analyse très présente en ergonomie et en psychologie du travail (Bobillier Chaumon, 2021; Hoc et Darses, 2004), cette problématisation identifie qu'une application donnée de l'IA peut à la fois être une contrainte et une ressource pour une activité de travail donnée. Par exemple, un outil de gestion des connaissances peut simultanément être une contrainte pour des comptables professionnels(les) (son utilisation contraint les employés(es) à révéler des savoirs jusqu'alors cachés) et une ressource pour ces mêmes employés(es) (son utilisation fait gagner en réputation les employés qui ont révélé ces savoirs) (Brivot et Gendron, 2011). Dans le cadre de cette problématisation, il nous a semblé particulièrement utile de développer le point de vue d'un syndicat (la CSN) qui, en tant que représentant des salariés (théorie des parties prenantes), affronte cette dualité dans sa négociation des différents accords autour de l'apparition de l'IA dans le travail des professionnels(les).

Enfin, les résultats de ces délibérations éthiques, ressources humaines et organisationnelles doivent *in fine* s'incarner dans des dimensions techniques, celles des caractéristiques des algorithmes qui fondent l'activité de l'IA dans l'activité de ces professionnels(les). Nous concluons cette discussion problématisée en présentant des recherches en informatique qui mettent en évidence que l'explicabilité est le cœur des débats actuels sur les effets de l'IA dans les activités des professionnels(les).

4.1 Problématisation par la recherche en éthique des affaires : pour quel groupe l'IA est-elle développée? Qui consent volontairement à l'IA?

Les résultats de l'étude de portée que nous venons de présenter montrent que l'IA peut fragiliser certains parcours professionnels, augmenter le stress, manquer de transparence, entraîner une externalisation de la main-d'oeuvre, bouleverser certaines tâches professionnelles dans le domaine du droit ou augmenter la complexité des tâches administratives. Du point de vue de la recherche en éthique des affaires, ces résultats posent la question : quelle est l'obligation morale de l'entreprise devant ce type de conflit : favoriser les intérêts des professionnels(les) ou ceux des propriétaires de l'entreprise?

Ces résultats soulèvent le débat de la bonne gouvernance des entreprises et la priorisation des intérêts des professionnels(les) par rapport à d'autres groupes. En effet, le développement technologique peut amener une entreprise à être plus concurrentielle et à créer plus de valeur économique, tout en ayant un impact négatif sur certains de ses employés(es) (Stout, 2012). En éthique des affaires, la « primauté actionnariale » (ou primauté des propriétaires) est l'approche selon laquelle on doit prioriser les intérêts des actionnaires (ou des propriétaires) de l'entreprise, ce qui revient souvent à faire augmenter sa valeur économique sur le long terme (Stout, 2012; Friedman 1962, 1970). Il existe plusieurs déclinaisons de cette théorie et plusieurs justifications différentes (Hasnas, 1998). On peut défendre cette approche dans le but de protéger la liberté économique et les droits de propriété qui y sont rattachés. Puisque les employés(es) ne sont pas propriétaires de l'entreprise, leurs intérêts ne sont pas prioritaires. On peut aussi défendre l'approche dans le but de favoriser l'efficacité économique si on croit que les entreprises gérées conformément à la primauté actionnariale seront plus concurrentielles ou plus productives.

Appliquée à la question de l'utilisation de l'IA, la norme de la primauté actionnariale accorde peu d'importance aux intérêts des professionnels(les) employés(es) par une entreprise. Les intérêts des actionnaires sont prioritaires et comme ceux-ci ont souvent intérêt à faire augmenter la valeur de l'entreprise sur le long terme, on tiendra compte des intérêts des professionnels(les) seulement s'il y a des conséquences économiques importantes pour l'entreprise. Par exemple, il peut être justifié d'adopter des technologies de l'IA qui augmentent le stress ou bouleversent certaines tâches si cela génère des réductions de coûts importantes.

La primauté actionnariale est la conception dominante dans plusieurs milieux d'affaires et dans les écoles de commerce (Moriarty, 2016). Toutefois, elle fait l'objet de nombreuses critiques et peu d'éthiciens des affaires la défendent de manière absolue. Une approche alternative est la « théorie des parties prenantes » (Freeman, 1994; 2010). Une partie prenante est un groupe d'individus impliqués ou touchés par les activités d'une entreprise, cela inclut les actionnaires ou les propriétaires de l'entreprise, mais aussi ses employés(es), fournisseurs, consommateurs(trices), différents groupes communautaires, syndicaux, ou tout autre groupe dont les intérêts sont affectés par l'entreprise. Selon cette approche, une entreprise doit servir les intérêts de différentes parties prenantes, pas seulement ceux des actionnaires. Par conséquent, on peut prioriser les intérêts des professionnels(les), même si cela ne sert pas directement les intérêts des actionnaires. On peut, par exemple, éviter de mettre en place certaines technologies plus profitables pour éviter de fragiliser certains parcours ou protéger certaines tâches, même si cela n'est pas directement avantageux sur le plan économique.

Toutefois, la théorie des parties prenantes fait aussi l'objet de nombreuses critiques (Heath et Norman, 2004; Boatright, 2006). Le principal problème touche la gestion des conflits entre différents groupes (Martin, 2013). Quels sont les intérêts à prioriser quand les intérêts des employés(es) ou des actionnaires sont en conflit, ou quand il existe des conflits entre différents groupes d'employés(es)? La réponse est simple dans les situations où il existe des solutions gagnant-gagnant, c'est-à-dire qu'on trouve de nouvelles solutions pour servir les intérêts de plusieurs groupes de parties prenantes sans qu'il y ait un impact



économique important sur l'entreprise. Mais les cas difficiles, malheureusement assez fréquents, sont les situations où ces différents intérêts ne concordent pas. Par exemple, un nouveau système automatisé de gestion et surveillance des employés(es) peut faciliter le travail des gestionnaires d'une entreprise et prévenir certaines fraudes, mais cela peut avoir un impact négatif sur d'autres employés(es). Un système d'aide à la décision dans une clinique privée peut améliorer la qualité d'un diagnostic, mais il peut aussi générer des problèmes de transparence, de bonne reddition de comptes, et rendre l'expérience impersonnelle pour un(e) patient(e).

Dans la recherche en éthique des affaires, ce débat sur la pertinence de la théorie des parties prenantes a été enrichi par une réflexion complémentaire sur la légitimité du consentement. Les travaux ont montré que ce n'est pas parce qu'il existe une capacité formelle pour les professionnels(les) de consentir à travailler dans un milieu où on utilise certaines technologies de l'IA qu'il y a pour autant un choix volontaire. L'absence de volontarité du choix peut se manifester de deux façons : premièrement, un(e) professionnel(le) peut ne pas avoir accès à d'autres emplois équivalents où il(elle) ne serait pas contraint(e) d'utiliser ces technologies de l'IA. Il existe de nombreuses situations documentées où l'absence d'accès à l'emploi ou la précarité économique peuvent générer différentes formes d'exploitation même si les individus ne sont pas contraints légalement (Olsaretti, 2004; Werhane et Radin, 2004).

Deuxièmement, un(e) professionnel(le) peut ne pas faire un choix volontaire s'il(elle) subissait des coûts importants en n'utilisant pas ces technologies. On peut penser à des coûts en matière de productivité, de reconnaissance professionnelle ou d'inclusion sociale. Par exemple, il est difficile de fonctionner de manière efficace dans bien des milieux professionnels sans utiliser différents médias sociaux ou différentes suites infonuagiques comme celles de l'entreprise Microsoft ou d'Alphabet/Google, même si l'utilisation de ces logiciels n'est pas obligatoire. Le développement de l'IA et son intégration dans les milieux professionnels peuvent générer des contraintes similaires.

Les recherches en éthique des affaires posent donc des questions centrales qui, en fonction des réponses qui sont données, ont un impact direct sur le travail, l'organisation du travail et la gestion des ressources humaines (GRH) dans les organisations qui utilisent l'IA.

4.2 Problématisation par la GRH : le défi de la dualité de l'IA

Les résultats de l'étude de portée que nous avons présentés montrent que l'IA a des effets nombreux mais très diversifiés. Elle peut conduire à l'amélioration de la réalisation de certaines tâches (p. ex., la recherche automatique dans le contenu de documents), au développement du pouvoir d'agir de certains professionnels (p. ex., prédire et donc parfois prévenir le décrochage de certains élèves) mais aussi à la réduction de celui-ci (p. ex., lorsque l'IA contraint le jugement des professionnels en proposant un nombre limité et hiérarchisé de solutions).

Même après avoir répondu aux questions que l'éthique des affaires pose en préalable, même si on a convenu, par exemple, que l'IA devait répondre aux intérêts des différentes parties prenantes (et pas seulement à ceux des actionnaires) et qu'on a mis en place des protections des employés(es) afin d'obtenir leur consentement le plus éclairé possible, on reste aux prises avec des effets clairement antagoniques de l'IA.

Les recherches en GRH conduisent donc à problématiser les effets de l'IA sous l'angle de la dualité (Meijerink et Bondarouk, 2022). Leçon majeure des anthropologues (Godelier, 1984) et de l'analyse des technologies (Woodward, 1968), très largement présente en ergonomie et en psychologie du travail (Bobillier Chaumon, 2021; Hoc et Darses, 2004), les effets de l'IA sont caractérisés par leur dualité : ils ne

sont pas homogènes, ils varient en fonction des contextes, des usages, des individus et des collectifs. Plus particulièrement, ils peuvent être positifs ou négatifs.

Sur le versant positif, l'IA, pour les professionnels(les), offrirait la possibilité de la prise en charge des tâches répétitives, le plus souvent administratives. Si un(e) professionnel(le) doit faire des actes particulièrement complexes (p. ex., l'annonce d'un diagnostic à un(e) patient(e)), en revanche, une partie de son travail correspond à des tâches plus répétitives et à plus faible valeur ajoutée (p. ex., rédaction de rapports). Il s'agit d'une des caractéristiques fondamentales des bureaucraties professionnelles (Pichault et Nizet, 2000). Les organisations dans lesquelles on retrouve le plus souvent les professionnels(les) mobilisent les savoir-faire hautement qualifiés des professionnels(les), mais sollicitent aussi sur de nombreuses tâches administratives. Comme le montre la recension, l'IA pourrait ainsi prendre en charge des opérations de traitement de textes, de rédaction de comptes rendus ou encore de classement de dossiers.

Au-delà de décharger les professionnels(les) de ces tâches répétitives à faible valeur ajoutée, l'IA pourrait également prendre en charge des tâches pour lesquelles elle a la capacité d'être plus efficace que les professionnels(les) eux-mêmes. Dès qu'il y a beaucoup d'information à analyser, les capacités cognitives des professionnels(les) apparaissent limitées (Simon, 1959). L'IA apparaît comme une technologie qui serait en mesure d'amplifier ce mouvement en leur permettant professionnels(les) d'appuyer leur jugement sur le traitement d'un grand nombre de données qu'eux-mêmes auraient des difficultés à traiter (analyse de photos, identification de jurisprudence, etc.).

Toujours sur le versant positif, il ne faut pas non plus sous-estimer une autre leçon fondamentale de l'analyse des technologies (Orlikowski, 2010) : les technologies n'imposent pas des « impacts » uniformes aux employés(es), au contenu du travail et au fonctionnement des organisations. Les utilisateurs(trices), individuellement mais aussi collectivement, sont capables d'interpréter les technologies et de les recréer à l'occasion. Il est donc parfaitement envisageable que les professionnels(les) soient en mesure de s'approprier les algorithmes de l'IA et les utiliser dans des directions imprévues et, potentiellement, favorables à l'expression de leur professionnalisme.

Du côté des impacts négatifs de l'IA, on peut craindre que, dans les bureaucraties professionnelles, les personnes employées pour réaliser une part du travail administratif des professionnels(les) ne soient progressivement remplacées par des algorithmes d'IA. On peut penser aux applications identifiées dans la recension où des applications permettent de retranscrire des comptes rendus enregistrés verbalement ou effectuer de la traduction simultanée. Des nuances doivent être apportées puisque l'IA ne fait que poursuivre une dynamique engagée depuis longtemps par l'informatique (Orlikowski et Scott, 2016). Cette technologie a permis une automatisation qui a contribué à faire disparaître de nombreux emplois administratifs peu qualifiés tout en créant un petit nombre d'emplois plus qualifiés consacrés au pilotage des dispositifs informatiques automatisés. Mais surtout, les travaux sur l'IA (Levy, 2022) mettent en évidence qu'elle ne semble pas en mesure de remplacer purement et simplement des emplois complets. La technologie s'oriente actuellement plutôt vers l'hybridation où certaines parties de l'activité des professionnels(les) pourraient être remplacées ou renforcées par l'IA, mais l'humain conserverait l'activité centrale de mettre en lien et de construire le sens de ces différentes activités (Levy, 2022). On se trouverait alors dans un schéma relativement classique en matière de technologie : analyser les relations entre l'IA et l'activité des professionnels(les) nécessiterait de comprendre comment les différentes formes de cette technologie s'insèrent dans leurs activités de travail. Toutefois, parmi les impacts plus complexes de l'IA, on trouve des dimensions spécifiques liées à la capacité de jugement des professionnels(les). Les algorithmes d'IA traitent un très grand nombre de données et présentent ensuite des résultats agrégés aux professionnels(les). Les dynamiques sont diverses et les algorithmes peuvent « évaluer » différentes options, en « recommander » certaines ou les « classer » (Kellogg, 2020). Ces applications de l'IA sont celles qui sont



le plus en rupture avec les technologies passées puisqu'elles qu'elles concourent à une finalité : influencer la prise de décision du ou de la professionnel(le). Ainsi, si l'option de l'hybridation se confirmait, l'IA impliquerait une réflexion particulière sur la question de la construction du jugement des professionnels(les). La prochaine section discute d'ailleurs plus en détail de cet enjeu du jugement professionnel.

4.3 Problématisation par la GRH : le défi de la dualité de l'IA, le point de vue d'un syndicat

Lorsque la perspective syndicale analyse l'impact de l'IA sur les conditions de travail et d'emploi des professionnels(les), la première idée qui émerge est qu'il ne s'agit pas de la première fois qu'une technologie vienne bouleverser le travail et les travailleurs(euses). Depuis plus d'un siècle, le monde du travail a subi diverses transformations technologiques. Au fil des époques, elles ont bouleversé la relation des travailleurs(euses) avec leur métier. Nous pouvons penser à l'adaptation de l'artisan qui s'est vu dépossédé de son autonomie professionnelle par les chaînes de montage; de l'ouvrier par l'automatisation et la robotisation des usines ou du personnel de bureau par l'informatisation des tâches. Si ces technologies ont parfois eu des effets bénéfiques en éliminant certaines tâches plus pénibles, elles ont aussi provoqué des mises à pied à l'avantage des employeurs qui ont ainsi réduit leur coût de main-d'œuvre.

Ces transformations technologiques ont généralement aussi ébranlé le sens que les travailleurs donnent à leurs tâches : le produit final leur échappant des mains. Au cours des 100 ans de la CSN, nous avons été des témoins privilégiés des différents bouleversements dans les milieux de travail. Les syndicats affiliés à la CSN mettent maintenant de l'avant diverses clauses de conventions collectives protégeant les travailleurs(euses) contre les préjudices causés par l'arrivée de nouvelles technologies, mais aussi, de manière plus proactive, des clauses portant sur leur participation aux décisions sur les changements à apporter à l'organisation du travail. Il est donc essentiel que les travailleurs(euses) soient informés(es) en amont de l'implantation des technologies issues de l'IA.

L'intelligence artificielle pourrait aussi transformer l'organisation du travail de l'entreprise. La participation des travailleurs(euses) est alors nécessaire pour que la transformation se déroule de manière respectueuse et optimale. Par ailleurs, l'IA utilise souvent une somme importante de données qui pourraient poser des enjeux de confidentialité (p. ex., données de productivité, données biologiques). Finalement, pour accomplir le travail intégrant l'IA, les travailleurs(euses) auront besoin d'être formés(es). La complexité de l'IA demande un temps de formation souvent plus long que celui accordé pour les technologies implantées antérieurement. Cette formation du personnel est un incontournable pour comprendre les tenants et aboutissants de l'IA.

L'IA est une nouvelle étape technologique qui n'est pas étrangère aux précédentes en ce sens où elle provoque un changement dans la relation travailleurs(euses)–travail. Une des différences se trouve ici dans le fait que l'IA est une technologie qui peut solliciter particulièrement les professionnels(les) en raison de leur jugement professionnel basé par leur formation spécifique et leur cumul d'expérience. Ce jugement est aussi encadré par des ordres professionnels. L'arrivée de l'IA fait craindre à plusieurs une perte de la plus-value de leur jugement. Par exemple, un comptable évaluant les états financiers d'une entreprise pour l'émission d'un prêt pourra-t-il toujours ne se fier qu'à son expertise ou devra-t-il se plier aux conclusions d'une application de l'IA d'aide à la décision? Si l'analyse professionnelle est différente de celle de l'IA, sera-t-il toujours protégé par son ordre? Les dernières innovations informatiques avaient pour but d'appuyer les professionnels(les), et non pas de leur indiquer la marche à suivre. La définition du rôle-conseil devra être ainsi repensée.

Les professionnels(les) perdront de leur autonomie s'ils n'en viennent qu'à exécuter des conclusions tierces. Ils se trouveront en conflit de loyauté entre leurs valeurs, leur professionnalisme et le devoir d'appliquer une décision dite rationnelle issue de l'IA. Déjà, certains logiciels sont utilisés pour soutenir les professions en relations d'aide. Par exemple, pour l'évaluation des besoins en soins à domicile par des personnes travailleuses sociales, ces dernières ont des profils de personnalité orientés vers le contact humain et la communication. Quelle sera la satisfaction des professionnels(les) lorsqu'un « froid » logiciel pourra décider à leur place des services à offrir? Il existe également des logiciels permettant le diagnostic psychologique. Comment se fier à des algorithmes alors que le domaine de la psychologie est traversé de différents courants de pensée et où le jugement subjectif est un atout essentiel? Les différentes écoles de pensées pourront-elles encore exister dans cet univers? Avec l'arrivée de l'IA, il faudra donc trouver un juste équilibre entre la logique rationnelle qu'elle propose et le respect du jugement professionnel.

4.4 Problématisation par la recherche en informatique : le débat autour de l'explicabilité

Les résultats présentés dans le cadre de notre étude de portée, comme les axes de problématisation discutés ci-dessus, constituent autant d'enjeux pour les recherches en informatique. En effet, le défi est que les choix qui seront faits en matière d'éthique, de RH ou d'organisation du travail parviennent finalement à s'incarner dans la technique informatique, dans les algorithmes qui seront développés, entraînés et implémentés.

En particulier, les enjeux identifiés ci-dessus peuvent être problématisés sous l'angle de l'explicabilité. Contrairement à la programmation « standard » où la personne qui écrit le code décrit explicitement dans celui-ci les étapes de l'algorithme que la machine doit suivre pour résoudre un problème particulier, l'objectif de l'apprentissage machine est de permettre à l'ordinateur d'apprendre à résoudre un problème à partir d'exemples (Murphy, 2021). Plus précisément, lors de la « phase d'entraînement », l'algorithme d'apprentissage prend en entrée un ensemble de données d'entraînement (p. ex., des profils de personnes avec le salaire associé) et produit en sortie un modèle d'apprentissage. Ce modèle pourra être ensuite utilisé lors de la « phase d'inférence » pour faire des prédictions sur de nouvelles données (p. ex., dans le cadre du recrutement prédictif en proposant un salaire à l'embauche à un profil particulier).

Un des principaux enjeux reliés à l'utilisation de l'apprentissage machine dans des systèmes d'aide à la décision est l'explicabilité. Cet enjeu émerge principalement lorsque les prédictions de modèles sont utilisées pour prendre des décisions qui impactent des êtres humains, et donc qu'on souhaiterait pouvoir fournir une explication sur le pourquoi de cette décision. Ainsi, lorsqu'un algorithme présente des options, qu'il a possiblement ordonnées, à un(e) professionnel(le) pour le guider dans sa décision, il transforme son activité professionnelle donc il rencontre l'enjeu de l'explicabilité.

Dans l'explicabilité (Lipton, 2018), le défi provient du fait que certains modèles d'apprentissage, tels que les réseaux de neurones profonds (qu'on vulgarise souvent sous le terme de « réseaux de neurones »), sont très performants sur le plan de la précision de leurs prédictions, mais ont une structure trop complexe pour pouvoir facilement expliquer leurs prédictions, ce qui amène parfois à les considérer comme des « boîtes noires ». Cette opacité rend difficile la possibilité pour un individu de comprendre comment une prédiction particulière a été développée, et donc limite aussi la possibilité de remettre en cause cette décision ou encore de valider que le modèle fonctionne effectivement comme attendu (Pasquale, 2016). Une des solutions à cet enjeu est de préférer l'utilisation de modèles interprétables, comme les arbres de décision ou les modèles linéaires, qui sont potentiellement un peu moins performants que les réseaux profonds (modèles « boîtes noires ») mais dont la structure permet d'expliquer directement la prédiction.

La notion d'explicabilité est sujette à discussion pour ce qui est de sa définition, mais est souvent exprimée par des critères tels que la simulabilité ou la décomposabilité (Lipton, 2018). Ainsi, la simulabilité correspond à la possibilité pour un modèle d'être décrit d'une manière concise qui permettrait à un humain



de calculer lui-même la prédiction alors que la composabilité fait référence au fait que tous les composants du modèle (p. ex., attributs, antécédents d'une liste de règles ou poids de modèles linéaires) sont compréhensibles par un humain.

L'autre approche possible à l'enjeu d'explicabilité est d'expliquer *a posteriori* la prédiction du modèle « boîte noire » par des méthodes d'explication post-hoc (Guidotti, 2018). Par exemple, les explications globales cherchent à expliquer la logique complète du modèle boîte noire par l'entraînement d'un modèle substitut qui est interprétable par conception tout en maximisant sa fidélité avec les prédictions de la boîte noire. Autrement dit, l'objectif est que le modèle substitut réalise des prédictions très fidèles au modèle boîte noire, mais pouvant facilement les expliquer du fait que ce modèle est interprétable. Il existe de nombreuses autres méthodes d'explication post-hoc, dont certaines visent à quantifier le poids de chacun des attributs d'entrée dans la prédiction (autrement dit à caractériser l'importance qu'un attribut particulier revêt dans la prédiction faite par le modèle) ou encore à suggérer un contrefactuel, qui est un profil proche du profil original donné au modèle en matière de similarité, mais qui aurait reçu une prédiction différente.

Avec l'utilisation d'algorithmes dans différentes organisations (p. ex., pour le recrutement chez Amazon), on a de plus constaté que les réseaux profonds ont souvent tendance à reproduire et accentuer les biais inconscients et les phénomènes de « discrimination » présents dans les données d'entraînement (Lacroux et Martin-Lacroux, 2021). Le processus est alors le suivant. Pour des raisons historiques et sociétales, les données d'entraînement sont biaisées de manière inhérente envers un groupe de la population (p. ex., les données qui ont été utilisées par Amazon pour entraîner leur algorithme de recrutement sous-représentaient très fortement les femmes). Le modèle appris par l'algorithme d'apprentissage reproduit et souvent amplifie ce biais négatif. Il risque alors de défavoriser de manière importante ce groupe dans les prédictions effectuées.

L'approche naïve de ce problème consiste à considérer qu'il suffit d'enlever dans l'ensemble des données l'attribut sensible qui peut amener à cette discrimination directe. D'une part, c'est souvent difficile parce que, par exemple, dans un CV, les indices du genre se trouvent dans de très nombreux éléments de nature très variée. Mais, d'autre part, la discrimination indirecte reste toujours possible à cause des corrélations qui existent entre l'attribut sensible et les autres attributs. Afin de remédier à cela, plusieurs notions d'équité ont été développées dans les dernières années qui peuvent être prises en compte dans l'entraînement d'un modèle d'apprentissage (Narayanan, 2018). Ainsi, « l'équité de groupe » cherche à produire des modèles d'apprentissage approximant la parité pour une certaine mesure statistique parmi les sous-groupes de la population définis par l'attribut sensible. Une autre approche, se basant sur « l'équité individuelle », requiert que le modèle produise la même prédiction pour des profils qui sont similaires (p. ex., pour deux profils qui ne diffèrent qu'en matière de l'attribut sensible).

5. Limites et études futures

Bien que l'étude de portée soit une façon pertinente de montrer comment les connaissances se construisent autour d'une question, elle est tributaire de la documentation sur laquelle elle porte. Ainsi, malgré un effort d'exhaustivité, il est possible que de la documentation n'ait pas été identifiée dans le processus en raison des bases et sources de données utilisées ou de la langue de publication. Ainsi, des applications, enjeux ou impacts sur le terrain sont peut-être sous-estimés. Ceci montre l'importance de conduire davantage d'études empiriques quant aux enjeux et impacts soulevés afin d'offrir des pistes de réponses davantage ancrées dans les milieux professionnels. Ces recherches sont importantes pour aller au-delà des questionnements en documentant, par exemple, comment le travail des professionnels(les) est modifié lorsqu'une application de l'IA est introduite dans un milieu ou une profession. Une plus étroite collaboration entre développeurs(euses), utilisateurs(trices) et chercheurs(euses) est souhaitable pour répondre aux questions soulevées et prévenir les effets négatifs anticipés.

Conclusion

Les effets que les applications de l'IA ont et pourraient avoir sur l'activité des professionnels(les) commencent à être de mieux en mieux connus. L'étude de portée que nous avons présentée permet de repérer les différentes finalités de ces applications ainsi que les enjeux et les risques qu'elles comportent. On commence même à disposer d'études qui identifient les impacts de l'IA sur le bien-être des employés(es) (Pereira et al., 2021). Une analyse pluridisciplinaire nous a permis de problématiser ces résultats. Au terme de cet article, il apparaît que, bien qu'elle soulève des préoccupations importantes chez les professionnels(les) en raison d'aspects sensibles du travail qu'elle est susceptible d'influencer, l'IA appelle une réflexion beaucoup plus large sur la conception même du travail des professionnels(les) et de leur rapport à la tâche, à l'utilisateur et à leur profession.

RÉFÉRENCES

- Àgnes, J. S. (2022). Gaining and Training a Digital Colleague: Employee Responses to Robotization. *The Journal of Applied Behavioral Science*, 58(1), 29–64. <https://doi.org/10.1177/00218863211043596>
- Abord de Chatillon, E. et Desmarais C. (2012). Le nouveau management public est-il pathogène? *Management international*, 16 (3), 10-24. <https://id.erudit.org/iderudit/1011413ar>
- Agrawal, A., Gans, J., & Goldfarb A. (2018). *Prediction Machines: The Simple Economics of Artificial Intelligence*, Boston, Massachusetts: Harvard Business Review Press.
- Arksey, H. et O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International journal of social research methodology*, 8 (1), 19-32. <https://doi.org/10.1080/1364557032000119616>
- Boatright, J.R. (2006). What's Wrong—And What's Right—With Stakeholder Management. *Journal of Private Enterprise*, 21(2), 106–130.
- Bobillier-Chaumon, M.-E. (2021). *Digital Transformations in the Challenge of Activity and Work: Understanding and Supporting Technological Changes*. Oxford : STE-Wiley
- Brivot, M. et Gendron, Y. (2011). Beyond panopticism: On the ramifications of surveillance in a contemporary professional setting. *Accounting organizations and society*, 36(3): 135-155. <https://doi.org/10.1016/j.aos.2011.03.003>
- Brynjolfsson, E., Mitchell, T. et Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, 108, 43-47.
- Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1-12. DOI: 10.1177/2053951715622512
- Cardon, D., Cointet, J.P. et Mazières, A. (2018). La revanche des neurons. *Réseaux*, 5(211), 173-220. <https://doi.org/10.3917/res.211.0173>
- Carlini, N., Liu, C., Erlingsson, U., Kos, J. et Song, D. (2019). The secret sharer: Evaluating and testing unintended memorization in neural networks. 28th USENIX Security Symposium, 267-284.
- Clot, Y., & Leplat, J. (2005). La méthode clinique en ergonomie et en psychologie du travail. *Le travail humain*, 68(4), 289-316. <https://doi.org/10.3917/th.684.0289>
- Freeman, R. E. (1994). *A Stakeholder Theory of the Modern Corporation*. In Beauchamp, T.L. et Bowie N.E. (dir). Ethical theory and business, Upper Saddle River: Pearson: Prentice Hall, 66–76. Freeman, R. E., Harrison, J.S., Wicks A.C., Parmar B. & de Colle S. (2010). *Stakeholder Theory, The State of the Art*. Cambridge: Cambridge University Press. <https://doi.org/10.1080/19416520.2010.495581>
- Freeman, R. E., Harrison, J.S., Wicks A.C., Parmar B. et de Colle S. (2010). *Stakeholder Theory, The State of the Art*. Cambridge: Cambridge University Press. <https://doi.org/10.1080/19416520.2010.495581>
- Fréour, L., Pohl, S. et Battistelli A., (2021). How Digital Technologies Modify The Work Characteristics: A Preliminary Study. *The Spanish Journal of Psychology*, 24, e14. <https://doi.org/10.1017/SJP.2021.12>



- Freyssenet, M. (1992). Processus et formes sociales d'automatisation Le paradigme sociologique. *Sociologie du travail*, 34(4), 469-496. <https://doi.org/10.3406/sotra.1992.2611>
- Friedman, M. (1962) *Capitalism and freedom*. Chicago: University of Chicago Press.
- Friedman, M. (1970). The Social Responsibility of Business Is to Increase Its Profits, *The New York Times Magazine*.
- Gherardi, S. et Landri, P. (2014). "I Sign, Therefore I am" (Un)stable Traces of Professional Practices. *Professions and Professionalism*, 4(2), 1-11. <https://doi.org/10.7577/pp.618>
- Godelier, M. (1984). *L'idéal et le matériel*. Paris: Fayard.
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F. et Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM computing surveys (CSUR)*, 51(5), 1-42. <https://doi.org/10.1145/3236009>
- Hasnas, J. (1998). The normative theories of business ethics: A guide for the perplexed. *Business Ethics Quarterly*, 8(1), 19-42. <https://doi.org/10.2307/3857520>
- Heath, J. et Norman W. (2004). Stakeholder Theory, Corporate Governance and Public Management: What can the History of State-Run Enterprises Teach us in the Post-Enron era? *Journal of Business Ethics*, 53, 247-265.
- Hoc, J.M. et Darses, F. (Eds.) (2004). *Psychologie ergonomique : tendances actuelles*. Paris : PUF, Collection Le Travail Humain. <https://doi.org/10.4000/activites.1635>
- Kellogg, K.C., Melissa A.V. et Christin A. (2020). Algorithms at work: the new contested terrain of control. *Academy of Management Annals*, 14(1), 366-410. <https://doi.org/10.5465/annals.2018.0174>
- Kohler, D. et Weisz J.D. (2016) Industrie 4.0: comment caractériser cette quatrième révolution industrielle et ses enjeux? *Annales des Mines-Réalités industrielles*, 4, 51-56. <https://doi.org/10.3917/rindu1.164.0051>
- Kuiper, M. (2018). Connective Routines: How Medical Professionals Work with Safety Checklists. *Professions and Professionalism*, 8(1), e2251. <https://doi.org/10.7577/pp.2251>
- Lacroux, A. et Martin-Lacroux C. (2021). L'Intelligence artificielle au service de la lutte contre les discriminations dans le recrutement: nouvelles promesses et nouveaux risques. *Management Avenir*, 2(122), 121-142.
- Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I. et Kasper, G. (2019). The challenges of algorithm-based HR decision-making for personal integrity. *Journal of Business Ethics*, 160(2), 377-392. <https://doi.org/10.1007/s10551-019-04204-w>
- Leonardi, P. M., Nardi, B. A. et Kallinikos, J. (Eds.). (2012). *Materiality and organizing: Social interaction in a technological world*. Oxford University Press.
- Levy, K.e.c. (2022). *Data Driven: Truckers, Technology, and the New Workplace Surveillance*. Princeton: Princeton University Press.
- Lipton, Z.C. (2018). The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue*, 16(3), 31-57.
- Malcolm, I. et Plowman, L. (2014). Knowledge, Technology and the Professional Learning of Localisers. *Professions and Professionalism*, 4(1), 1-15.
- Martin, D. (2013). The Contained-Rivalry Requirement and a "Triple Feature" Program for Business Ethics. *Journal of Business Ethics*, 115(1), 167-182. <https://doi.org/10.1007/s10551-012-1369-4>
- Meijerink, J. et Bondarouk, T. (2023). The duality of algorithmic management: Toward a research agenda on HRM algorithms, autonomy and value creation. *Human resource management review*, 33(1), 100876.
- Mintzberg, H. (1989). *Le management : voyage au centre des organisations*. Paris: Éditions d'Organisation.
- Mira, J. M. (2008). Symbols versus connections: 50 years of artificial intelligence. *Neurocomputing*, 71(4-6), 671-680.
- Moriarty, J. (2016). Business Ethics. In Zalta E.N. (Ed.) *The Stanford Encyclopedia of Philosophy*. Stanford: Metaphysics Research Lab, Stanford University.

- Murphy, K.P. (2021). *Probabilistic Machine Learning: An introduction*. Boston: MIT Press.
- Musk, E. (2014). I hope artificial intelligence is nice to us. *New Perspectives Quarterly*, 31(1), 51-55. <https://doi.org/10.1111/npqu.11427>
- Narayanan, A. (2018). Translation tutorial: 21 fairness definitions and their politics. Conference *Fairness Accountability Transparency*, New York.
- Office des professions du Québec (2021). *Code des professions*, tiré de <http://legisquebec.gouv.qc.ca/fr/document/lc/c-26>
- Olsaretti, S. (2004). *Liberty, desert and the market: A philosophical study*, Cambridge: Cambridge University Press.
- Orlikowski W.J. et Scott S.V. (2016). *Digital work: a research agenda*. In Czarniawska, B. (Ed), *A research agenda for management and organization studies*, Cambridge: Edward Elgar Publishing, p. 88-95.
- Orlikowski, W.J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organization science*, 3(3), 398-427. <https://doi.org/10.1287/orsc.3.3.398>
- Orlikowski, W.J. (2010). The sociomateriality of organisational life: considering technology in management research. *Cambridge Journal of Economics*, 34(1), 125-141. <https://doi.org/10.1093/cje/bep058>
- Pasquale, F. (2016) *The Black Box Society, The Secret Algorithms That Control Money and Information*. Harvard: Harvard University Press.
- Pereira, V., Hadjielias, E., Christofi, M. et Vrontis, D. (2021). A systematic literature review on the impact of artificial intelligence on workplace outcomes: A multi-process perspective. *Human Resource Management Review*, 33, 100857. <https://doi.org/10.1016/j.hrmr.2021.100857>
- Pfadenhauer, M. et Kirschner, H. (2017). From Dyad to Triad: Mediatization and Emerging Risks for Professional Autonomy. *Professions and Professionalism*, 7(1), e1646. <https://doi.org/10.7577/pp.1646>
- Pichault, F. et Nizet J. (2000). *Les pratiques de gestion des ressources humaines*. Paris: Seuil.
- Rabardel, P. (2005). *Instrument, activité et développement du pouvoir d'agir*. In Lorino P. et Teulier R. (eds). *Entre connaissance et organisation: l'activité collective*. Paris : La découverte, p. 251-265. <http://www.cairn.info/entre-connaissance-et-organisation-l-activite-coll--page-251.htm>
- Rissanen, O., Pitkänen, P., Juvonen, A., Rähä, P., Kuhn, G. et Hakkarainen, K. (2017). How Has the Emergence of Digital Culture Affected Professional Magic? *Professions and Professionalism*, 7(3), e1957. <https://doi.org/10.7577/pp.1957>
- Roll, M. et Ifenthaler D. (2021). Learning Factories 4.0 in technical vocational schools: can they foster competence development? *Empirical Research in Vocational Education and Training*, 13(1), 1-23. <https://doi.org/10.1186/s40461-021-00124-0>
- Schaeffer, E. (2017). *Industry X. 0: Realizing digital value in industrial sectors*. London: Kogan Page Publishers.
- Shokri, R., Stronati, M., Song, C., & Shmatikov, V. (2017). Membership inference attacks against machine learning models. *IEEE Symposium on Security and Privacy (SP)*, 3-18. <https://doi.org/10.1109/SP.2017.41>
- Simon, H.A. (1959). Theories of Decision Making in Economics and Behavioral Science. *The American Economic Review*, 49(3), 223-283.
- Stout, L. (2012). *The shareholder value myth: How putting shareholders first harms investors, corporations, and the public*. San Francisco: Berrett-Koehler Publishers.
- Tambe, P., Cappelli P. et Yakubovich, V (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review*, 61 (4), 15-42. <https://doi.org/10.1177/000812561986791>
- Trottier, M. et Comeau-Vallée, M. *Intelligence artificielle et travail : les enjeux organisationnels et managériaux*. 87^e Congrès de l'ACFAS, Ottawa, Canada
- Werhane, P.H. et Radin T.J. (2004). Employment at Will and Due Process. In Beauchamp, T.L. et Bowie N.E. (Eds), *Ethical theory and business*. 7^e éd, Upper Saddle River: Pearson: Prentice Hall.
- Woodward J. (1965). *Industrial Organization: Theory and Practice*, Oxford : Oxford University Press.