RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

anact
agence nationale pour l'amélioration des conditions de travail

OCTOBRE 2026

N° 18

La Revue des conditions de travail

TRAVAILLER SUR ET AVEC L'IA : QUELS EFFETS SUR LE TRAVAIL ET LES TRAVAILLEUSES ET TRAVAILLEURS ?

Jean-Noël Amato
Flore Barcellini
Ranya Bennani
Janine Berg
Marie Bernard
Marc-Éric Bobillier Chaumon
Ludovic Bugand
Priscilla Capovilla
Odile Chagny
Maxime Cornet
Cyril Cosme
Corinne Dubois
Yahya Fallah
Yann Ferguson
Pierre Flandrin
Éric Fourboul
Myriam Fréjus
Tamari Gamkrelidze
Marion Gilles
Nayla Glaise
Pawel Gmyrek
Pauline Gourlet
Olivier Grisvard
Zacharie Hellouin de Cenival
Elisabeth Leblanc
Clément Le Ludec
Denis Lemaître
Vincent Mandinaud
Philippe Marcadent
Luca Mourgaud
Anne-Marie Nicot
Grégory Plançon
Aída Ponce del Castillo
Maria Sol Romero Goldar
Benoit Rottembourg
Gauthier Roussilhe
Franca Salis-Madinier
Christophe Teissier
Nicolas Vispi
Elisabetta Zibetti
Moustafa Zouinar

Coordination scientifique : Caroline Datchary et Vincent Mandinaud

OCTOBRE 2026

N° 18

La Revue des conditions de travail

TRAVAILLER SUR ET AVEC L'IA : QUELS EFFETS SUR LE TRAVAIL ET LES TRAVAILLEUSES ET TRAVAILLEURS ?

anact

Destinée à un public de chercheurs, de praticiens et de consultants, *La Revue des conditions de travail* ambitionne d'approfondir et de renouveler le regard porté sur les évolutions du travail.

La Revue propose de cheminer sur deux voies :

- La première empruntera le très riche matériau issu des expérimentations menées par l'Anact et des consultants en entreprises, comme autant d'expériences originales de conduite concertée des transformations du travail.
- La seconde prendra appui sur des contributions scientifiques issues des nombreuses disciplines convoquées par le travail et les conditions de son exercice.

Le pari de *La Revue des conditions de travail* est de faire discuter les constats singuliers avec les problématiques générales, les acquis théoriques avec les connaissances issues de la pratique… Face à la dispersion des savoirs, mais aussi devant la nécessité d'en créer de nouveaux, la rencontre de ces deux voies est seule à même de dépasser la déploration des dysfonctionnements et de proposer des pistes d'amélioration légitimes et pertinentes.

La Revue des conditions de travail s'adresse deux fois par an à une large communauté de chercheurs, praticiens et intervenants engagés dans la compréhension des enjeux et des modes d'action associés à la transformation du travail et l'amélioration de ses conditions d'exercice.

Pour cela, la Revue se veut :

- Un **creuset** alimenté par les travaux de capitalisation issus de l'expérimentation de l'Anact, les études, mais aussi par des productions — enquêtes, travaux de chercheurs et de consultants — de différents milieux.
- Un **incubateur d'idées** propice à la controverse et à la mise en débat.
- Un **espace pluridisciplinaire** et **interinstitutionnel**, nourri par des apports issus de la psychologie du travail, la sociologie, les sciences de gestion, l'ergonomie, etc.

L'objectif ultime de *La Revue des conditions de travail* consiste à proposer des pistes de réflexion sur ce qui, aujourd'hui, conditionne la qualité du travail et de son environnement. En effet, le travail est traversé par de nombreuses controverses sur son devenir et son organisation. Ces controverses sont structurantes. Il est important d'en faire état pour s'approcher au plus près des réalités de terrain et examiner comment les acteurs s'emparent des questions du travail et élaborent des solutions satisfaisantes. L'enjeu est d'importance, tant pour l'avenir des entreprises que pour la santé et les conditions de travail des travailleuses et travailleurs.

Nota : Le contenu des articles et les propos relatés n'engagent la responsabilité que de leurs auteurs.

SOMMAIRE

{avant-propos}

6 TRAVAILLER SUR ET AVEC L'IA : QUELS EFFETS SUR LE TRAVAIL ET LES TRAVAILLEUSES ET TRAVAILLEURS ?
Caroline Datchary, Marion Gilles, Vincent Mandinaud

{Les IA en situation : recomposition des activités et des métiers}

16 SYSTÈME D'IA CONVERSATIONNEL POUR LA RELATION DE SERVICE DANS L'HÔTELLERIE : PISTES DE RÉFLEXION ET DE RÉPONSES « NON AUTOMATISÉES »
Pierre Flandrin

25 PLATEFORMES NUMÉRIQUES DE LIVRAISON DE REPAS : QUAND LE MANAGEUR DEVIENT UNE MACHINE
Anne-Marie Nicot

36 LA LIVRAISON DE REPAS : LE JEU ALGORITHMIQUE DES PLATEFORMES
Elisabeth Leblanc

46 LA RADIOLOGIE À L'ÈRE DES « NOUVEAUX » SYSTÈMES D'INTELLIGENCE ARTIFICIELLE : USAGES ET TRANSFORMATIONS DU TRAVAIL
Tamari Gamkrelidze, Moustafa Zouinar, Flore Barcellini

60 DES IA EXPLICABLES POUR DES TECHNOLOGIES ACCEPTABLES ET SOUTENABLES POUR LE TRAVAIL : ENJEUX ET CONTOURS DE L'EXPLICATION DANS ET POUR L'ACTIVITÉ
Marc-Eric Bobillier Chaumon, Ranya Bennani, Luca Mourgaud, Myriam Frejus, Denis Lemaître, Olivier Grisvard

{Travailler les IA : enquêter, concevoir, discuter, réguler}

73 PARTICIPER À L'IA AUTREMENT : RETROUVER PRISE SUR UN OBJET FUYANT PAR L'ENQUÊTE COLLECTIVE
Pauline Gourlet

87 CONCEVOIR DES SYSTÈMES D'IA ATTENTIFS AUX CONDITIONS DE TRAVAIL : RETOUR RÉFLEXIF SUR UNE DÉMARCHE PARTICIPATIVE ET INTERDISCIPLINAIRE
Maria Sol Romero Goldar, Zacharie Hellouin de Cenival

96 POURQUOI LE DIALOGUE SOCIAL TECHNOLOGIQUE NE PEUT-IL PAS FAIRE DE MIRACLES À L'HEURE DES IAS ? RETOUR SUR UNE SÉRIE DE PROJETS PORTÉS, SOUTENUS OU ACCOMPAGNÉS PAR L'ANACT
Vincent Mandinaud, Ludovic Bugand, Nicolas Vispi, Grégory Plançon

105 INTELLIGENCE ARTIFICIELLE ET PRÉVENTION DES RISQUES PSYCHOSOCIAUX : UNE DÉMARCHE RÉFLEXIVE DE NANTES MÉTROPOLE
Marie Bernard, Priscilla Capovilla

112 DIALOGUER SUR L'IA ? RETOURS SUR UNE EXPÉRIMENTATION ORGANISATIONNELLE COLLABORATIVE : LE PROJET DIAL-IA
Odile Chagny, Christophe Teissier

{IA, travail et emploi : Trajectoires annoncées et effets discutés}

121 L'IA GÉNÉRATIVE, L'EMPLOI ET LES CONDITIONS DE TRAVAIL : QUELQUES CONSIDÉRATIONS POUR LES DÉCIDEURS PUBLICS, LES PARTENAIRES SOCIAUX ET LES TRAVAILLEURS
Janine Berg, Pawel Gmyrek

132 APPRÉHENDER L'IMPACT DE L'IA SUR LE TRAVAIL À TRAVERS UNE APPROCHE PAR LES CHAÎNES DE VALEURS GLOBALES. LES CONSÉQUENCES DE LA DIVISION INTERNATIONALE DU TRAVAIL DE L'IA SUR LA PRÉCARISATION DU TRAVAIL
Clément Le Ludec, Maxime Cornet

142 L'IA GÉNÉRATIVE AU TRAVAIL : QUELLES TRANSFORMATIONS À L'HORIZON ?
Vincent Mandinaud, Yann Ferguson, Aída Ponce del Castillo

153 L'IA À LA RÉUNION : PERCEPTIONS ET STRATÉGIES DES PARTENAIRES SOCIAUX
Jean-Noël Amato, Corinne Dubois, Elisabetta Zibetti

{Entretiens}

Points de vue experts et institutionnels : fabriquer, réguler et évaluer les impacts écologiques

164 DE LA TRANSPARENCE À LA RÉGULATION DES ALGORITHMES
Entretien avec Benoit Rottembourg

171 UNE CONVENTION DE L'OIT POUR RÉGULER LE TRAVAIL DES PLATEFORMES
Entretien avec Cyril Cosme et Philippe Marcadent

179 QUELLES MÉTHODES POUR MESURER LES IMPACTS DE L'IA ?
Entretien avec Eric Fourboul et Gauthier Roussilhe

Points de vue de partenaires sociaux : quelles conditions pour un progrès économique et social ?

187 LA NÉCESSITÉ D'UNE RÉGULATION DE L'IA AU TRAVAIL
Entretien avec Franca Salis-Madinier

193 ENCADRER L'IA AU TRAVAIL PAR UNE DIRECTIVE EUROPÉENNE
Entretien avec Nayla Glaise

200 FAIRE DE L'IA, UN LEVIER D'INNOVATION POUR LES ENTREPRISES
Entretien avec Yahya Fallah

{avant-propos}

TRAVAILLER SUR ET AVEC L'IA : QUELS EFFETS SUR LE TRAVAIL ET LES TRAVAILLEUSES ET TRAVAILLEURS ?

Caroline Datchary (Université de Toulouse Jean-Jaurès, LISST-CNRS), **Marion Gilles** (Anact), **Vincent Mandinaud** (Anact)

La notion d'*intelligence artificielle* (IA) a été inaugurée en 1956 par un groupe d'informaticiens américains ; elle n'est donc pas nouvelle. Depuis cette date néanmoins, ses contours ont évolué, avec une visibilité variable selon les époques, aussi bien sur le plan technique que dans les discours, les représentations et les usages qui lui sont associés. Et si l'IA est, depuis une vingtaine d'années, sur le devant de la scène médiatique, politique, scientifique, mais aussi professionnelle et domestique, il n'existe pas pour autant de définition commune. Un consensus émerge toutefois dans la sphère scientifique considérant que le terme d'« IA » laisse dans l'ombre la diversité qu'elle recouvre : il n'y a pas *une* IA mais *des* IA, tout comme il n'y a pas « un » numérique, mais une diversité d'outils, de dispositifs et d'usages numériques. Relevant d'une informatique avancée, l'IA rassemble en effet une grande variété d'outils, de machines, d'algorithmes, de données, d'approches, mobilisés pour des finalités variées. Les IA endossent également des fonctions multiples : la recherche, l'analyse d'informations, les prédictions, les diagnostics, les recommandations, l'exécution dite « autonome » de tâches ou encore la génération de contenus nouveaux. Raisonner sur « l'IA » au singulier comporterait dès lors un risque de déterminisme technique : celui de considérer que la technologie porterait en elle-même des effets sur le travail, les emplois ou les compétences, indépendamment des choix organisationnels, économiques et sociaux qui accompagnent son introduction. Une telle approche conduit à considérer les transformations du travail uniquement comme des conséquences quasi mécaniques des évolutions techniques, alors qu'elles résultent aussi de décisions, de rapports de force, de régulations et de manières très diverses de concevoir ces technologies et de les intégrer dans les situations de travail. Ainsi les dénominations et définitions de l'IA ne sont-elles pas neutres : elles sélectionnent certaines capacités techniques, invisibilisent d'autres dimensions – notamment le travail humain nécessaire à leur conception, leur entrainement et leur maintenance – et orientent ainsi les attentes, les politiques publiques et les modalités de leur déploiement. Ces enjeux de définition sont donc cruciaux : selon la manière de définir les IA, s'engagent des représentations différentes de la technologie, du travail, des organisations et des populations qui participent à leur production ou à leur usage. Aussi, les débats sur les IA instruisent autant sur ces technologies que sur la manière dont elles sont conçues, appréhendées, utilisées et projetées. Par exemple, les débats des années 2010 insistaient sur l'automatisation des tâches répétitives, la suppression massive des emplois, l'avènement d'une nouvelle division du travail (aux IA reviendraient exécution, gestion, organisation, et à l'humain création, imagination,

coopération et émotions). À travers ces débats et ces pratiques des IA, ce sont en fait des politiques du travail qui se jouent. Aujourd'hui, ces débats se renouvellent et s'orientent davantage vers l'idée d'une transformation des activités – y compris les activités cognitives de haut niveau – plutôt que vers une simple substitution. Ce déplacement reflète l'évolution des technologies et leur diffusion massive, à l'image de celle des IA génératives, mais aussi la montée en puissance d'approches qui analysent les transformations des marchés de l'emploi, des activités, des organisations et des conditions de travail, en les reliant à l'introduction et au déploiement de ces nouvelles technologies.

Par ailleurs, les intelligences artificielles ne constituent pas des objets entièrement nouveaux qui viendraient bouleverser, à elles seules, le monde du travail. Elles prennent davantage la forme de nouveaux agencements sociotechniques où se recomposent les divisions du travail, les rapports entre humains et techniques, les modalités de production des connaissances, de répartition du pouvoir d'agir et de création de valeur. Leur développement dépend autant des données, des infrastructures de calcul, des modèles économiques ou des capacités énergétiques que des contextes organisationnels, institutionnels et sociaux dans lesquels elles sont conçues, déployées et appropriées. Cette perspective conduit à raisonner et à interroger les trajectoires des IA : favorisent-elles des formes d'innovation fondées sur une intensification des capacités de calcul, de la consommation énergétique et de la concentration des ressources, ou ouvrent-elles également la possibilité d'innovations plus sobres, davantage orientées vers le développement de la qualité du travail, des collectifs et des capacités d'agir ?

C'est à la lumière de cette approche combinée, des techniques, des pratiques et des débats à leur propos, que ce numéro de la *Revue des conditions de travail* aborde les IA au travail. Ce dossier rassemble des articles qui proposent des clés de compréhension sur la fabrication et le développement des applications à base d'IA, les contextes de leur introduction en milieu de travail, leurs usages effectifs, leurs effets sur l'emploi, les conditions de travail, le contenu et l'organisation du travail, ou encore leurs effets environnementaux. Il propose en outre des perspectives pour renforcer la capacité des organisations et des collectifs de travail à garder ou (re)prendre la main sur les processus de transformation de leur milieu de travail. Différents secteurs et métiers sont ainsi traversés : l'hôtellerie-restauration, le nucléaire et la défense, une administration territoriale, les métiers de la radiologie, de la livraison de repas, de l'information. Plusieurs fonctions endossées par des IA sont exposées : des systèmes d'IA conversationnelle, des systèmes d'IA intégrés à des dispositifs de management algorithmique, destinés à optimiser l'allocation des ressources, la prévision de la demande ou l'évaluation des performances, des systèmes de reconnaissance visuelle, de transcription écrite de la voix, des systèmes d'aide au diagnostic par machine learning. Ce sont aussi les méthodes et pratiques d'enquête, d'accompagnement, de dialogue sur ces techniques qui sont interrogées. Des expérimentations mettent en lumière le nécessaire recul critique sur ces techniques, mais aussi le besoin de méthodes et d'outils d'analyse, d'évaluation et d'action pour mettre en discussion les projets de transformation numérique sur la base du travail réel et du fonctionnement et dysfonctionnement ordinaires des organisations.

Pris ensemble, les articles réunis dans ce dossier invitent ainsi à déplacer le regard : plutôt que de considérer les IA comme des technologies dont les effets seraient donnés d'avance, ils montrent que leurs usages, leurs effets et leurs trajectoires se construisent dans des situations de travail, des organisations et des rapports sociaux qui peuvent eux-mêmes devenir des espaces de discussion, d'enquête et d'action. Finalement, les contributions réunies dans ce dossier invitent à déplacer les questions habituellement posées à « l'IA ». Il s'agit moins de se demander jusqu'où ces technologies pourront se développer que d'interroger les trajectoires qu'elles dessinent : quelles formes d'organisation du travail, quels modèles de développement, quelles capacités d'action individuelle et collective, mais aussi quels enjeux de soutenabilité soulèvent-elles ?

Ce dossier n'épuise évidemment pas les questions soulevées par les IA au travail. Au moins deux dimensions demeurent notamment en retrait. La première concerne les activités de conception, d'entrainement et de maintenance des systèmes d'IA, qui restent peu documentées alors même qu'elles conditionnent largement leurs usages ultérieurs. La seconde porte sur les biais des IA et sur les formes d'inégalités qu'elles peuvent produire ou renforcer. Ces questions constituent certainement des chantiers de recherche essentiels pour prolonger les analyses réunies dans ce numéro.

La première partie, intitulée ***Les IA en situation : recomposition des activités et des métiers***, fait le choix d'aborder les IA à partir des situations de travail. Elle s'intéresse aux transformations qu'elles produisent dans l'activité quotidienne des professionnels, dans les métiers et dans les collectifs de travail. Les contributions montrent que les systèmes d'IA s'insèrent dans des organisations, des règles, des savoir-faire et des coopérations préexistants qu'ils contribuent à reconfigurer. Loin de faire disparaitre le travail humain, ils déplacent souvent les tâches, redistribuent les responsabilités, transforment les modalités de contrôle et de coordination et créent de nouvelles exigences d'ajustement, de supervision ou d'interprétation. En mettant au centre l'analyse de l'activité réelle, cette partie permet ainsi de saisir les effets parfois invisibles des IA sur le travail, les travailleuses et travailleurs.

À partir d'une enquête menée dans un hôtel indépendant, l'article de **Pierre Flandrin** interroge les effets concrets de l'introduction d'un chatbot sur le travail des réceptionnistes. Alors que ces systèmes d'intelligence artificielle sont fréquemment présentés comme des outils capables d'automatiser les tâches répétitives et d'améliorer la qualité de service, l'analyse de l'activité réelle met en évidence un écart entre les promesses d'automatisation et les usages effectifs. L'étude montre que ces dispositifs génèrent chez les réceptionnistes de nouvelles tâches de supervision, de paramétrage et de reprise des interactions, tout en réorientant leur activité vers des objectifs de collecte et d'exploitation des données clients. Elle éclaire ainsi les transformations du travail de service induites par l'IA, les formes de travail invisibilisées qui soutiennent son fonctionnement, les tensions qu'elle introduit dans l'organisation du travail et la perte de sens qu'elle génère chez les professionnels.

Au-delà de l'introduction d'outils d'IA dans des situations de travail particulières, certaines contributions interrogent les transformations plus profondes des modes d'organisation et de management. À partir d'une comparaison entre le travail des livreurs de plateformes numériques et celui de coursiers encadrés par des managers de proximité, **Anne-Marie Nicot** analyse les transformations des fonctions managériales et leurs effets sur les possibilités de régulation, de coopération et de reconnaissance au travail. L'étude montre que le management algorithmique repose sur une forte opacité des règles, une individualisation du pilotage de l'activité et une faible possibilité de discussion ou de contestation des décisions. La comparaison met en évidence que, dans le cas étudié, le management qui ne repose pas sur des algorithmes s'appuie davantage sur des relations interpersonnelles, des ajustements négociés et la construction de rapports de confiance. L'article met ainsi en lumière le rôle central, dans la régulation du travail, des interactions directes et des possibilités de délibération, et souligne les risques que fait peser leur effacement sur la santé, le sentiment de justice organisationnelle et les conditions de travail.

Toujours dans le secteur des plateformes numériques, l'article d'**Élisabeth Leblanc** déplace l'analyse des formes de management vers l'activité réelle des travailleurs. Il montre comment les algorithmes, associés à un système d'IA, structurent l'organisation du travail, encadrent les marges d'action et redéfinissent les conditions dans lesquelles les livreurs peuvent mobiliser leur expérience, leur jugement et leur intelligence pratique. L'étude met en lumière les efforts constants déployés par les livreurs pour comprendre les règles, souvent opaques et évolutives, qui gouvernent l'attribution des courses, l'évaluation des performances et les possibilités de rémunération. Elle éclaire ainsi les tensions produites

par des dispositifs de gestion qui reposent sur l'autonomie formelle des travailleurs tout en exerçant un contrôle étroit sur leur activité, avec des effets délétères sur leurs conditions de travail et leur santé.

Cette attention portée à l'activité réelle se poursuit dans un tout autre univers professionnel avec l'article de **Tamari Gamkrelidze, Moustafa Zouinar et Flore Barcellini** qui analyse les transformations du travail et de l'organisation des soins associées à l'introduction de l'IA dans le secteur médical. Il s'appuie pour cela sur l'étude de deux systèmes d'intelligence artificielle déployés en radiologie : un dispositif de dictée à reconnaissance vocale et un outil d'aide à la détection des fractures. Les auteurs montrent que ces technologies sont fréquemment mobilisées comme des réponses à des difficultés structurelles, notamment l'augmentation de la charge de travail et les tensions sur les effectifs. L'analyse des usages réels met cependant en évidence une réalité plus ambivalente. Si ces systèmes peuvent constituer des ressources pour l'activité et contribuer à sécuriser certaines tâches, ils modifient également les coopérations professionnelles, déplacent les charges de travail et soulèvent des enjeux de fiabilité, de confiance et de maintien des compétences. L'article plaide ainsi pour une approche de l'IA centrée sur le travail, attentive aux conditions concrètes de son intégration et à ses effets sur les collectifs professionnels.

Les enjeux de conception et d'intégration des systèmes d'IA sont approfondis par l'article de **Marc-Éric Bobillier Chaumon**, **Ranya Bennani, Luca Mourgaud, Myriam Fréjus, Denis Lemaître et Olivier Grisvard** qui s'intéresse à leur mise en œuvre dans des environnements à haute fiabilité. À partir de deux études de cas menées dans les secteurs du nucléaire et de la défense, les auteurs défendent une approche centrée sur l'activité réelle des professionnels et sur les transformations organisationnelles associées à l'introduction de ces technologies. Ils montrent que les enjeux de performance ne peuvent être dissociés des questions d'explicabilité, de compréhension et d'appropriation des systèmes d'IA par leurs utilisateurs. L'article souligne ainsi l'importance de concevoir des dispositifs qui préservent le pouvoir d'agir et les capacités de jugement des travailleurs, ce qui contribue au développement favorable de l'activité et des sujets au travail. Il propose enfin plusieurs principes de conception visant à développer des systèmes d'IA capables d'assister l'activité humaine sans la déposséder de ses dimensions décisionnelles et professionnelles, ouvrant ainsi la réflexion sur les conditions d'une intégration soutenable des systèmes d'IA dans les milieux de travail.

Si l'IA transforme le travail, elle est aussi un objet que l'on peut travailler. C'est le fil conducteur de cette seconde partie du dossier : ***Travailler l'IA : enquêter, concevoir, discuter, réguler.*** Les contributions réunies dans cette partie ont en commun de considérer que les trajectoires de développement des IA ne sont pas données d'avance, loin des discours d'inéluctabilité des technophobes et de la fascination technologique des technophiles. Elles montrent qu'il est possible d'enquêter sur ces technologies, d'en discuter collectivement les usages, de les concevoir autrement et de construire des dispositifs de régulation capables d'en orienter les effets sur le travail. Ces contributions proposent des méthodes et des outils pour faire de l'IA un objet d'investigation, de délibération et d'action collective.

Tout d'abord, **Pauline Gourlet** analyse les mécanismes qui invisibilisent les activités, les rapports sociaux et les formes de participation à travers lesquels l'IA se développe. Puis, à partir d'une enquête collective menée avec et par des praticiens, issus de différents univers professionnels et mobilisant ou produisant des IA, l'article montre comment les situations de travail peuvent devenir des espaces d'enquête et d'action pour reprendre prise sur les trajectoires technologiques. Son apport est ainsi également méthodologique : l'auteure propose un dispositif original d'enquête collective offrant de problématiser l'IA à partir des préoccupations et des expériences concrètes des acteurs concernés.

À cette réflexion sur l'enquête comme moyen de reprendre prise sur les trajectoires technologiques répond l'article de **Maria Sol Romero Goldar et Zacharie Hellouin de Cenival** qui montre comment une démarche interdisciplinaire et participative peut créer les conditions d'une véritable mise en discussion des choix de conception des systèmes d'IA attentive à la qualité de vie et aux conditions de travail. S'appuyant sur un projet de recherche participatif associant chercheurs en informatique, sociologues et professionnels, les auteurs analysent les ajustements, tensions et apprentissages produits par la coopération entre chercheurs en informatique, sociologues et acteurs de terrain. Au-delà de ses enseignements sur la conception des technologies, l'article constitue un retour d'expérience méthodologique sur les pratiques de recherche participative et interdisciplinaire appliquées aux systèmes d'IA.

La question de la mise en discussion des technologies se poursuit avec l'article de **Vincent Mandinaud, Ludovic Bugand, Nicolas Vispi et Grégory Plançon**, qui interroge les conditions d'une démocratisation des choix technologiques. À partir d'une revue de littérature et de plusieurs expérimentations suivies ou conduites par l'Anact, les auteurs montrent que les démarches d'enquête partant du travail réel, de discussion et de dialogue social technologique peuvent contribuer à faire de l'IA un objet de délibération collective. Ils soulignent toutefois que cette capacité à orienter les trajectoires technologiques ne peut reposer sur le seul dialogue social, mais suppose aussi de renforcer les institutions et les dispositifs permettant d'articuler durablement expertise technique, expérience du travail et représentation des intérêts.

L'article suivant montre comment ces enjeux peuvent être traduits au sein d'une collectivité territoriale. **Marie Bernard et Priscilla Capovilla** témoignent d'une démarche innovante de prévention et d'anticipation des effets de l'intelligence artificielle sur les conditions de travail. À partir de l'expérience de Nantes Métropole, elle montre de quelle façon l'introduction de l'IA peut être encadrée par des dispositifs combinant expertise, dialogue social et réflexion éthique. L'article met notamment en lumière la création d'un observatoire dédié aux impacts de l'IA sur le travail et propose une grille d'analyse des risques psychosociaux rendant possible d'identifier, en amont, les opportunités et les risques associés à ces technologies.

Au-delà d'une organisation particulière, l'article d'**Odile Chagny et Christophe Teissier** s'intéresse à une expérience originale de construction collective de repères pour le dialogue social autour de l'intelligence artificielle. À travers le projet Dial-IA, leur article retrace la manière dont des organisations syndicales, des représentants du personnel, des employeurs, des chercheurs et des experts ont élaboré collectivement des repères pour débattre des effets de l'IA sur le travail et l'emploi. Au-delà des livrables produits, l'intérêt de l'article réside dans l'analyse du processus de co-construction lui-même, des controverses qu'il a suscitées et des apprentissages collectifs qu'il a autorisés.

Des entretiens menés auprès d'experts et de partenaires sociaux alimentent également les réflexions de cette seconde partie du dossier.

Tout d'abord, l'entretien avec **Benoit Rottembourg** de l'INRIA (Institut national de recherche en sciences et technologies du numérique) prolonge ces réflexions en ouvrant la « boite noire » des systèmes d'intelligence artificielle. En revenant sur les familles d'algorithmes et leurs modes de développement, il rappelle que les propriétés des systèmes d'IA sont le produit de choix humains à toutes les étapes de leur développement (choix des données, des critères d'optimisation, de l'organisation du travail des développeurs, de la sous-traitance, des métriques de performance, etc.). L'entretien met par ailleurs en évidence les limites des approches fondées sur la seule transparence des algorithmes, et plaide pour le développement de véritables dispositifs de régulation, associant expertise technique, représentants

des travailleurs et capacités d'explicitation des décisions algorithmiques. Il apporte ainsi des repères précieux pour penser les IA comme des objets de conception, de discussion et de régulation collective.

L'entretien avec **Cyril Cosme et Philippe Marcadent** de l'Organisation internationale du travail (OIT) ouvre, quant à lui, la réflexion sur les possibilités de régulation des transformations technologiques à l'échelle internationale. À partir de l'adoption, en juin 2026, de la Convention concernant le travail décent dans l'économie des plateformes, ils reviennent sur le processus de négociation, les principes qui structurent cette nouvelle norme et les évolutions réglementaires qu'elle pourrait impulser au niveau des États. De plus, l'entretien montre comment l'OIT entend prendre part aux débats sur la gouvernance de l'intelligence artificielle, en veillant à ce que le travail et ses acteurs y occupent une place à part entière.

Quant aux trois entretiens menés avec des représentants d'organisations syndicales et patronales, ils éclairent les débats contemporains sur les cadres juridiques, les politiques publiques, le dialogue social et les trajectoires de développement des IA. Ils apportent également un éclairage sur les enjeux de régulation de ces technologies à différentes échelles – de l'entreprise aux instances européennes.

Forte de son expérience au sein du Comité économique et social européen et de la commission française sur l'IA, **Franca Salis-Madinier** défend l'idée que les cadres réglementaires actuels, en particulier l'AI Act (règlement européen sur l'IA), demeurent insuffisants pour appréhender les effets de ces technologies sur les relations et conditions de travail. L'entretien plaide pour un renforcement des droits à l'information, à la consultation, à la formation et à l'expertise des représentants des travailleurs, ainsi que pour la garantie d'un contrôle humain effectif sur les décisions algorithmiques. Et souligne que la régulation de l'IA au travail suppose d'articuler évolution du droit, dialogue social et capacité des acteurs à débattre collectivement des usages et de leurs effets.

À partir de son expérience d'Eurocadres (organisation syndicale européenne), **Nayla Glaise** soutient que l'IA ne peut être laissée aux seuls choix des concepteurs ou des employeurs : elle doit devenir un objet de négociation collective, inscrit dans le dialogue social et appuyé par un cadre juridique adapté. L'entretien appuie sur l'importance d'interroger les besoins auxquels répondent ces technologies, leurs effets sur l'organisation du travail, les marges de manœuvre des salariés et leurs conditions d'emploi. Selon Nayla Glaise, le dialogue social doit être soutenu par une régulation européenne spécifique, mieux adaptée aux enjeux du travail que les dispositifs existants.

L'entretien avec **Yahya Fallah** apporte le point de vue d'un entrepreneur engagé dans le développement de solutions d'intelligence artificielle et représentant des Entrepreneurs (anciennement Confédération des petites et moyennes entreprises). Il défend les potentialités de ces technologies pour accompagner les entreprises dans la gestion des compétences, la formation ou encore l'amélioration des pratiques de ressources humaines, tout en insistant sur le rôle déterminant de la qualité des données et de l'intervention humaine dans leur utilisation. Son rôle au sein de la commission IA des Entrepreneurs l'amène en outre à soutenir le développement des conditions de l'innovation, en particulier dans les PME et les ETI ; ces conditions relevant selon lui de plusieurs ordres : marges de manœuvre économiques des entreprises, politiques publiques de soutien, diffusion des compétences, confiance dans les acteurs économiques et vision industrielle.

La troisième partie, intitulée ***IAs, travail et emploi : Trajectoires annoncées et effets discutés***, propose un changement de focale. Elle s'intéresse davantage aux dynamiques d'ensemble qui accompagnent le développement de l'intelligence artificielle dans le monde du travail : évolutions de l'emploi, transformations des activités, mondialisation de la chaine de production des IA, représentations autour de

ces technologies. Cette perspective invite à se décaler des débats souvent polarisés qui entourent l'IA et qui tendent fréquemment à opposer des visions excessivement optimistes ou alarmistes. Les contributions réunies ici cherchent au contraire à éclairer la diversité des situations, des usages et des trajectoires observables, en replaçant les technologies dans les contextes sociaux, organisationnels et institutionnels qui façonnent leurs fabrications, déploiements et effets.

En guise d'introduction à cette troisième partie du dossier, l'article de **Janine Berg et Pawel Gmyrek** propose un état des lieux des connaissances les plus récentes sur les effets de l'IA générative sur l'emploi et le travail, à partir des travaux de l'Organisation internationale du travail. S'appuyant sur une analyse mondiale de l'exposition des professions à l'IA générative, cet article contribue à dépasser les débats polarisés la plupart du temps entre promesses et menaces de l'intelligence artificielle. En distinguant exposition, automatisation et transformation des emplois, il montre que les principaux effets de l'IA générative devraient se traduire par une recomposition du travail davantage que par une substitution généralisée des travailleurs. Cette perspective conduit à porter une attention particulière aux conditions dans lesquelles ces transformations se déploient, à leurs effets sur la qualité de l'emploi et au rôle du dialogue social dans leur régulation.

Dans la deuxième contribution de cette partie, le propos se déplace des effets attendus de l'IA sur l'emploi vers les conditions concrètes de sa production. À rebours des discours annonçant une automatisation généralisée du travail, l'enquête met en lumière le rôle central d'un travail humain souvent invisible : celui des annotateurs de données qui participent à l'entrainement, à la vérification et à la maintenance des systèmes d'IA. À partir d'une recherche menée entre la France et Madagascar, **Clément Le Ludec et Maxime Cornet** montrent que l'essor de l'IA s'appuie sur des chaines mondialisées de sous-traitance où se combinent innovations technologiques, dépendances économiques et fortes asymétries dans la répartition de la valeur. Loin de faire disparaitre le travail humain, l'IA en reconfigure les formes, les localisations et les hiérarchies. Une contribution qui rappelle que les trajectoires de l'IA se jouent autant dans les choix économiques et les rapports sociaux que dans les avancées technologiques elles-mêmes.

En revenant aux usages de l'IA dans les organisations, l'article de **Vincent Mandinaud, Yann Ferguson et Aída Ponce del Castillo** prolonge cette remise en discussion des récits dominants sur l'automatisation. L'analyse déplace le débat sur l'IA au travail de la question des emplois supprimés vers celle des transformations concrètes du travail et de ses usages. L'analyse met en regard deux dynamiques complémentaires : d'un côté, les initiatives portées par les employeurs, généralement guidées par des objectifs de rationalisation et dont les résultats demeurent contrastés ; de l'autre, les usages développés par les salariés eux-mêmes, souvent en dehors d'usages encadrés par leur organisation, et qui mobilisent des IA pour améliorer leur efficacité, leur autonomie ou la qualité de leur travail. En soulignant les tensions, les opportunités mais aussi les risques associés à ces usages, l'article plaide pour le développement d'un véritable dialogue social technologique, capable d'orienter et de réguler collectivement l'introduction de l'IA dans les organisations.

En élargissant encore la réflexion aux conditions territoriales et institutionnelles du déploiement de l'IA, la contribution de **Jean-Noël Amato, Corinne Dubois et Elisabetta Zibetti** propose un éclairage territorial sur les perceptions de ces transformations. À partir d'entretiens réalisés auprès de partenaires sociaux et d'acteurs institutionnels de La Réunion, elle examine les représentations, les attentes et les préoccupations suscitées par ces technologies, ainsi que les conditions d'un dialogue social susceptible d'accompagner leur déploiement dans un contexte socio-économique singulier.

L'entretien avec **Éric Fourboul et Gauthier Roussilhe** contribue enfin à la compréhension des dynamiques d'ensemble proposée dans cette partie en développant un autre regard. Il élargit la réflexion sur les trajectoires de l'IA en tant qu'elles sont inséparables des infrastructures énergétiques, des capacités de calcul et des modèles économiques qui en organisent le développement, et qu'ainsi elles ne sont pas seulement façonnées par les usages ou les choix organisationnels. L'un des messages forts de cet entretien : les impacts environnementaux – dont certaines méthodes d'évaluation sont décrites par les deux interviewés – restent largement absents des mécanismes qui orientent aujourd'hui ces trajectoires, alors même qu'ils en constituent l'une des principales conséquences.

{Les IA en situation : recomposition des activités et des métiers}

16 SYSTÈME D'IA CONVERSATIONNEL POUR LA RELATION DE SERVICE DANS L'HÔTELLERIE : PISTES DE RÉFLEXION ET DE RÉPONSES « NON AUTOMATISÉES »

Pierre Flandrin

25 PLATEFORMES NUMÉRIQUES DE LIVRAISON DE REPAS : QUAND LE MANAGEUR DEVIENT UNE MACHINE

Anne-Marie Nicot

36 LA LIVRAISON DE REPAS : LE JEU ALGORITHMIQUE DES PLATEFORMES

Elisabeth Leblanc

46 LA RADIOLOGIE À L'ÈRE DES « NOUVEAUX » SYSTÈMES D'INTELLIGENCE ARTIFICIELLE : USAGES ET TRANSFORMATIONS DU TRAVAIL

Tamari Gamkrelidze, Moustafa Zouinar et Flore Barcellini

60 DES IA EXPLICABLES POUR DES TECHNOLOGIES ACCEPTABLES ET SOUTENABLES POUR LE TRAVAIL : ENJEUX ET CONTOURS DE L'EXPLICATION DANS ET POUR L'ACTIVITÉ

Marc-Éric Bobillier Chaumon, Ranya Bennani, Luca Mourgaud, Myriam Fréjus, Denis Lemaître, Olivier Grisvard

SYSTÈME D'IA CONVERSATIONNEL POUR LA RELATION DE SERVICE DANS L'HÔTELLERIE : PISTES DE RÉFLEXION ET DE RÉPONSES « NON AUTOMATISÉES »

Pierre Flandrin
(Université catholique de Lille – laboratoire ETHICS EA 7446)

Dans l'hôtellerie, les systèmes d'intelligence artificielle (SIA) conversationnels sont présentés comme susceptibles d'alléger le travail de réception et d'améliorer la relation client. Menée sur deux ans dans un hôtel indépendant équipé d'un chatbot, l'étude de cas mobilise l'analyse de l'activité des réceptionnistes pour interroger cet écart. Les résultats montrent que le SIA complexifie les demandes, multiplie les reprises humaines et s'accorde mal avec une activité fragmentée, mobile et désancrée spatialement, consacrée au back-office. Loin de convertir les conversations en réservation, le chatbot capte puis exploite les données clients, inscrivant la relation de service dans une gouvernance par les nombres qui réoriente le métier vers la relance commerciale, au prix d'une perte de sens du métier. Son fonctionnement repose sur un travail invisible, externalisé puis assumé par les réceptionnistes eux-mêmes, déplaçant la finalité vers la gestion.

Mots-clés : intelligence artificielle conversationnelle, chatbot, relation de service, hôtellerie, analyse de l'activité, travail invisible

— INTRODUCTION

L'essor des systèmes d'intelligence artificielle (SIA) dans les services ouvre de nouvelles perspectives pour la gestion de la relation client (Ukpabi et al., 2019). Parmi les dispositifs employés, les SIA conversationnels ou chatbots, en tant qu'agents capables de simuler le langage humain (Dubois et al., 2019 ; Velkovska et Zouinar, 2020), sont désormais utilisés dans une diversité de contextes : aussi bien pour l'information client, la recommandation touristique, que pour des tentatives d'automatisation de la réception (Parmar et al., 2019). Leur développement s'inscrit dans un contexte hôtelier marqué par

des tensions structurelles persistantes : difficultés de recrutement, taux élevé de rotation, attractivité limitée des métiers et conditions de travail contraignantes. Le secteur de l'hébergement-restauration est identifié comme l'un des plus exposés à ces pénuries de main-d'œuvre (DRIEETS, 2023). Face à ces enjeux, les SIA sont fréquemment présentés comme des solutions techniques susceptibles de soulager les équipes en prenant en charge les tâches répétitives et peu valorisantes (Liu-Lastres et al., 2023). Ce discours technodéterministe, largement relayé par les acteurs de l'innovation, s'appuie sur la promesse d'une intelligence artificielle dite « forte » (Alexandre, 2023), censée libérer du temps pour des activités considérées comme à plus forte valeur ajoutée, souvent envisagées du point de vue des directions, telles que la personnalisation de la relation client.

Si les SIA conversationnels ont d'abord été déployés dans les grandes chaines hôtelières aux ressources importantes, leur diffusion vers des structures plus petites et indépendantes est aujourd'hui rendue possible par les interfaces de programmation (API), mais soulève dans le même temps de nouvelles interrogations : comment s'articulent-ils aux pratiques existantes dans des contextes marqués par l'incertitude, la polyvalence et la pression temporelle ? Comment ces systèmes sont-ils perçus, adoptés ou contournés par les professionnels en situation ? Et surtout, comment les travailleurs de première ligne s'approprient-ils ou non ces dispositifs dans le cadre concret de leur activité, et quels sont les effets concrets de cette intégration sur le travail réel en réception ?

1. MÉTHODOLOGIE : ÉTUDE DE CAS AU SEIN D'UN HÔTEL INDÉPENDANT

L'étude s'appuie sur une recherche longitudinale de deux ans réalisée au sein d'un hôtel indépendant, classé parmi les très petites entreprises (TPE), ayant récemment introduit un agent conversationnel à sa réception. Ce système, fourni par un prestataire de SIA tel que défini par l'AI Act[1] (Marcellin, 2025, p. 147), est conçu exclusivement pour le secteur hôtelier et équipe plusieurs milliers d'établissements hôteliers à travers le monde. Sa spécialisation sectorielle, fondée sur l'apprentissage supervisé, le distingue des IA génératives généralistes de type Large Language Models (LLM), destinées à des contextes d'usage variés (Bender et al., 2021), et limite sa transposabilité à d'autres domaines ou secteurs.

L'analyse réalisée porte sur les conditions concrètes d'appropriation de ce dispositif par les réceptionnistes et s'appuie sur une démarche d'analyse de l'activité (Guérin et al., 2021). Deux méthodes principales ont été mobilisées dans cette démarche. Les entretiens semi-directifs, tout d'abord : quatre avec le dirigeant de l'hôtel, quatre avec des réceptionnistes, deux avec le fournisseur de SIA, et un avec un professionnel chargé de l'alimentation et de l'optimisation des réponses du chatbot (entraineur). Ces entretiens ont permis d'éclairer à la fois les logiques de déploiement, les modalités d'usage et les enjeux techniques et organisationnels du système.

En complément des entretiens, 30 heures d'observation *in situ* ont été réalisées au poste de réception, dans une posture d'observation non participante. Une partie de ces observations (six heures) a été enrichie par l'usage d'un dispositif oculométrique, permettant de suivre l'activité visuelle et spatiale des réceptionnistes et de conduire, à partir de ces traces, des auto-confrontations afin d'affiner l'analyse de leur expérience. L'ensemble du corpus a été analysé selon une approche thématique, croisant les perspectives des acteurs impliqués et les régularités observées dans les situations de travail.

1 Le fournisseur est une personne physique ou morale, une autorité publique, une agence ou un autre organisme qui développe ou fait développer un système d'IA ou un modèle d'IA à usage général et le met sur le marché ou le met en service sous son propre nom ou sa propre marque, que ce soit à titre onéreux ou gratuit (p. 147).

Un épisode singulier constitue le point d'entrée empirique de notre étude : lorsqu'un client demande, de manière anecdotique, au chatbot « *Qui est Napoléon ?* », celui-ci, incapable de répondre à une requête sortant du domaine hôtelier, génère automatiquement la réponse suivante : « *Ok ! J'ai transféré votre demande à mes collègues qui reviendront vers vous au plus vite.* » Cet échange, bien que trivial, nous a conduits à interroger les logiques de délégation et les reconfigurations du travail qu'il révèle, ainsi qu'à orienter l'investigation vers les situations d'arbitrage réalisées par les réceptionnistes.

— 2. RÉSULTATS

2.1. Des objectifs d'automatisation affichés

Nous débutons la présentation des résultats par un aperçu du contexte organisationnel et des motivations ayant conduit à l'intégration du chatbot. Conçu spécifiquement pour le secteur touristique, le chatbot visait initialement, selon le fournisseur, trois objectifs prioritaires qui ont séduit le dirigeant : (1) augmenter le taux de conversion des interactions en réservations, (2) améliorer la personnalisation du service client par un gain de réactivité et de productivité, et (3) déléguer aux machines les demandes répétitives, afin de soulager les réceptionnistes. Il s'agit ainsi de libérer du temps pour les interactions en face à face avec les clients, en réduisant le « *travail des coulisses* » (Pinna, 2018). L'examen de la documentation commerciale produite par le fournisseur, mobilisée comme support lors du premier entretien, révèle un objectif explicite d'automatisation de 85 % des interactions jugées sans valeur ajoutée. Ces interactions relèvent surtout du domaine informatif, comme répondre sur l'horaire des check-in et check-out, la disponibilité d'un parking gratuit, d'une piscine. Les 15 % d'interactions restantes sont envisagées comme relevant d'un processus de redirection instantanée vers le service compétent, grâce aux capacités de réorientation assignées au dispositif conversationnel. Dans cette perspective, le rôle des réceptionnistes se trouve circonscrit à la prise en charge des demandes qualifiées de complexes ou à plus forte valeur ajoutée. Cette distribution théorique des tâches s'ancre dans une réalité économique difficile induite par la crise sanitaire au commencement de notre enquête.

2.2. De l'automatisation pensée au blocage de la conversation perçue

Le retour sur les observations filmées révèle que l'automatisation n'atteint pas les résultats escomptés selon les réceptionnistes. D'ailleurs, le chatbot n'est pas relié à une baisse concrète de la charge de travail, mais génère de nouvelles exigences, comme le commente une réceptionniste en auto-confrontation : « *50 % du temps, on doit quand même récupérer derrière. Quand je vois ici, c'était une conversation avec 10-15 messages, il faut vraiment reprendre tout depuis le début et au final, ça prend autant de temps, voire plus, que si j'avais eu le mec par mail et qu'il t'avait posé toutes ces questions dans un mail et que toi, t'avais répondu en une fois. Donc c'est pour ça que c'est vraiment oui et non* » (Réceptionniste 1). Cet extrait montre que le chatbot, malgré la gestion automatisée d'échanges longs, contraint le réceptionniste à un effort cognitif important au moment de la reprise de la conversation.

2.3. Une gestion du SIA à l'origine de tensions et de compromis

Un principe affiché du système repose sur la détection automatique des intentions formulées en langage écrit (quand un client demande par exemple à quelle heure est le check-out, le chatbot détecte l'intention « check-out »), tout en se targuant d'offrir une certaine souplesse aux hôteliers pour ajuster manuellement les réponses associées à chaque intention. Cette adaptabilité, présentée comme un atout par le fournisseur, rencontre toutefois des limites pratiques.

En effet, cette logique de précision s'est rapidement révélée couteuse en temps et en maintenance du système, en raison de la nécessité de mettre à jour fréquemment les contenus pour garantir leur pertinence. Le dirigeant en tire ce constat sans équivoque : « *Plus on est précis dans nos réponses, plus*

il nous faut être rigoureux dans la gestion de l'outil. Au début, on a voulu encoder l'ensemble de nos bières, sauf que finalement, nos relations avec les brasseurs changent souvent, donc si vous voulez donner la bonne réponse au client, ça demande d'ajuster constamment nos réponses. Désormais, notre chatbot propose une réponse assez générique que l'on a nous-même écrite et personnalisée, mais qui ne permet pas de satisfaire totalement le client dans sa version automatisée » (Entretien Dirigeant, n° 4). Toute tentative d'augmenter la précision des réponses accroit la rigueur exigée pour la gestion du système, et expose à des effets de bord : les ajustements manuels peuvent entrer en conflit avec d'autres intentions préexistantes, rendant la configuration peu efficace et instable. La situation précédente souligne ainsi les tensions inhérentes entre automatisation et personnalisation et illustre le compromis opéré entre qualité de l'information délivrée et charge de travail induite pour la maintenance du chatbot.

2.4. Des usages du SIA mis en perspective de l'activité spatio-temporelle des réceptionnistes

La disponibilité aux clients et l'immédiateté de réponse sont désignées tant par le dirigeant que par les réceptionnistes comme la principale transformation des exigences professionnelles dans l'hôtellerie. L'analyse ergonomique de l'activité de réception, menée à partir d'observations systématiques, indique un décalage significatif sur cet aspect. L'observation d'un réceptionniste sur un créneau de deux heures révèle que seules 24 minutes sont consacrées à des interactions directes avec les clients, réparties sur 14 échanges brefs. La majorité du temps de travail des réceptionnistes est en réalité absorbée par des tâches de back-office, souvent peu visibles mais essentielles à la continuité du service : gestion des réservations, traitement des avis clients, coordination avec les équipes de nettoyage et de maintenance, et même réassort de linge ou aide ponctuelle à la préparation des repas. Une séquence particulièrement significative concerne la répartition des chambres à nettoyer, à laquelle une réceptionniste consacre près de 20 minutes en début de shift. Si cette activité ne relève pas de son périmètre prescrit, elle témoigne d'un engagement dans la qualité du service rendu et d'une forme de coopération inter-métiers visant à préserver l'efficience collective, notamment en limitant les déplacements des agents d'entretien. Ce travail de coordination inter-métiers se cristallise dans la phase de « *handover* », période de transmission des informations entre l'équipe de nuit et celle du matin, qui mobilise également des échanges avec les techniciens et les femmes de ménage. Loin de l'image d'un poste centré sur l'accueil du client, l'activité de réception apparait ainsi comme un nœud organisationnel, dans lequel se joue la régulation de nombreuses interrelations humaines et techniques.

Dans ce contexte, l'usage du chatbot se heurte à une réalité de terrain, celle d'une activité fragmentée, mobile et multicanale. L'analyse des interactions montre que la reprise manuelle d'une conversation bloquée est fréquemment asynchrone et différée : dans 8 cas sur 14, les réceptionnistes sont engagés dans un autre espace, souvent dans des situations de coopération avec des collègues. Dans ces cas, le délai moyen de reprise est alors de 5 minutes 30, et atteint 8 minutes lorsqu'un échange en face à face est en cours. Cette situation devient particulièrement marquée lors des shifts réalisés seul, rendant le suivi du chatbot difficile. Les résultats, une fois quantifiés, montrent que le contexte multicanal et mobile limite l'efficacité du chatbot dans la reprise en main des interactions. Cette difficulté à user du chatbot est plus importante pour une personne seule à la réception (poste du soir, poste de nuit), une situation de plus en plus fréquente (Figure 1).

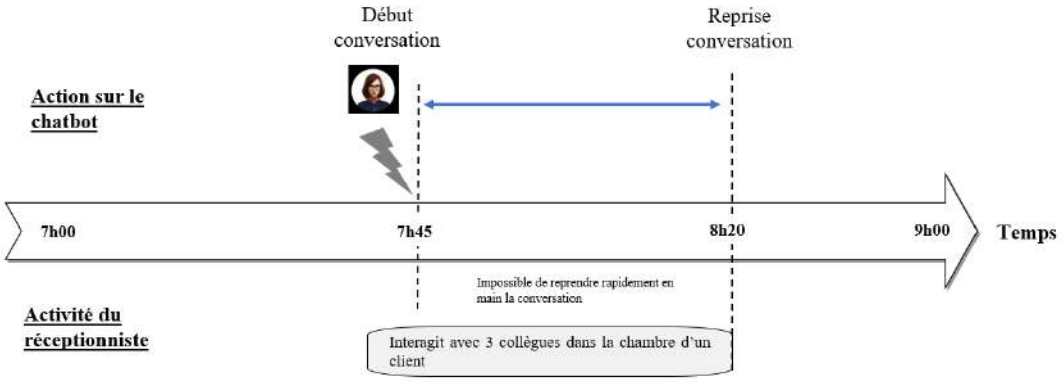


Figure 1. Chronique d'activité simplifiée associant action sur le chatbot et activité du réceptionniste

Cette complexité, rarement prise en compte dans les discours sur l'automatisation, pose la question des effets du système sur une activité déjà marquée par une obligation de réactivité.

2.5. Un fonctionnement du SIA nécessitant du travail humain externalisé et internalisé

Si l'on se penche plus spécifiquement sur le fonctionnement du chatbot, on comprend que son efficacité repose en réalité sur un ensemble de tâches humaines, souvent invisibilisées, fragmentées et en partie délocalisées. Un maillon essentiel consiste d'abord en un travail humain de classification des intentions pour que le chatbot puisse attribuer une intention à des formulations variées, comme, dans le cas présent, avec le repérage de l'intention « prix » : « *Quel est le prix pour une nuit en chambre double ?* », « *Combien coute une chambre pour deux personnes ?* », ou encore « *Tarif d'une chambre double, s'il vous plait* » (Entretien Entraineur). Ce travail de préparation est ici externalisé vers des pays aux contextes socio-économiques éloignés du terrain d'usage, où les couts de main-d'œuvre sont plus faibles et les capacités de compréhension linguistique jugées suffisantes par le fournisseur. Chargés d'annoter les corpus d'interactions, ces travailleurs accomplissent des tâches essentielles au développement des SIA, bien que ces opérations demeurent largement ignorées dans le discours de ceux qui visent leur déploiement.

À cette première couche de travail dit invisible s'ajoute une autre activité cette fois internalisée au sein de l'hôtel, directement inscrite dans l'activité des réceptionnistes. La paramétrisation du chatbot est en effet une activité couteuse en temps et en investissement pour les réceptionnistes novices ou en formation qui doivent configurer les réponses, ajuster les scripts et tester les liens entre les différentes intentions spécifiques à l'hôtel. De simples changements liés à la tarification ou la rénovation d'une chambre entrainent d'importantes reconfigurations (activer ou désactiver certaines intentions) et le réencodage manuel des intentions. Certaines questions fréquemment posées ne sont par exemple pas préexistantes, ce qui implique pour le personnel de devoir inscrire directement dans l'interface du chatbot la question et la réponse à apporter : « *Juste pour exemple cette semaine j'en ai eu un, la personne a demandé quand l'hôtel a été construit. (Le chatbot) a répondu que l'hôtel était réouvert depuis le 19 juin. Donc il y a quand même des bugs* » (Entretien Réceptionniste 1, n° 2). Des questions et des demandes qui n'auraient sensiblement jamais été posées directement doivent désormais être traitées par les réceptionnistes. Une partie essentielle de ce travail repose sur les réceptionnistes en formation, dont il est attendu qu'ils testent le chatbot tout en se formant à leur propre fonction. En cas d'erreur ou d'incompréhension du chatbot, ils sont chargés d'en identifier la cause, de la remonter, voire d'en proposer une reformulation. Ce travail d'interrogation active du dispositif constitue une forme de travail invisible de conception *in situ* à l'hôtel.

2.6. Vers plus de réservations ou une meilleure quantification

Sur le plan des résultats, l'analyse des interactions générées depuis le déploiement du chatbot rappelle qu'aucune réservation n'a été réalisée avec son utilisation. Si, à première vue, cette absence de conversion pourrait être perçue comme un indicateur d'inefficacité, au regard de la direction, elle révèle un autre type de valeur, plus discrète mais stratégique, liée à la collecte et à la capitalisation de données clients. Interrogé sur la finalité du système, le fournisseur le présente comme un outil capable d'orchestrer un processus de suivi « proactif » : « *Tout en guidant les clients à travers le processus de réservation, l'intelligence artificielle obtient des données exploitables (ainsi que le consentement RGPD), permettant de relancer le client par courriel ou de transmettre des informations stratégiques à l'équipe de réservation, qui prendra le relais par téléphone. Ce type de processus conduisant mécaniquement à une conversion plus élevée* » (Entretien Fournisseur, n° 2). À travers cette configuration, le chatbot fonctionne comme un artefact médiateur de la mise en données du client où les chiffres produits ne servent pas seulement à évaluer, mais aussi à fournir les conditions d'une activité future jugée plus efficace et rentable du point de vue de la direction hôtelière. Le process d'interaction engagé est alors en réalité bidirectionnel : « *En même temps que le client va recevoir de l'information, il va nous en donner. C'est-à-dire que le client ne se rend même pas compte qu'il nous donne ces informations : son nom, son mail. L'idée c'est quoi, c'est que si le client… imagine, il fait une demande, mais il ne termine pas sa réservation. En fait, on va pouvoir le réengager par téléphone aussi. Donc autant prendre l'information que l'on chope sur le client. Là, par exemple, tu vas mettre ton nom, ton e-mail, tu vas accepter notre politique de privacy. GDPR. C'est-à-dire que s'il ne termine pas son process de réservation, eh bien on va pouvoir le relancer* » (Entretien Fournisseur, n° 1).

Dans ce cas d'usage précis du chatbot au sein de cet hôtel, le chatbot se positionne non pas comme un support transactionnel au sens strict, c'est-à-dire capable de finaliser une transaction, mais comme un dispositif de collecte des informations visant à mieux encadrer et réorienter l'activité des réceptionnistes. Le chatbot vient pallier certaines limites du système de gestion de la relation client (CRM), en automatisant l'enregistrement d'informations personnelles dès la première interaction (nom, prénom, adresse e-mail). Il réoriente l'activité des réceptionnistes sur une dimension qui se veut plus commerciale et moins reliée à une logique métier de service. Cette redéfinition des usages et de l'usage des données traduit une évolution stratégique : le chatbot sert de mécanisme d'enrichissement de la base de données clients tout en reconfigurant la posture des réceptionnistes vers une approche résolument différente et questionnante du point de vue du métier et du rapport à ce dernier.

— 3. DISCUSSION CONCLUSIVE

Cette recherche met en évidence un décalage structurant entre les promesses associées à l'introduction d'un SIA conversationnel dans la relation de service et les transformations effectives de l'activité de travail des réceptionnistes. Présenté comme un système d'automatisation visant à alléger la charge de travail et à améliorer la personnalisation des échanges avec les clients, le chatbot participe en réalité d'une reconfiguration plus profonde du métier, des formes d'organisation du travail et des finalités assignées à la relation de service.

L'analyse met ainsi en lumière plusieurs effets majeurs liés à l'introduction du chatbot : (1) une complexification des demandes clients à traiter, (2) une nécessité fréquente de reprises humaines, (3) une difficulté accrue à organiser et prioriser les tâches, (4) une orientation renforcée de l'activité vers des objectifs commerciaux, et (5) une extension des relations de service au-delà des interactions dites « traditionnelles » en réception (Vanlair et Largier, 2017). Plus largement, le décalage observé entre la promesse d'un SIA palliatif et la réalité de ses usages montre que le chatbot, conçu comme une

prothèse cognitive pour les réceptionnistes, ne s'intègre que partiellement aux logiques réelles de leur activité (Bobillier Chaumon, 2021). Au sein de cette TPE hôtelière, nos résultats contribuent à montrer que les reprises humaines restent fortement contraintes, tant par l'intensification du travail que par la polyvalence imposée aux réceptionnistes, dans un environnement où il faut arbitrer en permanence entre plusieurs types de demandes clients.

L'étude met également en évidence une incompatibilité entre, d'une part, les enjeux d'immédiateté et de personnalisation portés par le discours technologique et, d'autre part, les réalités d'une organisation du travail flexible. Les réceptionnistes ajustent en permanence leur activité en fonction de la présence physique sur site, du nombre de clients et des tâches à accomplir. L'activité de réception apparait dès lors largement désancrée spatialement, à la différence d'autres formes de médiation technologique déjà étudiées (Gras Gentiletti et al., 2022). Cette désorganisation spatiale et temporelle contribue à retarder les reprises en main du chatbot.

Nos observations ont révélé un décalage significatif entre les attentes et la réalité du travail constaté : les réceptionnistes passent la majorité de leur temps en back-office, à des activités collectives et organisationnelles peu visibles (coordination interne, gestion logistique), limitant paradoxalement leur disponibilité effective pour les clients. La mise en place du SIA conversationnel a ainsi généré une complexification des demandes, nécessitant régulièrement des reprises manuelles de la part des employés. Le chatbot n'a pas démontré de capacité directe à convertir les interactions en réservations immédiates. Son utilité se perçoit plutôt en termes de collecte et d'exploitation de données clients, alimentant des pratiques de rationalisation statistique (Benedetto-Meyer, 2011).

Le dispositif apparait moins comme un outil d'automatisation du service que comme une stratégie commerciale proactive, réorientant insidieusement le métier traditionnel de réceptionniste vers une fonction de suivi et de relance commerciale. Cette reconfiguration du métier s'accompagne d'un travail supplémentaire d'alimentation, de paramétrisation et d'optimisation du chatbot, réalisé par les réceptionnistes eux-mêmes, mais au demeurant invisible aux yeux des clients (Tubaro, Casilli et Coville, 2020). La supervision continue du chatbot, comprenant l'ajustement des réponses aux intentions, le suivi des interactions et la correction des erreurs, s'ajoute à la gestion simultanée d'un ensemble de canaux numériques. Ces activités constituent des formes de travail pourtant essentielles au fonctionnement des systèmes d'intelligence artificielle, mais souvent répétitives, peu valorisées et invisibilisées par les dispositifs qui les rendent possibles.

Si l'introduction du chatbot autorise une meilleure connaissance des clients, elle a aussi entrainé une perte de sens pour les réceptionnistes. L'outil utilisé pour produire des données quantifiables sur les clients et sur l'activité même des travailleurs (nombre d'interactions, mises à jour du CRM, etc.) inscrit pleinement le dispositif dans une logique de gouvernance par les nombres (Supiot, 2015). À cet égard, il participe d'une transformation profonde de la relation de service, en substituant à la relation intersubjective une logique de démarchage rendue possible par la collecte et l'exploitation de données personnelles.

Ce basculement, qui s'inscrit dans une logique de rationalisation de l'activité à des fins de pilotage, s'accompagne de nouvelles formes de travail invisibilisées. Du côté de la conception, des professionnels plutôt qualifiés sont chargés d'entrainer, superviser et adapter le chatbot. Du côté opérationnel, les réceptionnistes prennent en charge la paramétrisation continue du SIA (ajustement des réponses aux intentions, suivi des interactions), tout en devant gérer en parallèle un ensemble croissant de canaux numériques. Ces formes de travail, pourtant essentielles au fonctionnement des systèmes d'intelligence artificielle, relèvent souvent de tâches répétitives et peu valorisées et invisibilisées par ceux qui fournissent ces systèmes (Tubaro, Casilli et Coville, 2020).

À travers cette étude de cas, nous avons cherché à montrer dans quelle mesure les choix de conception assignés à un système d'IA « déterminent », d'une certaine manière, sa place dans l'activité des réceptionnistes et les possibilités d'usage, autrement dit son acceptation *in situ* (Bobillier Chaumon, 2021). Au-delà de sa faible contribution à l'amélioration ou à la personnalisation de la relation de service, le système étudié dans ce cas s'inscrit dans une logique plus large de capitalisation des données clients, enrichissant les dispositifs de pilotage de la relation client. Ce mouvement marque un déplacement de la finalité initiale du service vers une perspective davantage orientée vers la gestion et l'optimisation.

Enfin, cette étude plaide pour une prise en compte renforcée des usages réels aux niveaux micro et méso de l'activité (Zouinar, 2025), ainsi que pour un dialogue accru entre fournisseurs de SIA, décideurs stratégiques et travailleurs de première ligne dans la conception de ces outils. Si les résultats mettent en évidence les limites et contraintes associées à l'introduction du chatbot dans ce type d'environnement socio-technique, ils montrent également que cette technologie rend visibles des choix organisationnels et des orientations du travail jusqu'alors peu formalisés. À ce titre, le chatbot peut aussi constituer une porte d'entrée pour une discussion collective sur les pratiques de travail et les finalités de l'activité.

BIBLIOGRAPHIE

- Alexandre, L. (2023), *La Guerre des intelligences à l'heure de ChatGPT*, Paris, JC Lattès, 480 p.
- Bender, E., McMillan-Major, A., Gebru, T. et Mitchell, M. (2021), « On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? », Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. p. 610-623.
- Benedetto-Meyer, M. (2011), « Des statistiques au cœur de la relation clients : l'accès aux "données clients", leur effet sur l'organisation du travail et les relations clients/vendeurs en boutique », in *Sociologies pratiques*, 22(1), 49-61.
- Bobillier Chaumon, M.-É. (2021), « Technologies émergentes et transformations digitales de l'activité : enjeux pour l'activité et la santé au travail », in *Psychologie du travail et des organisations*, 27(1), 17-32.
- Direction régionale et interdépartementale de l'économie, de l'emploi, du travail et des solidarités d'Île-de-France [DRIEETS] (2023), *Synthèse thématique n° 109 : Marché du travail francilien – Des tensions sur le marché du travail accentuées par la reprise post-crise*, ministère du Travail, du Plein Emploi et de l'Insertion, https://idf.drieets.gouv.fr/sites/idf.drieets.gouv.fr/IMG/pdf/synthese_thematique_109_vf.pdf
- Dubois, C., Salotti, J., Seminel, D. et Simonazzi, N. (2019), « Le chatbot : un outil de la relation aux clients », in *Hermès, La Revue*, 84, 95-97.
- Gras Gentiletti, M., Bourmaud, G., Fréjus, M. et Decortis, F. (2022), « Concevoir pour des activités instrumentées par des chatbots », in *Activités*, 19(1).
- Guérin, F., Pueyo, V., Béguin, P., Garrigou, A., Hubault, F., Maline, J. et Morlet, T. (2021), *Concevoir le travail, le défi de l'ergonomie*, Toulouse, Octarès, 424 p.
- Liu-Lastres, B., Wen, H. et Huang, W.-J. (2023), « A reflection on the Great Resignation in the hospitality and tourism industry », in *International Journal of Contemporary Hospitality Management*, 35 (1), 235-249.
- Marcellin, S. (2025), *Droit de l'intelligence artificielle* (1re éd.), Paris, Gualino, 270 p.
- Pinna, G. (2018), « Servir sur la scène ou dans les coulisses ? Une étude dramaturgique des tensions au sein des collectifs de travail dans l'hôtellerie de luxe », in *L'Homme et la Société*, 206(1), 317-339.
- Parmar, S., Meshram, M., Parmar, P., Patel, M. et Desai, P. (2019), « Smart hotel using intelligent

chatbot: A review », in *International Journal of Scientific Research in Computer Science Engineering and Information Technology*, 5(2), 823-829.

— Supiot, A. (2015), *La Gouvernance par les nombres, Cours au Collège de France (2012-2014)*, Paris, Fayard, 512 p.

— Tubaro, P., Casilli, A. et Coville, M. (2020), « The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence », in *Big Data & Society*, 7(1).

— Ukpabi, D.C., Aslam, B. et Karjaluoto, H. (2019), « Chatbot adoption in tourism services: A conceptual exploration », in *Robots, Artificial Intelligence and Service Automation in Travel, Tourism and Hospitality* (p. 105-121), Leeds, Emerald Publishing Limited, 274 p.

— Vanlair, J. et Largier, J. (2017), « Une relation de service étendue : le cas de l'équipement en smartphones des agents d'escale de la SNCF », in Benedetto-Meyer, M., Cihuelo, P. et Largier J. (dir.), *L'Organisation face aux transformations de la relation de service* (p. 57-87), Paris, L'Harmattan, 194 p.

— Velkovska, J. et Zouinar, M. (2020), « Les relations aux machines "conversationnelles" : vivre avec les assistants vocaux à la maison », in *Réseaux*, 220-221(2-3), 47-79.

— Zouinar, M. (2025), « Intelligence artificielle au travail », in *Chroniques du travail* (15).

PLATEFORMES NUMÉRIQUES DE LIVRAISON DE REPAS : QUAND LE MANAGEUR DEVIENT UNE MACHINE

Anne-Marie Nicot
(Anact)

Le développement du management algorithmique pose de nombreuses questions en termes d'impact sur les conditions de travail. Au-delà des effets déjà documentés en termes d'intensification et de limitation de l'autonomie dans le travail, cet article analyse plus précisément ses effets sur les dynamiques psychosociales. Il s'appuie pour cela sur la présentation de deux configurations de travail qui présentent des similitudes suffisantes pour soutenir la mise en contraste. Dans l'une, le management algorithmique est développé à son maximum et les interactions avec un manager humain sont quasi inexistantes. Dans l'autre, à l'inverse, le manager humain est au cœur de l'organisation et de la régulation du travail. Il ressort de cette analyse que, au-delà des problématiques d'intensification ou de surveillance liées aux choix de conception des outils numériques, être managé par des algorithmes implique une interaction dans laquelle les deux parties ne peuvent pas partager une expérience commune du réel et du monde pour stabiliser des règles pratiques et sociales de travail.

Mots-clés : management algorithmique, régulation du travail, plateforme, livraison, intensification.

INTRODUCTION

Les outils numériques sont utilisés depuis de nombreuses années dans le management du travail, que ce soit pour planifier les horaires de travail ou attribuer des tâches. Depuis une dizaine d'années, un seuil de développement de ces outils a été franchi, notamment avec l'apparition des plateformes numériques de mobilité (les plus connues étant Uber, Uber Eats et Deliveroo). Celles-ci ont en effet mis en place un dispositif technico-organisationnel pour diriger le travail des livreurs et chauffeurs sans intervention directe d'un manageur humain. Depuis, de nombreuses recherches ont été conduites sur ce qui est désormais désigné sous le terme de « management algorithmique », que l'OCDE définit comme « *le recours à des outils technologiques, qui peuvent intégrer l'intelligence artificielle (IA), pour automatiser partiellement ou totalement des tâches généralement dévolues aux cadres* » (Milanez et al. 2025). L'utilisation de ces outils peut être plus ou moins intensive, laissant plus ou moins de place au management humain.

Le modèle organisationnel des plateformes numériques de livraison de repas, avec une automatisation quasi complète des processus managériaux, constitue la forme la plus intensive de déploiement du management algorithmique (Fernández-Macías et al., 2023). De nombreux travaux ont déjà mis en évidence le fait que l'utilisation de ces nouveaux dispositifs était associée à une augmentation des principaux facteurs de risques psychosociaux reconnus (Martin et al., 2025). Cependant, « *l'impact du management algorithmique sur les travailleurs est façonné non seulement par ses caractéristiques techniques, mais aussi par la manière et les objectifs pour lesquels il est conçu et utilisé*[1] » (ibid., p. 17).

Dans cet article nous proposons de décrire la manière dont sont organisées deux activités relativement proches, l'une réalisée au moyen des algorithmes des plateformes[2], l'autre par des manageurs humains. L'article interroge ainsi les transformations de la fonction managériale dans l'organisation du travail, en comparant une configuration fondée sur un management algorithmique et une configuration reposant sur des manageurs humains de proximité, afin d'en analyser les effets sur les conditions de travail et les dynamiques psychosociales. L'objectif n'est pas ici de proposer une comparaison terme à terme entre deux configurations qui seraient strictement équivalentes, mais de mobiliser ces deux situations comme des contrastes analytiques permettant d'éclairer les transformations des interactions entre manageurs et travailleurs, et de leurs effets sur les conditions de travail.

1 « *The impact of [algorithmic management] on workers is shaped not only by its technical features but also by how, and for what purposes, it is designed and used.* » Notre traduction.

2 Cette analyse porte sur les plateformes numériques de livraison à vélo de nature commerciale uniquement (pas sur les modèles alternatifs et coopératifs, tels que développés autour de CoopCycle par exemple, qui réservent encore une forte place aux interactions humaines dans leur fonctionnement et leur management).

Méthodologie

Pour cette comparaison, nous nous appuierons sur une revue de littérature réalisée dans le cadre de l'expertise collective « *Évaluation des risques sanitaires pour les travailleurs des plateformes numériques de livraison de repas en France* » (Anses, 2024) à laquelle nous avons contribué.

Les études scientifiques mobilisées pour cette expertise ont été identifiées par différents canaux :
— recherche bibliographique sur la base de données Scopus[3] ;
— recherche bibliographique sur d'autres bases de données et site éditeur (notamment Cairn.info[4] pour les sciences humaines) ;
— injection de toutes publications pertinentes connues des experts du groupe de travail.

Cette revue bibliographique a été complétée par des auditions d'acteurs institutionnels et une étude exploratoire sur la santé physique et mentale des livreurs.

La première partie de cet article, sur les livreurs de repas via les plateformes numériques, reprend de manière synthétique les résultats de l'expertise collective sur l'organisation de cette activité, et les effets du management algorithmique.

La deuxième partie s'appuie sur une publication, identifiée dans le cadre de cette expertise, qui présentait les résultats d'une étude ergonomique basée sur des entretiens et des observations de situations de travail, réalisée en 2004 sur le métier de coursier à deux-roues motorisés.

Les corpus mobilisés pour cette comparaison sont donc de nature différente. Et chacun comporte des limites. L'étude ergonomique, notamment, ne porte que sur deux entreprises de taille intermédiaire et ne développe ses analyses que par rapport à son objet d'étude, à savoir « *les liens qui existent entre, d'un côté, la structuration de la profession, dans ses aspects économiques, organisationnels et réglementaires, et de l'autre le risque encouru par les salariés et la place laissée au développement de pratiques de prévention par les entreprises employantes* » (Van Belleghem et Bourgeois, 2004, p. 5). D'autres aspects du fonctionnement de ces entreprises n'ont donc pas été décrits.

Il n'en reste pas moins que les enjeux et pratiques mis en évidence dans cette étude trouvent une certaine résonance avec le métier de livreurs de repas via les plateformes, à la fois dans les similitudes (cf. infra §31) et dans les différences. Nous nous intéresserons ici particulièrement au contraste entre le rôle central du manageur humain dans un cas, et son absence totale dans l'autre cas, et sur les incidences que cela comporte sur les conditions psychosociales du travail.

— 1. LES PLATEFORMES : UN SYSTÈME SOCIOTECHNIQUE

Les premières plateformes de livraison de repas sont apparues en 2014 en France. Elles ont introduit plusieurs bouleversements dans l'organisation historique de la livraison de repas. Il s'agit, en premier lieu, d'un *bouleversement serviciel*, puisque ces plateformes proposent des applications, avec un large choix de restaurants et la promesse d'une livraison rapide et à cout très réduit. Le bouleversement est

3 Scopus est une base de données transdisciplinaire de résumés et de citations de publications scientifiques lancée par l'éditeur scientifique Elsevier. Elle référence environ 25 000 journaux scientifiques et intègre chaque année près de 3 millions de nouvelles références : articles scientifiques, publications industrielles, collections d'ouvrages, actes de conférence.

4 Cairn.info est un portail Web francophone consacré aux sciences humaines et sociales (et plus récemment aux sciences, technologie, ingénierie, mathématiques et médecine). Le contenu ainsi rassemblé comprend plus de 360 000 articles issus de plus de 600 revues.

également *technologique*, ces services reposant sur des algorithmes sophistiqués de géolocalisation, d'allocation des commandes et de tarification. Dans ce cadre, l'intelligence artificielle renvoie ici principalement à des systèmes d'apprentissage automatique permettant d'ajuster en continu les modalités d'appariement et de tarification à partir du traitement massif de données (Turakia, 2017 ; Zhu et Laptev 2017). Les degrés d'automatisation et d'apprentissage de ces dispositifs peuvent toutefois varier selon les plateformes et les usages. Enfin, il s'agit aussi et surtout d'une *mutation des formes d'organisation du travail,* puisque ces plateformes se présentent comme un simple intermédiaire numérique par lequel s'établit la relation commerciale entre le client, le restaurant et le livreur.

1.1. Mobiliser une « foule » de travailleurs indépendants

Les livreurs sont donc des travailleurs indépendants – ils ne sont salariés ni des plateformes ni des restaurants (comme c'était le cas avant l'apparition des plateformes).

Comme l'indique explicitement Uber dans sa notice d'introduction en Bourse[5], le fait de recourir à des travailleurs indépendants est un pilier du modèle d'affaires des plateformes de mobilité (livraison de repas, transport de personnes). En effet, travailler avec des salariés impliquerait l'application de l'ensemble des dispositions légales, réglementaires et conventionnelles sur le temps de travail, les salaires (minimum horaires, majorations, etc.), les cotisations sociales, la fourniture et l'entretien des équipements de travail, la formation, la santé et sécurité au travail, etc.

Plus encore, comme la demande pour ce type de prestation est très fluctuante (selon les jours de la semaine, selon les heures de la journée, selon les aléas climatiques ou autres), pour réaliser leur « promesse » auprès des clients (livrer rapidement et pas cher), les plateformes ont besoin de pouvoir mobiliser beaucoup de livreurs lorsqu'il y a une forte demande – *sans avoir à les payer sur le reste du temps*.

C'est pourquoi, dans leur communication auprès de potentiels livreurs, les plateformes de livraison de repas mettent en avant la facilité pour accéder au travail. Et de fait, les barrières à l'entrée sont réduites autant que possible, car les plateformes ont besoin d'une masse importante de livreurs à mobiliser sur les heures de pointe.

Ce recours à des travailleurs indépendants (et non des salariés qui sont tenus d'obéir aux prescriptions de leur employeur) pourrait présenter des risques pour les plateformes : les livreurs pourraient se connecter aux horaires qui leur conviennent et non quand il y a une forte demande, refuser les courses qui ne les intéressent pas, ne pas livrer dans les délais, etc. C'est pourquoi les plateformes mettent en place un ensemble de dispositifs, qui s'appuient largement sur le management algorithmique, pour orienter les comportements des livreurs dans le sens de leurs besoins.

1.2. La fabrique de comportements conformes

Si les livreurs sont des travailleurs indépendants, les plateformes maitrisent les aspects essentiels de leur activité, et donc de leurs revenus. Car ce sont elles qui, via leurs algorithmes, attribuent les tâches (livraisons) et en déterminent le prix. Elles peuvent aussi leur fermer l'accès au travail en déconnectant les comptes des livreurs (temporairement ou définitivement).

Pour attribuer les livraisons et en fixer le prix, elles utilisent des algorithmes d'appariement (ou « dispatch ») et de tarification qui, avec l'appui de l'intelligence artificielle, ont la capacité de « *traiter chaque*

5 Prospectus d'introduction en Bourse d'Uber, rubrique « facteurs de risque ».
https://www.sec.gov/Archives/edgar/data/1543151/000119312519103850/d647752ds1.htm#toc647752_2

livreur individuellement et instantanément, malgré leur grand nombre et leur dispersion » (Lemozy et Le Lay, 2021a, p. 101), en comparant leurs performances en temps réel.

Pour cela, les plateformes recourent à une collecte massive et continue de données. Toutes les interactions du livreur avec la plateforme via l'application laissent des « traces » numériques (lieu et heure de connexion, acceptation ou refus d'une course, délai de réponse à une notification, géolocalisation, etc.) qui autorisent un suivi continu et très précis de ce que font les livreurs. Cette collecte automatique de données est complétée par la notation du livreur par le client et, éventuellement, le restaurateur.

Sur leur application, les livreurs voient s'afficher toute une production statistique relative à leur compte. Ce qui est affiché peut varier selon les plateformes. Mais, *in fine*, les livreurs ne savent pas quelles sont les données utilisées par les algorithmes, ni comment elles sont utilisées, pour décider de ce qui est déterminant pour eux : l'affectation des livraisons et leur tarification. Car les plateformes maintiennent une opacité quasi complète sur les règles de construction des algorithmes[6]. De plus, les observations donnent à penser que ces algorithmes sont très mouvants : ils changent très fréquemment du fait de l'apprentissage automatique, voire ne sont pas les mêmes pour deux utilisateurs au même moment (en raison des pratiques d'A/B testing[7]).

Les livreurs n'étant payés que lorsqu'ils ont des livraisons à réaliser, ces algorithmes sont au centre de leurs préoccupations. La « *peur de ne pas rouler* » (Lemozy et Le Lay, 2021b, p. 185), c'est-à-dire de ne pas être en capacité de générer un chiffre d'affaires estimé convenable dans la journée, est un aspect déterminant dans le fonctionnement des plateformes de livraison. En effet, les livreurs vont chercher à répondre continuellement aux attendus réels ou supposés de l'algorithme (horaires de connexion, acceptation des livraisons, rapidité de livraison, satisfaction du client, etc.) dans l'espoir de produire des données qui leur permettront d'obtenir des livraisons plus nombreuses et mieux rémunérées.

Et comme les livreurs ne savent pas pourquoi les demandes de livraison leur sont affectées ou pas, ni comment les tarifs sont fixés, le fonctionnement des algorithmes est à l'origine de nombreux échanges entre les livreurs, chacun partageant ses hypothèses pour expliquer ce qui semble *a priori* inexplicable. Dans les entretiens qu'elle avait conduits avec des livreurs à Édimbourg, Gregory (2021, p. 326) constatait ainsi que « *presque chaque livreur avait une compréhension différente de la manière dont le travail était distribué, créant leurs propres histoires sur la manière dont le système fonctionne*[8] ». Leblanc (2024, p. 254-308) donne une large illustration de ces multiples hypothèses construites et partagées sur les réseaux sociaux par les livreurs : être connecté longtemps, ou tard le soir, ou tôt le matin, ou rester en mouvement plutôt que d'attendre à l'arrêt, « *mettre j aime aux restos que tu viens de livrer et aux clients qui ont commandé ; mettre des pouces a tous les commerçants* » (*sic*, p. 255), etc.

Chaque livreur tente en permanence de réguler son activité en fonction de « *règles inconnues, contingentes et aléatoires. Il tente sa chance dans le jeu du hasard* » (ibid., p. 263). Et « *l'incertitude permanente et un facteur de stress, de dévalorisation de soi, de perte de repères et de confiance en ses capacités, de conflit et de violence* » (ibid., p. 308).

6 Les pages documentation d'Uber Eats (https://www.uber.com/fr/blog/notre-systeme-de-proposition-de-commandes-en-toute-transparence/) et Deliveroo (https://riders.deliveroo.fr/fr/support/livrer-une-commande/comment-et-pourquoi-les-commandes-sont-proposees-) à ce sujet sont tellement évasives que la description restera conforme, y compris en cas de changements drastiques des algorithmes utilisés.

7 L'A/B testing consiste à proposer deux versions d'un algorithme à deux groupes d'utilisateurs répartis sur chaque version de manière aléatoire, pour identifier la manière dont une métrique prédéfinie est impactée. Une fois qu'est déterminée la version la plus efficace, celle-ci est conservée pour l'ensemble des utilisateurs.

8 « *Almost every rider had a different understanding of how work was distributed, creating their own stories of how the system functions.* » Notre traduction.

Ces effets de l'incertitude permanente sont renforcés par la faible qualité des interactions des livreurs avec la plateforme : il est quasiment impossible d'obtenir une discussion ou une révision des données et des décisions, même lorsqu'elles semblent injustifiées.

Le fonctionnement des plateformes de livraison (et plus globalement des plateformes de travail, cf. Bérastégui, 2021, p. 40-54) se traduit donc par plusieurs formes d'injustice organisationnelle : injustice distributive (les rétributions ne sont pas proportionnelles aux efforts fournis), injustice procédurale (les règles sont opaques, inconsistantes et aléatoires) et injustice interactionnelle (l'application de la plateforme est seulement conçue pour valider la réalisation des étapes de la livraison, sans interlocuteur humain pour répondre aux demandes des livreurs).

Au-delà de la précarité et de l'intensification, le management algorithmique développé par les plateformes de livraison produit ainsi un ensemble de caractéristiques organisationnelles qui créent un climat défavorable à la « confiance organisationnelle », telle que définie par Tan et Tan (2000), à savoir « *le sentiment de confiance qu'a un salarié dans le fait que l'organisation entreprendra des actions bénéfiques, ou du moins non préjudiciables, à son égard*[9] ». Il est assez significatif de ce point de vue que, en parallèle de la plainte au pénal déposée le 22 avril 2026 contre Uber Eats et Deliveroo pour « *traite d'êtres humains* », ces deux plateformes aient été mises en demeure au tribunal judiciaire de Paris, en vue d'une action de groupe pour « *discrimination algorithmique*[10] ».

2. UNE LIVRAISON À DEUX-ROUES ORGANISÉE ET ENCADRÉE PAR DES HUMAINS

Cette partie propose d'éclairer, en miroir des éléments analysés pour les plateformes, une configuration dans laquelle le travail est encadré par un management humain de proximité, afin de mettre en évidence les différences de régulation du travail. Comme nous l'indiquions plus haut, les plateformes de livraison de repas ne sont pas directement comparables aux services antérieurement proposés dans ce domaine – généralement par les restaurants eux-mêmes. C'est donc dans un autre secteur, celui des services de coursiers à deux-roues motorisés existant jusqu'au début des années 2010, que nous pouvons trouver une activité encadrée par des humains qui présente des similitudes suffisantes pour soutenir la comparaison : il s'agit aussi d'un service de livraison entre un point de retrait et un destinataire, dans un délai prédéfini, avec une multitude potentielle de points de retrait et de livraison et de nombreux coursiers à qui affecter les livraisons. Il existe une étude ergonomique sur ce travail (Van Belleghem et Bourgeois, 2004), sur laquelle nous nous appuierons largement dans cette partie, car elle donne à voir le rôle du management (humain) de proximité dans l'organisation du travail des coursiers.

2.1. Des similitudes servicielles, organisationnelles et sociologiques

Les services de coursiers à deux-roues motorisés des années 2000 présentent un certain nombre de points communs avec les plateformes de livraison de repas actuelles.

Tout d'abord, l'organisation du travail était aussi structurée autour du délai client. C'était un paramètre essentiel dans l'affectation des courses aux coursiers, mais aussi dans l'activité du coursier, notamment en raison des aléas auxquels il pouvait être confronté, à l'instar du livreur de repas (difficulté à trouver

9 « *An employee's feeling of confidence that the organization will perform actions that are beneficial, or at least not detrimental, to him or her.* » Notre traduction.

10 Cf. par exemple https://www.lemonde.fr/economie/article/2026/04/23/uber-eats-et-deliveroo-accuses-de-traite-d-etres-humains-la-situation-des-livreurs-ces-proletaires-du-xxi-siecle-a-de-quoi-indigner_6682781_3234.html

une adresse, problème d'accès chez le client, etc.). Cette question du délai était également un enjeu essentiel en termes de risques au travail dans la mesure où l'activité se réalisait essentiellement en circulation urbaine (comme le livreur).

Cette importance du délai et la variabilité de la demande impliquaient par ailleurs que l'entreprise puisse gérer des pics d'activité parfois imprévisibles.

Le profil des coursiers présentait quant à lui un certain nombre de points communs avec celui des livreurs de repas. Comme l'indiquent les auteurs, « *la population qui se destine à devenir et rester coursier est (…) atypique. (…) La plus grande partie du recrutement se fait dans la population jeune des banlieues. (…) Le motif, l'intention, l'envie de travailler tiennent essentiellement dans la facilité et la simplicité de la relation contractuelle (…). On ne pose pas de questions sur le parcours scolaire (en échec le plus souvent). (…) À cela s'ajoutent l'attrait des aspects d'autonomie organisationnelle (par exemple, démarrer sa journée de chez soi en obtenant les courses à faire par téléphone, autrement dit ne pas pointer, symboliquement, tous les jours) et l'attrait du mode de rémunération, totalement déconnecté d'une qualification initiale* » (ibid. p. 21-22). Leur mode de rémunération présentait lui aussi des points communs avec celui des livreurs de repas, car, bien qu'ils fussent salariés, les coursiers étaient rémunérés aussi en fonction du nombre et du type de courses réalisées (cf. infra). Et ils étaient propriétaires du véhicule avec lequel ils réalisaient les courses.

En revanche, la clientèle des sociétés de coursiers était principalement constituée d'entreprises, avec des horaires de travail correspondant généralement à ceux des bureaux.

2.2. De l'« urgence client » au « trajet coursier »

L'organisation des entreprises de coursiers à deux-roues présentait cependant des points majeurs de divergence avec les plateformes de livraison de repas. En effet, pour ces entreprises, l'enjeu était de maitriser *ensemble* la qualité du service client (le respect du délai) et la qualité du « trajet coursier », qui incluait plusieurs paramètres (cf. infra), parmi lesquels l'absence de contrainte de délai perçue par le coursier pour des raisons évidentes de sécurité (éviter que ceux-ci ne s'engagent dans une « course vitesse »).

Pour atteindre ce double objectif, en l'absence d'algorithmes sophistiqués, les entreprises jouaient principalement sur deux variables : la tarification et l'organisation des courses. D'une part, l'adoption d'une tarification différenciée selon le délai de livraison demandé par le client, incitait celui-ci à relativiser la notion d'urgence et à formuler une demande de délai en adéquation avec ses besoins réels (et donc souvent moins contraignante). D'autre part, les courses n'étaient pas réalisées les unes après les autres mais regroupées (plusieurs enlèvements dans le même secteur pour des livraisons dans un autre secteur) ou enchainées (plusieurs enlèvements et livraisons sur une même ligne) par un « dispatcheur » (ou « dispatch »). C'est ici un point de différence majeure avec les plateformes de livraison de repas. En effet, « *la stratégie du dispatch consiste à optimiser au mieux les trajets et l'attribution des courses pour que l'enchainement des enlèvements et des livraisons corresponde à une progression régulière et non précipitée du coursier. (…) Ce processus permet, lors de l'affectation des courses, de ne pas indiquer leur délai : le dispatch est censé l'avoir intégré dans l'enchainement des courses* » (ibid. p. 17). Ce « dispatch » était réalisé par des humains avec des outils relativement simples : cartes papier, téléphones pour joindre les clients et les coursiers et, le cas échéant, échanges directs avec d'autres dispatcheurs de l'entreprise. « *L'organisation de la course-trajet est la réponse trouvée par l'entreprise à la réalisation effective de la prestation de service vendue au client (la livraison du pli) autant qu'à l'optimisation de ses moyens de transport. La mesure de l'efficacité de cette organisation peut donc bien se mesurer à partir de la qualité du service rendu d'une part (le respect du délai pour le client)*

et de la qualité du trajet coursier d'autre part (dont l'absence, théoriquement, de contrainte de délai fait partie). Mais la maitrise de ce double enjeu ne s'obtient pas forcément facilement » (ibid, p. 19).

Les dispatcheurs occupaient dès lors une place centrale dans l'organisation de ces entreprises. Car leurs décisions permettaient de concilier (ou non) les objectifs de qualité pour le client, de performance de l'entreprise (optimisation du nombre de courses réalisées par les coursiers[11]) et de satisfaction pour le coursier (en termes de revenus, de sécurité, de conditions de travail). « *En tant que chefs d'équipe (légitimité par leur statut d'agent de maitrise), ils sont dans une position fonctionnelle, mais aussi hiérarchique vis-à-vis des coursiers : ils prescrivent le travail à réaliser. Concrètement, pour réaliser l'affectation des courses au coursier, ils doivent réaliser au préalable des arbitrages (à qui attribuer telle course ?) dont ils doivent par la suite assumer la responsabilité (faire entendre au coursier que les choix réalisés sont les meilleurs tant pour l'efficacité économique de l'entreprise que pour celle du coursier, mais aussi pour sa sécurité). La qualité de cette interface est fortement marquée par le sentiment de confiance, ou au contraire de manque de confiance, des coursiers vis-à-vis du dispatch, et dont les entretiens font largement part* » (ibid., p. 49).

2.3. Développer la confiance

La contrainte de temps était gérée par le dispatcheur et, pour les coursiers rencontrés dans le cadre de l'étude, n'intervenait pas comme une prescription. En revanche, « *les modes de rémunération d'une part, et les systèmes d'organisation entre dispatcheurs et coursiers d'autre part, se révèlent particulièrement importants dans la gestion sociologique de la population de coursiers et de leurs pratiques de courses* » (ibid., p. 22).

Cette gestion était fondée sur la connaissance interpersonnelle et visait à développer la confiance entre dispatcheurs et coursiers. « *Les coursiers estiment que les dispatcheurs connaissent leurs points forts et faibles. Les dispatcheurs que nous avons rencontrés parlent des coursiers de façon très personnelle, évoquant des logiques de courses-trajet très différentes entre eux. Ils ont une représentation des "façons de rouler" individuelles, leur permettant d'affecter les courses en tenant compte, dans la mesure du possible, des préférences du coursier ("untel aime bien le centre de Paris", "untel n'aime pas trop le périphérique"…) et de leurs compétences particulières ("untel se débrouille bien avec une densité de courses dans un quartier très localisé", "untel ne se perd jamais dans des banlieues éloignées"…)* » (ibid., p. 25). En retour, « *les coursiers évoquent des craintes personnelles ou situations particulièrement détestables à leurs yeux, qu'ils souhaitent éviter. Par exemple, untel exprime sa réticence, voire sa peur du périphérique ou de l'autoroute. Ces peurs, si elles constituent une légitimité, sont importantes à connaitre et à prendre en compte dans le dispatching des courses. Elles sont un élément de la constitution de la confiance entre coursiers et dispatcheurs* » (ibid., p. 28).

Bien entendu, dans l'affectation des courses, « *la maitrise (des) critères liés à la rémunération est importante pour le dispatcheur, parce qu'elle détermine la relation de confiance (ou de manque de confiance) qui s'établira entre lui et le coursier* » (ibid., p. 53). Les coursiers étaient rémunérés sur le principe d'une base fixe et d'une part variable, calculée en fonction du nombre de courses et de leur prix (selon leur degré d'urgence). Leurs attentes dans ce domaine pouvaient donc interférer, voire être contradictoires, avec l'objectif des dispatcheurs d'organiser des « courses-trajet » (sans notion de délai), qui conditionnent la performance de l'entreprise (cf. supra).

11 Les coursiers étant salariés, leur temps d'attente entre deux courses leur est payé. Ce qui fait une différence majeure avec les livreurs des plateformes, dont le temps d'attente n'est pas payé (cf. §21 supra).

Comme le signalaient les auteurs de l'étude, « *le risque que les modalités de rémunération favorisent la survenue de situation à risques est présent, et les entreprises doivent se doter de "garde-fous" pour éviter que ce risque ne se manifeste. (…) Tout dépend alors de la capacité des entreprises à mettre en place ces garde-fous, et à garantir leur efficacité à tout moment (…). Les garde-fous mis en place par les entreprises observées montrent que le management occupe une place centrale dans cet enjeu de prévention, tant dans la relation de confiance qu'il est en mesure de créer avec les coursiers que dans les modes d'attribution des courses aux coursiers* » (ibid., p. 60). En effet, ces entreprises ont fait évoluer le mode de calcul de la part variable de la rémunération des coursiers pour réduire les contradictions potentielles entre la rémunération des coursiers et l'organisation de « courses-trajet » qui répondent aux besoins de prévention des risques et de performance économique. L'une des entreprises observées dans l'étude est même allée plus loin, en spécialisant ses coursiers (grosses cylindrées en banlieue, coursiers attitrés à certains secteurs, etc.), car c'était plus simple et plus efficace pour les dispatcheurs. Mais comme cela impliquait des risques d'iniquité entre les coursiers, et de dégradation de la relation de confiance entre les coursiers et les dispatcheurs, de nouvelles modalités de calcul de la part variable ont été négociées au niveau de l'entreprise pour neutraliser l'impact de ces choix organisationnels sur la rémunération des coursiers.

Le travail des coursiers à deux-roues motorisés s'inscrivait, on le voit, dans une organisation très différente de celle des livreurs à vélo. La confiance entre le management et les coursiers constituait un aspect important du fonctionnement de ces entreprises. Et cette confiance se construisait dans la durée, par un mode de rémunération transparent et par des pratiques d'affectation des courses qui tenaient compte des attentes des coursiers en termes d'équité et de prise en compte des caractéristiques personnelles. Des dimensions éthiques fondamentales (équité, confiance) structuraient donc la relation qui se construisait au fil des interactions entre les coursiers et leurs manageurs. Ces dimensions éthiques pouvaient aussi être formellement mises en avant dans les négociations collectives. Pour leur bon fonctionnement, ces entreprises étaient ainsi conduites à porter une grande attention aux enjeux de justice organisationnelle, dans toutes ses dimensions (distributive, procédurale et interactionnelle), dans leurs pratiques managériales.

— CONCLUSION

L'analyse comparative proposée permet de mieux comprendre en quoi la substitution d'un management humain par un dispositif algorithmique transforme en profondeur les modalités de régulation du travail et leurs effets sur les travailleurs.

Les éléments décrits précédemment pour les livreurs à vélo, en particulier l'absence d'interactions directes, l'opacité des règles d'organisation du travail et la difficulté à élaborer des repères collectifs, peuvent être éclairés par les apports de la clinique du travail. En effet, les algorithmes des plateformes fonctionnent en complète extériorité au monde et à l'expérience du travail. Car « *le travail requiert une mobilisation du corps pour sentir la résistance de la matière ou de l'outil, mais également pour éprouver la perplexité, la peur, le doute, l'angoisse, le plaisir* » (Gernet et Dejours, 2009, p. 28). Comme l'ajoutent les auteurs, « *travailler suppose également de collaborer avec une hiérarchie et des collègues, qu'il va falloir apprendre à connaitre et avec lesquels il faudra pouvoir interagir pour atteindre l'objectif de production d'un bien ou d'un service* » (ibid., p. 28). Et c'est dans ces interactions que les éprouvés du travail permettent d'élaborer et de stabiliser des règles de travail qui ont une visée à la fois pratique et sociale, dans la mesure où elles organisent un « vivre ensemble ». Car « *travailler, ce n'est pas seulement produire, c'est également vivre ensemble* » (ibid., p. 29). C'est pourquoi les dimensions éthiques, qui tenaient une place centrale dans les relations entre les coursiers et leurs manageurs,

sont complètement absentes sur les plateformes de livraison de repas. Au contraire, le fonctionnement des algorithmes des plateformes génère incompréhensions, insatisfactions, sentiment d'injustice (voire de discrimination). L'absence de délibération, tant sur les questions techniques que sur les valeurs éthiques convoquées dans l'exercice du travail, oppose un déni à la réalité de l'expérience vécue par les livreurs dans leur travail. Or, « *quand on est privé de l'échange sur le travail, quand on ne peut plus dire ce qu'on fait, ce défaut de contenance symbolique renvoie à une confrontation en solitaire avec le réel, sans moyen pour l'apprivoiser, le signifier* » (Lhuilier, 2006, p. 91).

La disparition des interactions humaines dans le management du travail sur les plateformes de livraison de repas comporte donc un risque intrinsèque pour la santé des travailleurs, indépendamment des logiques d'intensification, de taylorisation ou de précarisation du travail qui peuvent être portées par les concepteurs des plateformes. Ces constats n'excluent pas l'existence de formes d'ajustement ou d'appropriation locales par les livreurs, ni des variations selon les plateformes et leurs modalités d'organisation, mais ils soulignent malgré tout des tendances structurelles fortes dans la transformation du management du travail.

BIBLIOGRAPHIE

— Anses (2024), *Évaluation des risques sanitaires pour les travailleurs des plateformes numériques de livraison de repas en France* (saisine n° 2021-SA-0045), Maisons-Alfort, Anses, 245 p. https://www.anses.fr/sites/default/files/AP2021-SA-0045RA.pdf

— Bérastégui, P. (2021), « Exposure to psychosocial risk factors in the gig economy: a systematic review, ETUI report 2021.01 ». https://www.etui.org/fr/publications/exposition-aux-facteurs-de-risque-psychosociaux-dans-la-gig-economy

— Fernández-Macías, E., Urzi, C., Wright, S. et Pesole, A. (2023), « The platformisation of work: Evidence from the JRC algorithmic management and platform work survey (AMPWork) », Publications Office of the European Union. http://dx.doi.org/10.2760/801282

— Gernet, I. et Dejours, C. (2009), « Évaluation du travail et reconnaissance », in *Nouvelle revue de psychosociologie*, 8(2), 27-36. https://doi.org/10.3917/nrp.008.0027.

— Gregory, K. (2021), « 'My Life Is More Valuable Than This': Understanding Risk among On-Demand Food Couriers in Edinburgh », in *Work, Employment and Society 2021*, vol. 35(2), 316-331. https://journals.sagepub.com/doi/10.1177/0950017020969593

— Leblanc, É. (2024), « Livreurs des rues. La *mètis* à l'épreuve de la plateformisation du travail », thèse de doctorat, Université Lyon 2. https://theses.hal.science/tel-04866596v2

— Lemozy, F. et Le Lay, S. (2021a), « Le rapport subjectif au travail dirigé par les algorithmes. Être livré à soi-même sur une plateforme capitaliste », in *Mouvements*, 2021/2, n° 106. https://shs.cairn.info/revue-mouvements-2021-2-page-99?lang=fr

— Lemozy, F. et Le Lay, S. (2021b), « Mieux comprendre le vécu des livreurs de plateformes de travail », in Salvoldelli, P. (coord.), *Ubérisation, et après ?*, Paris, Éditions du Détour, 272 p.

— Lhuilier, D. (2006), *Cliniques du travail*, Toulouse, Éditions Érès, 248 p.

— Martin, J. (coord.) (2025), « Algorithmic management and AI-based systems as a new form of work organisation: Psychosocial factors and implications for Occupational Safety and Health ». https://doi.org/10.23775/20250616

— Milanez, A., Lemmens, A. et Ruggiu, C. (2025), « Algorithmic management in the workplace: New evidence from an OECD employer survey », in *OECD Artificial Intelligence Papers*, n° 31, https://doi.org/10.1787/287c13c4-en

— Tan, H.H. et Tan, C.S. (2000), « Toward the differentiation of trust in supervisor and trust in organization », in *Genetic, Social, and General Psychology Monographs*, 126 (2), 241-260.
— Turakia, C. (2017), « Engineering More Reliable Transportation with Machine Learning and AI at Uber ».
https://www.uber.com/en-ES/blog/machine-learning/
— Van Belleghem, L. et Bourgeois, F. (2004), « Le métier de coursier et ses pratiques de prévention. Étude ergonomique pour la prévention des risques professionnels ».
https://www.researchgate.net/publication/282862199_Le_metier_de_coursier_et_ses_pratiques_de_prevention_Etude_ergonomique_pour_la_prevention_des_risques_professionnels
— Zhu, L. et Laptev, N. (2017), « Engineering Uncertainty Estimation in Neural Networks for Time Series Prediction at Uber ».
https://www.uber.com/en-ES/blog/neural-networks-uncertainty-estimation/

LA LIVRAISON DE REPAS : LE JEU ALGORITHMIQUE DES PLATEFORMES

Elisabeth Leblanc
(Anact)

À partir d'une étude de l'activité réelle des livreurs de repas travaillant pour des plateformes numériques, cet article analyse les effets du management algorithmique et son système d'IA sur le travail. Dans ce système de livraison associant plateforme, client, restaurant et livreur, les algorithmes orchestrent l'organisation du travail et encadrent les actions des livreurs par l'intermédiaire de leur smartphone. Classé « à haut risque » par la réglementation européenne dans le domaine des relations de travail, ce mode de management repose sur des règles opaques et évolutives. L'analyse montre que les livreurs doivent constamment chercher à comprendre le fonctionnement du système afin d'atteindre quatre buts d'activité, dont un seul est directement rémunéré. Pour faire face à cette incertitude, ils mobilisent une intelligence pratique de type mètis, faite de ruses, d'ajustements et d'expérimentations. Cependant, cette ingéniosité se heurte à un système algorithmique qui semble l'absorber et la neutraliser. L'article met ainsi en évidence les effets délétères du management algorithmique sur les conditions de travail et la santé de ces travailleurs le plus souvent vulnérables et en très grande précarité.

Mots-clés : livreurs des plateformes, management algorithmique, activité réelle, conditions de travail des livreurs, plateformisation du travail

— INTRODUCTION

Des plateformes proposent une mise en relation entre des offres de restauration et des demandes clients dans la sphère numérique. Puis elles génèrent un service de livraison pour des commandes effectives de clients. Le « *business model design* » des plateformes (Choudary, 2018, p. 2) a pour objectif de gérer des flux et de cadrer l'activité de service. Il est pensé sans interaction humaine (Anses, 2025). Le service de livraison proposé par la plateforme est quant à lui réalisé dans la sphère réelle, par des travailleurs humains pour d'autres humains. Les décisions d'organisation du travail des systèmes algorithmiques conçus par les plateformes sont prises à partir de l'analyse d'un ensemble considérable de données numériques collectées, puis transmises aux travailleurs. La relation de management avec le travailleur est essentiellement numérique et algorithmique. À quoi ressemble l'activité de travail des livreurs des plateformes managés par des systèmes algorithmiques ? Ont-ils des marges de manœuvre ? L'ingéniosité déployée dans toute activité de travail, sans cesse renouvelée dans son

évitement de la souffrance, participe à la santé des travailleurs (Dejours, 2013) : qu'en est-il dans cette situation ? L'étude clinique des interactions du management algorithmique (MA) avec les livreurs de repas des plateformes dessine le réel du travail. L'étude montre comment le système algorithmique d'une telle organisation met à l'épreuve l'intelligence pratique des travailleurs et leur ingéniosité nécessaire à la régulation de leur activité de travail (Leblanc, 2024b).

1. UN ENSEMBLE ORGANISATIONNEL ET FONCTIONNEL

Étudier l'activité réelle des travailleurs des plateformes induit de considérer un large ensemble de facteurs (Leblanc, 2024a et 2024b). Un comité spécialisé « Agents physiques et nouvelles technologies » de l'Anses[1] a constitué un groupe de travail « Livreurs des plateformes » pour procéder à une évaluation des risques sanitaires pour les travailleurs des plateformes numériques de livraison en France (Anses, 2025). Parmi les 14 experts scientifiques pluridisciplinaires sollicités (dont nous faisons partie) pour réaliser l'expertise, on compte deux spécialistes des algorithmes, de l'IA et de la protection des données numériques. Inclure ce qui touche aux nouvelles technologies pour étudier les conséquences sur la santé des travailleurs et leurs conditions de travail s'impose lorsqu'il s'agit de travail proposé ou généré par des plateformes numériques. Les plateformes développent leurs modèles d'affaires et la gestion de leurs activités avec les outils numériques comme le smartphone, la géolocalisation, la 4 et 5G et l'IA. À propos de la livraison, Tassinari et Maccarrone (2019) décrivent une configuration organisationnelle et relationnelle à quatre faces : plateforme, client, livreur, restaurant. Le système de gestion algorithmique de la plateforme (1re face) collecte et traite une masse de données considérable émanant des trois autres faces. Ce système s'enrichit et s'actualise en fonction des objectifs économiques et organisationnels du modèle d'affaires des plateformes. Selon le BIT[2], « *the management techniques used by labour platforms are an integral part of the framework that determines the propensity for workers to be either empowered or exploited*[3] » (Choudary, 2018, p. 6).

Concernant le travail des livreurs, il est donc bien organisé pour répondre à l'efficience des objectifs des plateformes via le management algorithmique (MA). Ce dernier prédit les conduites des travailleurs et prend des décisions automatisées ou semi-automatisées pour organiser individuellement et instantanément le travail d'un très grand nombre de livreurs (Anses, 2025). La rapidité de calcul et le nombre de données lui permettent de modifier et de se réajuster en permanence en fonction du nombre de demandes, de la variation des prix des courses, des intempéries, des exigences clients, etc. Les leviers d'optimisation en faveur du modèle d'affaires des plateformes sont les changements fréquents et rapides des décisions algorithmiques, ce qui les rend totalement opaques pour le travailleur (Choudary, 2018), qui ne peut anticiper son comportement.

Centré sur la planification et l'attribution des tâches, le MA peut déduire des variables sur les activités passées collectées, « *sur le temps de trajet ou sur la productivité, mais aussi effectuer automatiquement des appariements de tâches et de travailleurs afin d'optimiser au mieux des objectifs (…) (maximiser le nombre de courses, minimiser l'attente des clients, durée de la course à atteindre pour le livreur, etc.* » (Anses, 2025, p. 51). Il ne prend pas en compte (en tant que données possibles) ce que cela produit sur les individus et sur le travail. Les ordres sont transmis de façon descendante aux livreurs par le seul smartphone, l'interface du MA.

1 Anses : Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail.
2 BIT : Bureau international du travail.
3 « *Les techniques de gestion utilisées par les plateformes de travail font partie intégrante du cadre qui détermine la propension des travailleurs à être soit responsabilisés, soit exploités* » (notre traduction).

« […] *mon patron – ne s'est pas présenté, je ne sais pas ce qu'il regarde dans mon téléphone, comment il me juge, comment il collecte mes données ; j'ai l'impression que plus je carbure plus j'ai de commandes, j'ai l'impression de me faire pote avec l'ordinateur.* […] *l'ordinateur m'a vu prendre une pause : c'est plus mon pote* » (remarque d'un livreur, Leblanc, 2024b, p. 63).

Le patron est incarné par le smartphone, et son management algorithmique produit six mécanismes comportementaux, les « 6R : *Restricting, Recommending, Recording, Rating, Replacing, Rewarding*[4] », selon Kellogg, Valentine et Christin (2020, p. 367). Le MA limite le choix des travailleurs en n'affichant que peu d'informations ; il incite les travailleurs à faire les choix voulus par la plateforme ; il les enregistre et les surveille ; il mesure et évalue leurs performances ; il déconnecte ou sort les travailleurs et les remplace ; puis il récompense. En aucun cas les livreurs ne peuvent échanger verbalement, discuter ou négocier. Leur ressenti oscille entre incompréhension, injustice et sanction.

Avant d'examiner les effets de ce dispositif technique sur le travail et la santé des livreurs, il convient de préciser les relations entre intelligence artificielle et management algorithmique. L'article 3 du règlement de la Commission européenne daté du 2 février 2025[5] définit un système d'IA et ses fonctionnalités de la manière suivante : « *Un "système d'IA" désigne un système basé sur une machine qui est conçu pour fonctionner avec différents niveaux d'autonomie et qui peut faire preuve d'adaptabilité après déploiement, et qui, pour des objectifs explicites ou implicites, déduit à partir des entrées qu'il reçoit comment générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels.* »

Le MA correspond bien à cette définition. Les algorithmes sont définis comme « *une suite d'instructions – ou de règles – pour parvenir à un objectif* » (Gibert, 2021). « *Des machines de plus en plus puissantes peuvent faire fonctionner des algorithmes exploitant davantage de données (images, textes, vidéos, trajets, etc.) et fournir des traitements plus complexes* » (Anses, 2025, p. 50). Cette complexité et son évolution par apprentissage croissante rendent leur fonctionnement opaque et poussent à les qualifier de « boites noires ». Les formules comme « *algorithmes auto-apprenants, machine learning, réseau de neurones, intelligence artificielle* » (Anses, 2025, p. 50) entretiennent la confusion et le fantasme chez l'humain de la toute-puissance des machines, bien que son « *incarnation technique paradoxale (soit) redoutablement rivale de sa propre impuissance* » (Missonnier, 2020, p. 14). Il est nécessaire « *de ne pas se laisser distraire par la technologie au point d'ignorer les acteurs humains derrière les outils numériques, avec leurs complexités sociologiques, psychologiques et juridiques* » (Anses, 2025, p. 21), et de prendre en compte le classement « à haut risque » des systèmes d'IA dans les relations au travail par le règlement européen[6].

2. L'ACTIVITÉ RÉELLE DE LIVRAISON POUR PLATEFORMES

Rendre compte de l'activité réelle du travail de livraison des plateformes a consisté en l'analyse de 150 échanges écrits de livreurs (1 302 interlocuteurs) sur les réseaux sociaux (RS), collectés entre 2021 et 2023. À cela s'ajoute l'apport de 29 entretiens réalisés entre 2017 et 2023, d'analyses de documents et d'observations diverses. Puis l'analyse détaillée de cinq récits de vie d'une expérience professionnelle

4 Traduction des « 6R » : restriction (d'informations), recommandation, surveillance, notation, déconnexion et récompense.
5 https://artificialintelligenceact.eu/article/3/
6 https://artificialintelligenceact.eu/fr/annex/3/ : « *Les systèmes d'IA destinés à être utilisés pour prendre des décisions concernant les conditions des relations de travail, la promotion ou la résiliation des relations contractuelles liées au travail, pour attribuer des tâches sur la base du comportement individuel ou de traits ou caractéristiques personnels, ou pour surveiller et évaluer les performances et le comportement des personnes dans le cadre de ces relations* » (EU AI Act, annexe III, §4-b).

de la livraison pour plateformes balaie l'activité entre 2014 et 2023[7]. Le croisement des données met en évidence les récurrences de l'activité de travail quels que soient la période et le réel des effets perçus et vécus du « management algorithmique » par les livreurs. L'analyse de nos données prend appui sur les théories de la régulation et plus précisément sur la définition qu'en fait Canguilhem, schématisée par Leplat (2006) sur le modèle cybernétique. Comme représenté sur la figure 1, le livreur réalise une tâche (un but) en fonction d'un résultat estimé dans un cadre donné qui a ses règles. Le but est réévalué en fonction du résultat obtenu et des marges de manœuvre que le travailleur éprouve dans ses possibilités d'action (Clot et Simonet, 2015).

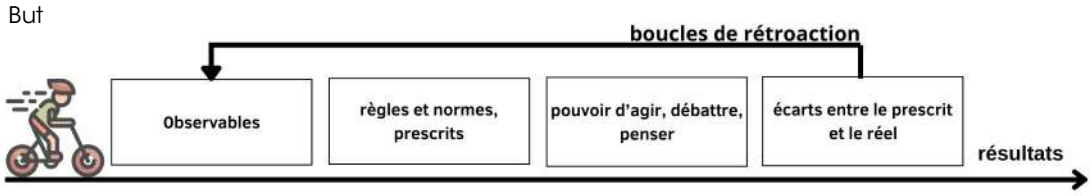


Figure 1. Illustration simplifiée d'une boucle de régulation de l'activité pour livreur (Leblanc, 2024b).

Répondre aux buts dans le cadre de la plateformisation du travail (Leblanc, 2024a) passe systématiquement par l'emprunt d'un chemin défini et imposé par le MA. Les résultats montrent que l'activité de la livraison se compose de quatre buts pour le travailleur (Figure 2), incontournables, indissociables et permanents. Ce découpage a été mis en évidence par le prisme de l'entrepreneur qui souhaite monter, développer et rentabiliser son affaire. En effet, la première condition obligatoire pour obtenir un travail de livreur des plateformes est d'être inscrit au registre des entreprises (micro-entreprise, entreprise individuelle) et donc d'être un travailleur indépendant avant même d'être un livreur. Trois des buts concernent l'activité de l'entrepreneur, et un seul des buts l'activité de livreur. Seul ce dernier génère une rémunération de la part de la plateforme. Ce constat n'est pas anodin dans la façon dont le travailleur va appréhender toute son activité de travail et la réaliser.

But 1 : obtenir le travail (activité d'entrepreneur)
But 2 : s'équiper pour le travail (activité d'entrepreneur)
But 3 : effectuer des livraisons (activité de livreur)
But 4 : rentabiliser son activité (activité d'entrepreneur)

Figure 2. Découpage par buts de l'activité de l'entrepreneur-livreur des plateformes (Leblanc, 2024, p. 232).

Les livreurs interrogés disent que, pour eux, ce travail, c'est être « *comme dans un jeu vidéo avec du vrai argent* » (Leblanc, 2024b, p. 230). Les buts de l'activité de travail pris dans le « jeu algorithmique » des plateformes peuvent-ils être atteints par les travailleurs ? Le premier but consiste à obtenir le travail. Cette étape doit prouver que les démarches préalables ont été effectuées (inscription comme travailleur indépendant, papiers en règle) et que le postulant au travail a déjà un équipement (véhicule, capacité de transport si besoin, sac ou cube). Elle demande également et obligatoirement de s'engager à être surveillé, à transmettre ses données et à accepter de recevoir de la publicité. Le but « obtenir le travail » est incontournable et récurrent de l'activité, car il n'est jamais atteint malgré le travail obtenu. Ce but se représente régulièrement pour réobtenir le travail à la suite de déconnexions

7 Tous les verbatims cités dans les encadrés sont issus de la recherche Leblanc, 2024b. Il est seulement mentionné le numéro de la page pour repère. Les échanges collectés sur les RS sont volontairement retranscrits à l'identique.

faites par la machine, ce qui est relativement courant. La réalisation des autres buts est continuellement sous-tendue par la déconnexion potentielle – qui signifie la perte de son travail – et cela, tout au long de l'expérience de la livraison avec les plateformes, ce qui génère un stress important lié à l'incertitude et à la peur de la sanction.

Ce premier but se fait par interaction directe entre le travailleur humain et les décisions du MA. L'inscription « facile et simple » pour obtenir le travail de livraison mise en avant par les plateformes sur leurs sites est en réalité fastidieuse et complexe. À ce stade de l'activité, l'IA présente dans le dispositif ajuste ses réponses pour chaque entrepreneur-livreur « postulant », de manière différente sans qu'il soit possible de savoir pourquoi, d'où vient l'erreur, s'il y en a une ou quel est le motif d'un refus ou d'une attente par exemple. En aucun cas il n'est possible d'interagir avec un humain en cas de problème. Malgré les bonnes réponses et tous les papiers fournis en bonne et due forme par l'entrepreneur, il n'est pas systématique qu'il obtienne le travail, comme le montrent certains témoignages[8].

> Un livreur partage un message reçu de la part de Deliveroo sur les réseaux sociaux (RS) : « *Deliveroo fait uniquement appel à des indépendants. Malheureusement, votre statut ne vous permet pas de créer une entreprise. Votre inscription a été rejetée.* » Le livreur partage : « *Et mon kbis a été acceptée par uber mais jsais pas c quoi le problème avc deliveroo* ».
> Des internautes-livreurs lui proposent des solutions : « *Tu dois rendre ton kabis visible* » ; « *tap sur Google comment rendre mon kabis difusible* ». D'autres seront plus suspicieux vis-à-vis de l'intention des plateformes : « *C fait exprès* » (p. 237).
> Un autre livreur demande des conseils : « *je remplis leur document comme ils veulent mais ils font que refuser* ».
> À propos de la validation du sac : « *Salut l'équipe, mon sac à été validé depuis ce matin mais je peu toujours pas passé en ligne, est ce normal ?* » Le livreur poste le message donné par la plateforme : « *Malheureusement, votre photo (de votre sac) ne respecte pas nos critères de validation, vous pouvez acheter un sac dans la boutique en ligne* » (p. 240).

Les réponses du système algorithmique sont faites à un moment précis, à partir de l'ensemble des données disponibles (inconnues et invisibles pour le travailleur), ce qui rend les décisions aléatoires, contingentes, incompréhensibles, sans logique et bien sûr sans recours pour le travailleur. « *Ces refus les plongent dans l'impuissance, l'injustice et l'incertitude* » (Leblanc, 2024b, p. 238). Les seules marges de manœuvre possibles pour réaliser ce premier but de l'activité, il les trouve dans l'illégalité, les combines ou le renoncement (Leblanc, 2024b).

> « *ben faut attendre c'est tout ; c'est ce que je me dit car bon Mr robot 🤖 de la messagerie répond toujours la même chose 🤣 j'ai plus cas attendre ...* » (p. 239).

Le deuxième but, « celui de s'équiper », est aussi relatif à la fonction d'entrepreneur. Le futur travailleur, s'il est libre de choisir le véhicule et la tenue qu'il souhaite, est responsable de son équipement, de son entretien, de sa sécurité (matérielle et corporelle) et de la réglementation en vigueur en fonction du véhicule choisi. Le choix du véhicule répond à ses priorisations (cout, vitesse, sécurité, confort, capacités physiques, adaptabilité, etc.). La réponse à ce deuxième but, qui semble laisser une grande place au

8 Les témoignages des livreurs pris sur Internet sont volontairement retranscrits à l'identique dans la thèse par souci d'authenticité (Leblanc, 2024).

choix, dépend indirectement et de façon stricte aux décisions algorithmiques. « *Le moteur s'est imposé "naturellement"* » face aux vélos « *notamment pour des raisons de performances* », explique un livreur (Leblanc, 2024b, p. 241). En réalité, les entrepreneurs-livreurs priorisent le besoin algorithmique pour aller plus vite afin d'avoir plus de commandes. Pourquoi ne pas livrer à vélo ? C'est « écolo », sportif, récréatif, partageaient les anciens livreurs. Indirectement, le MA les a poussés à une auto-accélération permanente sur fond de concurrence-compétition avec les autres livreurs et à opter pour le véhicule (et tout autre matériel) le plus performant en termes de vitesse.

> « *Tous ceux qui avaient des meilleures connexions parce qu'ils avaient le téléphone avec des processeurs un peu plus rapides, ils se retrouvaient avec les commandes avant nous* » (p. 245).

Le troisième but de l'activité de travail est celui d'effectuer des livraisons, c'est le seul qui génère une rémunération payée par la plateforme. Celui-ci concerne le travailleur en tant que livreur et dépend totalement des décisions du management algorithmique. Pour livrer, il doit obtenir une notification via son smartphone. Surveillé en permanence, il doit être « choisi » ou repéré par l'algorithme pour l'obtenir. La préoccupation du livreur (ou le but 3 du jeu) est de limiter au maximum son temps d'attente et d'avoir le plus de commandes possible.

> « *Comment avoir plus de chance d'avoir une course ? Auriez vous un conseil à me donner ou une astuce pour ça ? est ce en fonction de la rapidité du livreur? De sa note globale? Le plus proche du resto? Ou pur hasard* » ; « *Ça c'est la question que tout le monde se pose* » (p. 255).

Les livreurs cherchent à comprendre sur quels critères favorables « aux yeux de l'algorithme » ils pourront obtenir des notifications. Sans sans réponse ni repère, ils testent différents comportements. Il se crée alors des légendes urbaines qu'ils partagent sur les RS.

> « *Pas de secret, bien choisir son secteur ; j'ai un spot stratégique où je me pose ; Repérer une zone ou il y a 5-6 restaurant autour ; moi perso je me pause prêt d'un centre co ; Après minuit* 🕛 » (p. 255). « *Tourne. Ne reste pas sur place ; Ce qui est sure, c'est que Uber préfère les personnes en mouvement , etc. ; C'est une légende urbaine ; En 2ans et demi j'ai jamais réussi à voir une réel différence* » (p. 255). « *C'est à nous après d'être extrêmement rapides pour créer de la disponibilité le plus tôt possible et donc se faire récupérer par l'appli* » (p. 255).
> « *Mettre jaime aux restos que tu vien de livrer et aux client qui on commandé ; mettre des pouces a tout les commerçants car si tu met un pouce en bas tu sera pas prioritaire sur son restaurantn* » ; « *coller un restaurants juste pour avoir plus de commandes ; jouer les sangsus ; faire de la lèche vitrine pour gagner ma vie* » (p. 255). « *Les comptes créé avant Septembre 2018 sonnent mieux et payent mieux* » (p. 256) ; « *Ton temps de connexion et ton ancienneté dans ce boulot en grande partie pour le reste vaste légende ; change de zones ; éloigne toi de 7 km de la* », « *je me suis déplacer a 10km carrément, tjs rien* » (p. 257).

Dès que les livreurs imaginent avoir saisi un critère pour obtenir des notifications, celui-ci varie sans aucune information ou explication, ce qui les désoriente totalement.

Certains livreurs rapportent que des riverains excédés par leurs attroupements aux mêmes endroits ont réagi en accord avec les plateformes : « *À Lyon, la multiplication de "zones blanches" anti-livreurs Uber pour éviter les attroupements devant certains restaurants, dans certaines rues commerçantes où il y a des dépôts de plainte réguliers, parce que les scooters – parce que les discussions un peu fortes, éventuellement des embrouilles qui se développent entre coursiers – ils font en sorte que tu ne puisses pas capter de courses sur cette zone-là, sur un périmètre pour protéger la zone et que personne ne vienne squatter sur cette zone-là* » (p. 252) ; « *Il y avait un jeu avec les plateformes où quand ils voyaient qu'on se mettait tout le temps-là, au final, ils nous, ils nous déplaçaient. C'est-à-dire qu'ils mettaient une zone noire sur la carte, ce qui pour nous était un point stratégique. Et ben eux, ils en faisaient un point où il n'y a pas de commande justement* » (p. 254). Il serait donc possible que l'algorithme accepte des injonctions extérieures pour orienter ses décisions, mais celles des livreurs pour améliorer leurs conditions n'en font pas partie.

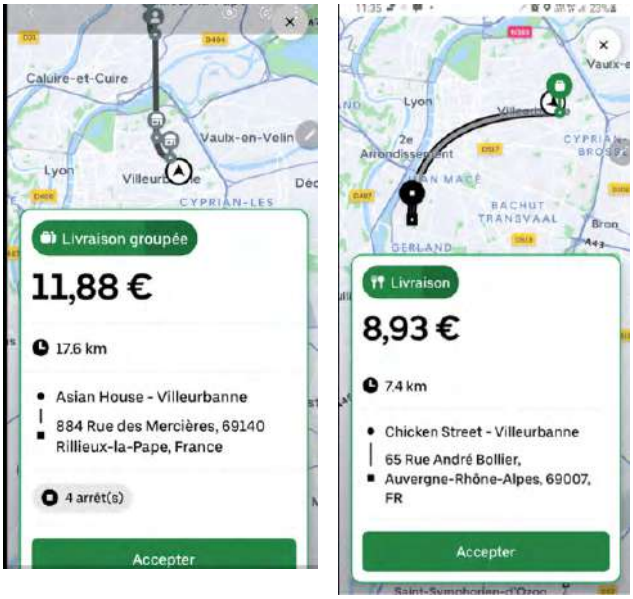


Figure 3. Comparaison des critères de deux notifications. Captures d'écran postées les 14 et 12 décembre 2023.

Une fois que le livreur reçoit une notification, il doit l'accepter (ou la refuser). Les informations que le smartphone transmet motivent sa décision (temps, tarif, trajet). Il a attendu parfois longtemps et constate que les tarifs ne sont pas toujours cohérents, comme les conditions pas toujours acceptables (exemple : Figure 3 *:* 0,67 euro/km pour une course longue ou 1,20 euro/km pour une courte). Mais comme le livreur ne sait jamais ce qu'un refus peut engendrer comme conséquences, il hésite à refuser. Il a peur d'être sanctionné, car il en a souvent fait les frais sans savoir exactement à quoi associer ladite sanction. Les enchainements de causes à effets partagés sur les RS par des livreurs sont un exemple de leur cheminement de pensée (encadré, Leblanc, 2024b, p. 260).

Uber propose une course de presque 18 km à un livreur à vélo pour un tarif d'environ 15 euros. « *Comment tu peut avoir une course aussi loin en velos c'est pas possible il te faudrais 2 h et t'es mort au retour son fou* », réagit un internaute. Un autre anticipe : « *Et si c'est froid t'a un pouce rouge.* » Puis, « *en réalité c'est 87 cts le km* » (p. 259).

Cet exemple montre que le MA dans son calcul ne considère pas l'information concernant le type de véhicule utilisé par ce livreur. Pour un cycliste, cette course serait épuisante, il perdrait du temps et donc de l'argent. Il connait le terrain et sait que le retour n'est pas payé. De plus, si le lieu de livraison est dans une zone éloignée, il n'a que peu de chances d'obtenir une autre course sur le trajet de retour. À vélo, la course longue aura toutes les chances d'être pénalisée par une mauvaise notation de la part du client (repas froid). Logiquement, il devrait la refuser dans son intérêt. Mais au regard de ce que cela peut générer pour l'algorithme, il sera certainement déconnecté et donc perdra son travail. Les livreurs se sentent dénigrés et déconsidérés, ils vivent le rapport au MA comme une injustice.

Autant les livreurs cherchent à se rendre visibles de l'algorithme, autant, dans cet exemple, c'est apparemment l'inverse. Ils cherchent, ils essaient, en mettant en place toutes sortes de combines, d'astuces, mais ne trouvent pas de réponses stables ou durables pour comprendre le fonctionnement du MA. Le MA repère, capte, contourne, rejette tout ce qui ne va pas dans le sens d'une corvéabilité sans faille afin de satisfaire les besoins de son modèle d'affaires. Le livreur sans marge de manœuvre devient la variable d'un ajustement permanent du MA (Leblanc et al., 2019) ; « *Le caractère pathogène du stress est donc lié à la perte du pouvoir d'agir* » (Davezies, 2021, p. 62).

Enfin le quatrième but de l'entrepreneur-livreur est de rentabiliser son activité. Il dépend là aussi des décisions du MA, directement et indirectement. Une notification lui indique une livraison et des conditions fixes imposées (Figure 3). Son intérêt en tant qu'entrepreneur est de prendre en compte la rentabilité de son activité à l'aune de son investissement matériel, physique et temporel. Il est le chef d'entreprise de lui-même en tant que livreur. Son fournisseur d'activité est la plateforme avec laquelle il ne peut rien négocier, ni les tarifs, ni le volume, ni le temps d'une course, ni le trajet, etc. Il n'a pas la main sur les variables qui pourraient faire augmenter son chiffre d'affaires. Il constate les incohérences des tarifs qui ne semblent pas considérer l'effort, la pénibilité, la durée du trajet, les horaires, les dénivelés, et qui peuvent varier de façon notable pour des courses similaires. Il se sent lésé. Il se plie au bon vouloir et au rythme du MA, tout en cherchant désespérément des solutions pour rentabiliser ses affaires, qu'il ne trouve jamais. Pour cela, il teste, par exemple, la rapidité sur la route et ne respecte pas les règles urbaines, il néglige la qualité de son travail pour aller plus vite, il se rend disponible continuellement. Il s'épuise, se place dans l'illégalité et se met en danger pour une rémunération aléatoire et faible. Il souffre en silence (Leblanc et al., 2019).

> « *Course à 1,64 euros pour 0,9km ; 13,68 pour 14,7km* » ou « *4.72 euros la double comande en pleine nuit* » (p. 259) ; « *C'est clairement un manque de respect moi j'ai arrêté de travaillé avec eux ça vaut vraiment pas le coup ! Si on compte le temps perdu l'essence pour certains et tout. 17 km pour même pas 5e* 🤦 *; Pareil pour moi hier soir (…) ; Et des 3 km à 2,67 € on en parle ! Presque 1.50€ de l'heure meme en chine tes mieux payé. L'esclavage putinnnn* » ou « *Moi ça va j'arrive à me faire 40€ voir 50€ par soir* » (p. 287) ; « *Pedalez plus pour qu'ils gagnent plus* 😊 » (p. 290).

Pour chacun des buts présentés de l'activité, ces situations décrites et ressenties par les entrepreneurs-livreurs sont permanentes. Les livreurs évoquent régulièrement des *bugs* de la machine pour se rassurer. Ils ont du mal à intégrer que le MA ne les prend en compte que dans son intérêt strict, sans état d'âme, sans humanité. Le MA n'organise pas le travail de l'entrepreneur-livreur mais le désorganise.

3. LA FAUTE À L'IA ?

Les livreurs des plateformes, ces travailleurs du réel, se heurtent au monde virtuel comme on cognerait inlassablement une vitre. Ce monde virtuel qu'ils ne peuvent comprendre, ils l'alimentent et l'entretiennent avec leurs propres données issues de leur réel. Un jeu où, à chaque seconde, les règles changent pour le livreur et où la notion de règle d'organisation du travail ne semble pas exister pour le système. Nos résultats montrent que l'activité de livraison des plateformes inclut la quête incessante d'une stabilité. Elle se caractérise par le fait de chercher en permanence comment elle s'organise et quelles sont les règles « du jeu ». Pour cela, ces travailleurs mobilisent leur intelligence pratique, habile, rusée, inventive et créative (Leblanc, 2022 et 2024b). Cette forme d'intelligence et de pensée communément mise au service de sa santé dans la réalisation de son activité de travail est incarnée par la mètis des Grecs (Dejours, 2013 ; Detienne et Vernant, 2018). Elle est ici mobilisée et éprouvée sans fin pour comprendre l'organisation du travail. Une des caractéristiques de la mètis est de précéder le savoir. Elle se découvre et s'éprouve par les « *ficelles de métier qu'on ne peut pas (…) apprendre* », c'est une intelligence enracinée dans le corps (Dejours, 2017, p. 29). Chez les livreurs des plateformes de notre étude, l'intelligence « mètis » se caractérise par cinq expressions récurrentes oscillant chacune sur un axe. Dans ce système instable, elle s'exprime sur l'axe de l'illusion (gagner de l'argent, gérer son temps, être libre, autonome, etc.) et de la désillusion (déceptions en chaine, perte de confiance, incertitude, précarisation) ; sur l'axe des relations incarnées (réagir avec des restaurateurs, clients, passants, etc.) et désincarnées (avec le MA) ; sur l'axe de se rendre visibles ou invisibles des algorithmes en fonction de leurs croyances (culpabilité, désorientation), mais aussi des clients, des restaurateurs, de la police, etc. ; sur l'axe de la légalité (se respecter, éviter les ennuis) et de l'illégalité (jouer avec les règles, les contourner pour gagner en pouvoir d'agir et pour s'en sortir) ; et enfin sur l'axe organisation/désorganisation en lien avec leurs multiples tentatives dans un système incompréhensible (Leblanc, 2024b). Le MA semble capter la mètis du travailleur pour en tirer profit jusqu'à l'aliénation totale de l'individu au système.

CONCLUSION

Ce MA est maltraitant et a toutes les chances d'engager les personnes fragiles ou fragilisées (comme les sans-papiers) vers des pathologies physiques et mentales lourdes. En ce sens, la plateformisation du travail (définition : Leblanc, 2024a) pour cette activité participe de la dégradation des conditions de travail et de la santé des travailleurs. L'intelligence « mètis » des entrepreneurs-livreurs est éprouvée par le MA qui finit par l'absorber dans ses calculs pour l'optimisation de son système rentable.

Les systèmes algorithmiques sont-ils réellement mis au service de l'humain afin de faciliter le travail ? Il est démontré, dans cette analyse de l'activité réelle des livreurs, que le système managérial algorithmique élaboré par des humains pousse d'autres humains (les travailleurs en bout de chaine) à se dispenser de leur intelligence « mètis » au profit des premiers pour mieux les asservir. Bernard Stiegler (2008) invite à penser l'objet numérique comme *pharmakon*, avec son pouvoir de guérison ou d'empoisonnement ; nous invitons à réfléchir à la plateformisation du travail de la même façon. Pour la livraison de repas des plateformes réalisée par des travailleurs indépendants, nos résultats montrent que les trois quarts de leur activité concernent celle de l'entrepreneur mais que la totalité de leur activité est organisée et imposée par les plateformes. Est-il possible de libérer le travailleur de cette forme d'indépendance-dépendance managée par les algorithmes ? Plus globalement, comment réguler avec les plateformes et leurs algorithmes ?

BIBLIOGRAPHIE

— Anses (2025), « Évaluation des risques sanitaires pour les travailleurs des plateformes numériques de livraison de repas en France » (saisine n° 2021-SA-0045).
— Choudary, S.P., (2018), « The architecture of digital labour plateforms: Policy recommandations on plateform design for worker well-being », International Labour Office.
— Clot, Y. (2008), *Travail et pouvoir d'agir*, Paris, Presses universitaires de France, 324 p.
— Clot, Y. et Simonet, P. (2015), « Pouvoirs d'agir et marges de manœuvre », in *Le Travail humain*, 78(1), 3152. https://doi.org/10.3917/th.781.0031
— Davezies, P. (2021), « L'activité entre assujettissement et subjectivation », in Dujarier, M-A., Gillet, A. et Lénel, P. (coord.), *L'Activité en théories : regards croisés sur le travail* (p. 47-86), Toulouse, Éditions Octarès, 360 p.
— Dejours, C. (2013), *Travail vivant 1 : Sexualité et travail*, Paris, Petite bibliothèque Payot, 224 p.
— Dejours, C. (2017), « L'engagement du corps dans l'intelligence à l'épreuve du travail vivant », in Amado, G., Bouilloud, J.-P., Lhuillier, D. et Ulmann, A.-L. (dir.), *La créativité au travail* (p. 61-80), Toulouse, Éditions Erès, 408 p.
— Detienne, M. et Vernant, J.-P. (2009), *Les ruses de l'intelligence. La mètis des Grecs*, Paris, Flammarion, 464 p.
— Gibert, M, (2021), *Faire la morale aux robots*, Paris, Éditions Climats, 168 p.
— Kellogg, K.C., Valentine, M.A. et Christin, A. (2020), « Algorithms at work: The new contested terrain of control », in *Academy of Management Annals*, 14(1), 366- 410.
— Leblanc, É., Cuvillier, B. et Rouat, S. (2019), « Coursier à vélo de l'uberisation : une profession qui souffre en silence », in *La Revue des conditions de travail*, 9, 44-51.
— Leblanc, É (2022), « L'ingéniosité du coursier à vélo de l'ubérisation », in *Canal Psy.*
— Leblanc, É. (2024a), « La plateformisation du travail », in Valléry, G., Bobillier Chaumon, M.-É. et Brangier, É. (dir.), *Psychologie du travail et des organisations : 120 notions clés* (p. 346-349), Malakoff, Dunod, 496 p.
— Leblanc, É. (2024b), « Livreurs des rues. La *mètis* à l'épreuve de la plateformisation du travail », thèse de doctorat présentée et soutenue le 21 novembre 2024, Université Lumière Lyon 2/Anact.
— Leplat, J. (2006), « La notion de régulation dans l'analyse de l'activité », in *Perspectives interdisciplinaires sur le travail et la santé*, 81. https://doi.org/10.4000/pistes.3101
— Missonnier, S. (2020), « La mutinerie annoncée des robots : transitionnalité d'un double créatif ou syndrome de Münchhausen par procuration ? L'humain et la technologie », in *Canal Psy,* 125, 13-20.
— Stiegler, B. (2008), *Réenchanter le monde*, Paris, Champs essais, 178 p.
— Tassinari, A. et Maccarrone, V. (2019), « Riders on the Storm: Workplace Solidarity among Gig Economy Couriers in Italy and the UK », in *Work, Employment and Society*, 34(1), 3554. https://doi.org/10.1177/0950017019862954

LA RADIOLOGIE À L'ÈRE DES « NOUVEAUX » SYSTÈMES D'INTELLIGENCE ARTIFICIELLE : USAGES ET TRANSFORMATIONS DU TRAVAIL

Tamari Gamkrelidze*, Moustafa Zouinar et Flore Barcellini***

(*CNAM - Centre de recherche sur le travail et le développement (CRTD), **Orange)

L'intégration et l'utilisation des systèmes d'intelligence artificielle (SIA) actuels dans les activités professionnelles suscitent de vives questions quant à leurs conséquences sur le travail. La radiologie est particulièrement concernée, car l'IA est utilisée dans ce secteur depuis longtemps. Il constitue ainsi un cas d'analyse exemplaire. Cet article présente une étude empirique des usages de deux SIA et les transformations du travail associées, à partir d'observations du travail et d'entretiens. Les résultats mettent en évidence une variabilité des transformations et des conséquences duales liées à l'utilisation des SIA, en fonction des contextes d'usage, des choix organisationnels et des propriétés des systèmes. Ils remettent en question les discours souvent techno-déterministes et polarisés caractérisant l'IA comme soit bénéfique, soit préjudiciable en radiologie, et soulignent la nécessité de tenir compte du caractère situé des usages et des transformations liées à l'IA. Les résultats montrent également la nécessité de mobiliser une approche systémique pour appréhender les conséquences de l'IA dans les contextes de travail.

Mots-clés : radiologie, intelligence artificielle centrée humain, usages, travail

— INTRODUCTION

La diffusion et les usages croissants dans les activités professionnelles des systèmes d'intelligence artificielle (SIA) – basés sur des techniques d'apprentissage automatique de plus en plus performantes, telles que l'apprentissage profond (Goodfellow et al., 2016) – suscitent de vives interrogations quant à leurs conséquences sur le travail et l'emploi.

Le secteur médical est particulièrement concerné (Mignot et Schultz, 2022 ; Vallès-Peris et Domènech, 2023), avec une multiplication de SIA destinés à ce secteur, une diversité d'usages potentiels annoncés et des promesses de bénéfices associés à leur introduction (Kempt et al., 2023 ; Mignot, 2025). La radiologie se distingue comme un domaine emblématique, en raison de l'adoption élevée des SIA (Machado et al., 2025) et leurs implications fortes sur le travail des radiologues (Ranschaert, 2018 ; Stoeklé et al., 2021). Discipline historiquement marquée par l'informatisation (Brunelle et Brunelle, 2019 ; Mayo et Leung, 2018 ; Tang et al., 2018), la radiologie a connu toutes les « ères » de l'IA (Fujita, 2020 ; Oakden-Rayner, 2019). Par exemple, les systèmes de dictée à reconnaissance vocale (RV), déployés dès les années 1980 (Al-Aiad et al., 2020), se sont largement diffusés depuis les années 2000 pour réduire les couts et les délais de réalisation des comptes rendus (CR) radiologiques (White, 2005). Plusieurs études soulignent des gains dans la mise à disposition rapide des CR, améliorant la prise en charge des patients (Krishnaraj et al., 2010 ; Ringler et al., 2017), mais aussi des effets négatifs comme la surcharge de travail liée à la correction des erreurs de transcription, une certaine insatisfaction et frustration des radiologues, ou encore la suppression d'emplois de transcripteurs (Al-Aiad et al., 2020 ; Pezzullo et al., 2008 ; Mayo et Leung, 2018 ; Ramaswamy et al., 2000 ; White, 2005).

Les CAD (*computer-aided diagnosis*, diagnostics assistés par ordinateur) représentent un autre type de système. Initialement conçus pour la mammographie et la détection de nodules pulmonaires (Castellino, 2005 ; Balleyguier et al., 2005), ils visaient à améliorer la précision diagnostique, en étant considérés comme « seconds lecteurs » (Shiraishi et al., 2011). Si certaines études ont montré une amélioration des performances dans le diagnostic médical (Sahiner et al., 2009), d'autres ont mis en évidence leurs limites : faux positifs, perte de spécificité, temps d'interprétation accru et faux sentiment de sécurité (Fujita, 2020 ; Kohli et Jha, 2018 ; Oakden-Rayner, 2019). La période actuelle de l'IA (Chartrand et al., 2017) relance les promesses de gain de temps, de performance ou encore d'amélioration des diagnostics avec des CAD plus puissants. Ces promesses se renouvellent progressivement en s'alignant sur les préoccupations des radiologues, qui participent ainsi à la reformulation des promesses (Arena et al., 2024). Elles nourrissent aussi des interrogations majeures : le métier de radiologue est-il destiné à être automatisé (Chockley et Emanuel, 2016 ; Obermeyer et Emanuel, 2016), transformé (King, 2018 ; Moehrle, 2018) ou « augmenté » (Liew, 2018 ; Lincoln et al., 2019) ? Après la prédiction de Geoffrey Hinton qui avait déclaré en 2016 qu'il fallait cesser de former les radiologues car l'apprentissage profond deviendrait meilleur qu'eux dans cinq ans[1], prédiction qui avait fortement inquiété les radiologues (en particulier les plus jeunes), la question du remplacement total des radiologues par des SIA parait de moins en moins réaliste et laisse place à la perspective « augmentative ». Plusieurs études cliniques et expérimentales (Oppenheimer et al., 2023 ; Reverberi et al., 2022) examinent ainsi l'efficacité de la « collaboration » entre les radiologues et les SIA par rapport à leur travail individuel, sans SIA. Elles montrent globalement que l'IA peut améliorer la précision du diagnostic et réduire la charge de travail lorsqu'elle assiste les radiologues, avec des bénéfices plus marqués pour les débutants ou moins expérimentés. Toutefois, ces gains dépendent de la fiabilité des systèmes et de la confiance accordée, une sur- ou sous-confiance pouvant compromettre l'efficacité (Reverberi et al., 2022). La majorité de ces travaux expérimentaux et centrés sur une vision partielle de la performance (gain de temps, rapidité, etc.) ne permettent pas d'analyser les transformations des métiers, car ils sont souvent réalisés hors contexte de travail, et sont liés à des scénarios théoriques d'utilisation de l'IA (Vallès-Peris et Domènech, 2023). De plus, ils n'examinent pas la réalité vécue du travail (avec ou sans IA), des usages réels et leurs conséquences organisationnelles (Anichini et Geffroy, 2021 ; Vallès-Peris et Domènech, 2023). Ils ne tiennent pas non plus compte de la dimension collective du travail.

1 https://www.youtube.com/watch?v=2HMPRXstSvQ

Dans les centres hospitaliers comme dans les cliniques, le travail repose sur de multiples interactions entre les acteurs du soin. Radiologues, secrétaires et personnels administratifs coopèrent, s'échangent des avis et s'entraident selon les situations. La prise en charge des patients résulte ainsi d'un travail collectif, dont la dimension est essentielle à considérer.

Néanmoins, les études empiriques sur l'introduction, les usages de SIA et les transformations du métier de radiologue en contextes réels se développent (Anichini et Geffroy, 2021 ; Gamkrelidze, 2022). Elles offrent de comprendre comment les SIA sont utilisés, et comment leur déploiement affecte les activités des travailleurs et l'organisation du travail. S'inscrivant dans cette perspective, cet article présente une étude qui avait pour objectif d'analyser les usages et les incidences organisationnelles sur le travail de deux SIA en radiologie.

— 1. MÉTHODOLOGIE

Cet article présente une étude empirique menée avec une approche ergonomique du travail (Daniellou et Rabardel, 2005). Deux systèmes ont été étudiés dans des contextes socio-organisationnels différents :

— Un système de dictée à reconnaissance vocale (RV) utilisé pour la rédaction des comptes rendus (CR) radiologiques mis en place dans un service d'imagerie médicale d'un hôpital universitaire. Ce système emploie des techniques de reconnaissance vocale à base d'IA pour transcrire du texte dicté par un radiologue (Figure 1). Il est constitué d'un logiciel de reconnaissance de la voix, d'une interface de traitement de texte et d'un micro filaire, et est intégré dans un système d'archivage et de transmission d'images (PACS).

Figure 1. Utilisation de la dictée RV par un radiologue

— Un système de détection des fractures (DF) sur des images radiographiques déployé dans les services d'imagerie et d'urgence d'hôpitaux publics et de cliniques privées. Ce système, basé sur l'apprentissage profond, analyse automatiquement les radiographies osseuses pour détecter d'éventuelles lésions et fournit trois types d'informations en fonction de son « niveau de certitude » : absence de fracture (Figure 2), une ou plusieurs fractures probables (Figure 3), ou la présence de fractures (Figure 4). Dans certaines structures médicales, le système est également disponible pour les médecins urgentistes qui ne sont pas des radiologues, ce qui fera l'objet d'un point de discussion.

Figure 2. Résultats du système DF : absence de fracture

Figure 3. Résultats du système DF : fracture probable (doute)

Figure 4. Résultats du système DF : présence d'une fracture

Le travail de terrain sur le système de dictée RV s'est déroulé en deux phases. La première (de décembre 2019 à septembre 2020) visait à comprendre le fonctionnement global du service de radiologie à travers neuf entretiens semi-directifs (radiologues, manipulateurs, cadres, secrétaires, informaticien) et 30 heures d'observation *in situ*. La seconde (septembre 2021) s'est focalisée sur l'usage de la dictée, l'effet sur les pratiques et les incidences organisationnelles, à partir d'entretiens avec quatre radiologues et trois secrétaires, complétés par des observations et enregistrements vidéo des activités de rédaction des comptes rendus.

Pour le système DF, des entretiens semi-directifs (cinq radiologues et deux médecins urgentistes, de 5 à 30 ans d'expérience) et des observations *in situ* (20 heures d'observations libres et 6 heures d'observations filmées avec verbalisations simultanées qui étaient spontanées ou provoquées[2]) ont été réalisés (juillet 2020 à juillet 2021) dans six cabinets-cliniques privés et six centres hospitaliers. L'objectif était d'analyser le processus d'introduction du système, d'identifier les intentions de sa mise en place et d'évaluer ses conséquences sur le travail, en se centrant sur les usages des radiologues, les urgentistes n'ayant pas pu être observés.

L'ensemble des données a été analysé avec MAXQDA en s'inspirant de la théorie ancrée (Paillé, 1994), et structuré autour de thèmes principaux (comme le contexte de déploiement du système, son implémentation dans l'organisation du travail, ses caractéristiques, les usages du système et les conséquences liées à son déploiement), issus à la fois de la littérature et du matériau empirique. Les enregistrements vidéo ont été analysés en articulation avec ces thèmes, notamment concernant l'activité et l'usage des SIA.

— 2. RÉSULTATS ET DISCUSSION

Les résultats mettent en lumière la manière dont les SIA sont mobilisés comme une « solution » à certaines difficultés socio-organisationnelles dans les services de radiologie. Ils mettent par ailleurs en évidence la diversité et le caractère systémique des transformations des activités et de l'organisation du travail liées à l'introduction des SIA, marquées par une dualité entre bénéfices et risques.

2 Cela signifie que le chercheur questionnait spécifiquement sur les événements survenus pendant l'observation.

2.1. Des conduites de projet et des déploiements de l'IA principalement orientés « solution »

Le domaine de la radiologie se voit confronté aujourd'hui à l'augmentation du volume d'examens que ce soit dans le secteur privé ou public, associée à une baisse de moyens, notamment humains. Face à ces problèmes, l'IA est vue comme une solution pour « faire avec » moins de professionnels (radiologues, secrétaire médicale, etc.).

Le projet de mise en place de la dictée RV en est une illustration. Face à la charge de travail élevée des secrétaires relative à la transcription des CR radiologiques, dictés et enregistrés par les radiologues, un projet de réorganisation du secrétariat (plus largement du service) est élaboré. Du point de vue de la conduite de projet, cette réorganisation a été accompagnée par la mise en place de groupes de travail impliquant les professionnels concernés, dont les radiologues et les secrétaires, pour penser une nouvelle organisation du secrétariat. Le déploiement de la dictée RV dans le service d'imagerie coïncide avec ce projet de réorganisation, et elle a été intégrée dans la nouvelle version du système d'archivage et de transmission d'images (PACS), déployée au moment du projet de réorganisation du secrétariat. Le projet implique donc à la fois des changements organisationnels et techniques (introduction de la dictée RV et nouvelle version du PACS). La dictée RV était vue comme une solution technique aux retards importants de livraison des CR, retards découlant de l'augmentation de la charge de travail. Comme le disent les secrétaires elles-mêmes, « l'option technologique » a ainsi été préférée à « l'option humaine » qui aurait pu être de recruter davantage de secrétaires. Ce choix relève ainsi d'une stratégie typiquement techno-solutionniste. Au-delà de ce cas, cette stratégie est aujourd'hui renforcée par les promesses d'automatisation qui sont attachées à l'IA. En termes de conduite de projet, bien que les secrétaires aient été impliquées dans les transformations du service à travers des groupes de travail, leur point de vue critique a eu peu de valeur, du fait que la décision de déployer la dictée RV avait été prise en amont.

La même perspective techno-solutionniste se retrouve dans le cas du système DF dont l'introduction, notamment dans les structures hospitalières, est également envisagée comme un moyen pour faire face à l'augmentation de la charge de travail due au manque de radiologues, et à ses conséquences en termes de dégradation des conditions de travail, de risque sur la santé, de risques d'erreurs et de baisse des performances dans l'interprétation des images – entendue uniquement en termes de temps de délivrance des CR –, et conséquemment, de la prise en charge des patients. Au moment de l'étude, les systèmes avaient été déployés à titre expérimental dans le cadre d'une utilisation opérationnelle, sachant que les radiologues et les urgentistes qui pouvaient y avoir accès dans certaines structures étaient libres d'utiliser ou non le système de DF. Contrairement à la dictée RV, les déploiements de ce système n'étaient donc pas réalisés dans le cadre de projets bien définis. Si, dans le premier temps, le système s'adresse aux radiologues, il est peu à peu mis à disposition des médecins urgentistes – particulièrement confrontés aux cas de traumatologie alors qu'il manque des radiologues aux urgences. Ces urgentistes sont fréquemment les premiers à devoir analyser les radiographies (notamment les nuits ou les week-ends, car les radiologues ne sont pas présents), voire les seuls dans certaines structures particulièrement concernées par ce manque.

Par ailleurs, suivre l'évolution et l'innovation liées à l'IA est vu comme un impératif pour rester compétitif et attractif dans le domaine de la radiologie, afin de ne pas être « *à la traine des évolutions* » (Radiologue 14). On voit apparaitre ici un autre objectif à l'origine des projets d'introduction des SIA : rester dans la course technologique, qui peut apporter des avantages compétitifs.

Dans les deux cas, les problèmes centraux, comme le manque structurel de personnel, restent cependant non traités, avec le risque que des conséquences – une intensification, par exemple – se reportent sur d'autres services ou personnels. En effet, le caractère systémique des problèmes organisationnels est souvent peu pris en compte dans les stratégies techno-solutionnistes.

Ces résultats illustrent comment l'IA apparait comme une forme de « solutionnisme technologique » (Morozov, 2013), qui met en avant la technique comme unique remède aux problèmes socio-organisationnels complexes qui touchent le secteur de la santé. Le recours à l'IA, en tant que « prothèse organisationnelle » (Gamkrelidze, 2022) ou « mesure palliative » (Mignot, 2025), est en priorité pensé dans les secteurs médicaux défaillants avec des conditions de travail dégradées.

2.2. De la dualité bénéfices/risques du travail avec l'IA en radiologie

Globalement, les radiologues associent les SIA à plusieurs bénéfices pour leur travail dans la rédaction de CR ou d'interprétation de radiographies. Cependant, l'étude révèle que le déploiement et l'usage de SIA en radiologie n'est pas sans risques et peut, dans certains cas, générer des effets problématiques non seulement pour les radiologues, mais également pour les secrétaires. Cette dualité bénéfices/risques s'observe dans les deux cas étudiés.

2.2.1. Le cas de la dictée RV : gain de temps vs appauvrissement du travail, intensification et entrave à la construction collective de la sécurité de soins liés au SIA

Le système de dictée RV est utilisé pour la réalisation d'un CR radiologique dont l'objectif est de donner « *tous les éléments utiles à la prise en charge de la pathologie suspectée* » (Radiologue 7). Pour les radiologues, un « bon » compte rendu radiologique est concis, répond précisément aux questions posées, sans ambiguïté quant au niveau de certitude du radiologue. La qualité d'un CR est évaluée sur des critères fonctionnels (pertinence et clarté du contenu, concision ou richesse des formulations, continuité et rigueur du diagnostic) et esthétiques (orthographe, style, mise en page). Ces dimensions, à la fois médicales et formelles, conditionnent la précision du diagnostic et la transmission efficace des informations issues des examens.

Avant la mise en place de la dictée RV, les radiologues dictaient leurs CR, que les secrétaires transcrivaient, dont elles vérifiaient la qualité en fonction de ces critères et qu'elles imprimaient pour validation et signature par les radiologues. La production d'un CR de qualité relevait donc d'une activité collective distribuée entre secrétaires médicales et radiologues. La mise en place de la dictée RV a conduit à repenser cette organisation du travail. Désormais, c'est le système qui effectue la tâche de transcription, et le radiologue corrige et vérifie souvent seul son CR. Les secrétaires n'interviennent plus que dans la phase d'envoi des comptes rendus (Figure 5).

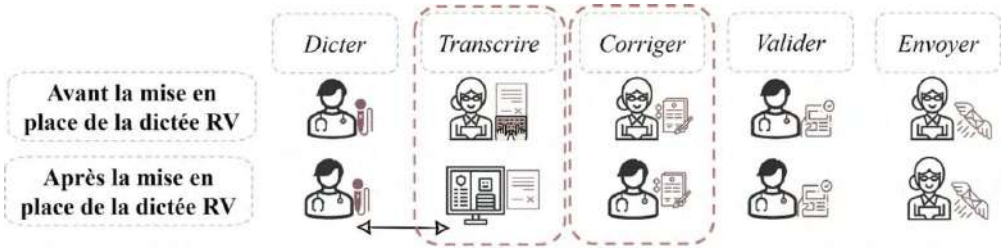


Figure 5. Comparaison de l'organisation du travail avec et sans dictée RV

La tâche de correction des CR aurait pu être laissée aux secrétaires, ce qui avait été fait pendant une courte période, mais afin d'atteindre les objectifs de gain de temps, il avait été décidé que le radiologue corrigerait directement son compte rendu sans passer par la secrétaire. De plus, certaines secrétaires relatent que relire un CR à la recherche des fautes sans l'avoir transcrit était très fastidieux, inintéressant et dévalorisant pour elles : « *Complètement inintéressant quand vous êtes là, en train de relire le compte rendu d'un médecin, franchement, on est là pour retrouver les fautes, corriger les fautes. Je vous avoue que… la déprime totale (…) et puis même, à force de chercher la faute, on ne la trouve plus* » (Secrétaire 4).

L'introduction de la dictée RV associée à cette transformation de l'organisation du travail s'est traduites par plusieurs conséquences, certaines vécues comme positives, d'autres comme négatives (Figure 6).

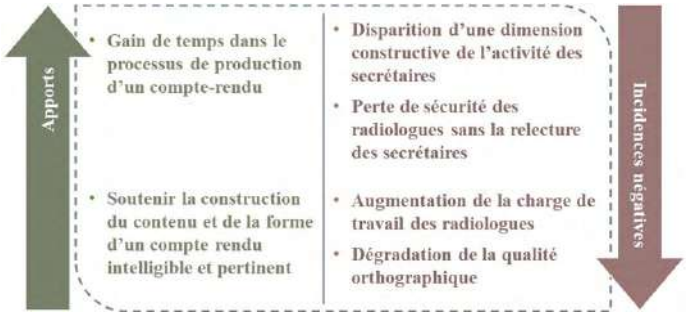


Figure 6. Bénéfices et incidences problématiques liés au système de dictée RV

Du côté des bénéfices, elle est vue comme améliorant la qualité des CR et autorisant un gain de temps dans leur production. Plus précisément, dans cette activité complexe et itérative qu'est la production d'un CR, l'usage de la dictée RV a plusieurs avantages selon les radiologues :

— en permettant de voir, d'organiser et de réorganiser le contenu du CR en temps réel, l'usage de la dictée permet de construire des CR clairs et pertinents ;
— elle aide à structurer et systématiser plus efficacement l'analyse radiologique en donnant la possibilité de modifier le CR au fur et à mesure de l'analyse des images ;
— l'utilisation de CR types, avec un contenu prédéfini, est devenue plus accessible avec l'introduction de la dictée RV ;
— l'utilisation de la dictée RV, combinée à une nouvelle répartition des tâches, semble offrir la possibilité d'une validation plus rapide du CR, ce qui évite les oublis potentiels et contribue à améliorer la qualité du diagnostic.

Du côté des incidences problématiques et des risques, le déploiement et l'utilisation du système affectent différents aspects du travail à la fois des radiologues et des secrétaires (Figure 6).

Tout d'abord, la dictée RV soulève plusieurs difficultés d'utilisation. Les radiologues doivent être attentifs aux erreurs d'orthographe ou de transcription de la dictée qui peuvent parfois modifier le sens des phrases ou les rendre inintelligibles, ce qui peut compliquer la tâche de correction en créant des incompréhensions ou entrainer une perte de temps et de l'agacement chez les radiologues. Ainsi, les radiologues ont tendance à produire des CR plus courts afin de réduire le risque d'erreurs de transcription. La dictée RV peine aussi à transcrire un langage naturel, obligeant les radiologues à adapter leur style de dictée (phrases courtes, rythme artificiel) et en surveillant la transcription du texte. Cet inconvénient est accepté comme un compromis pour une meilleure qualité de transcription : « *Quand on veut assurer, on parle un peu comme des robots. (…) En découpant bien, ça marche mieux. C'est un peu agaçant d'être obligé de parler comme ça. Mais on y gagne, au final, en qualité* » (Radiologue 1). L'utilisation du système implique donc une transformation non négligeable du travail des radiologues dans la production des CR.

Avec la dictée RV, la réorganisation du travail a supprimé la transcription des CR par les secrétaires. Cette suppression a deux conséquences. Premièrement, elle affecte les secrétaires dans la mesure où elle fait disparaitre une tâche qui était une source de développement de connaissances et d'enrichissement professionnel importante pour ces dernières. La rédaction des CR leur permettait d'élargir leurs connaissances médicales. Deuxièmement, elle se traduit, pour les radiologues, par une perte d'un

niveau de contrôle des CR. Alors qu'auparavant, la pertinence et la cohérence des CR faisaient l'objet d'un travail collectif avec un double regard – le radiologue puis la secrétaire, qui pouvait signaler des incohérences ou des formulations ambiguës –, cette étape de vérification repose désormais sur le seul radiologue. Cette disparition d'une « *barrière humaine de sécurité* » accroit le risque d'erreurs, oblige les radiologues à relire attentivement et longuement leurs CR, et modifie en profondeur l'organisation et leur charge de travail.

Si l'idée de substitution des médecins tels que des radiologues par des SIA parait aujourd'hui peu pertinente (voir aussi Mignot, 2025), nos résultats mettent en évidence que l'introduction et l'usage des SIA peuvent affecter d'autres métiers collaborant avec les médecins. Ils illustrent le fait que la mise en place de SIA en radiologie est le plus souvent pensée par rapport aux besoins des radiologues et à leur travail. Or, ce travail n'existe pas dans un espace organisationnel clos. Il implique des interactions avec d'autres acteurs dont le travail peut être affecté par le déploiement de SIA destinés aux radiologues. Cela souligne par ailleurs un manque de réflexion systémique sur l'IA dans les secteurs de la santé et plus largement au travail.

2.2.2. Le cas du système DF : « filet de sécurité » vs fiabilité limitée et risque d'excès de confiance

Comme l'illustre la figure 7, la dualité entre bénéfices et risques dans le cas du système DF est similaire sur ces certains points à celle relative à la RV (par exemple, gain vs perte de temps), mais s'en distingue sur d'autres aspects comme la confiance et l'acquisition de compétences, qui apparaissent comme deux problématiques centrales pour les radiologues.

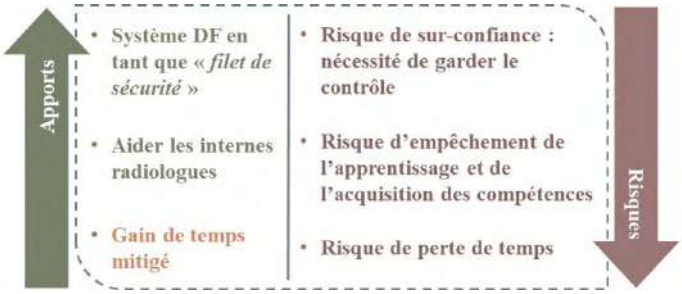


Figure 7. Bénéfices et risques liés au système DF

Entre « filet de sécurité » et besoin de construire l'interprétation des résultats du système DF

L'analyse met en évidence deux principaux types d'usage du système DF. Le premier, majoritaire et recommandé dans les structures étudiées, consiste à l'utiliser comme « *second lecteur* » : les radiologues ou urgentistes réalisent d'abord leur interprétation puis consultent les résultats du système, afin d'éviter des situations où « *on ne réfléchit plus* », ce qui peut par exemple être problématique pour l'apprentissage des internes en formation ou en cas de fatigue ou de surcharge. Toutefois, même dans ce cadre, le système peut orienter l'attention des médecins vers certaines zones de l'image : « *des fois, ça te facilite, au lieu de chercher partout, des fois elle te dit : il y a un truc à cet endroit-là – eh bien je vais aller voir à cet endroit-là précis* » (Médecin urgentiste 3).

Le second correspond à un usage en « première lecture », lorsque les résultats du système sont directement consultés, soit parce qu'ils apparaissent automatiquement à l'ouverture du dossier, soit dans des contextes de surcharge ou de fatigue. Ce type d'usage permet d'anticiper le contenu des comptes rendus types, validés ensuite après l'analyse d'images : « *Dans mon système d'information médicale, j'ai des petites imagettes et j'ai le résultat de l'analyse de l'intelligence artificielle, c'est marqué : "no fract". (…) Donc, j'ai déjà un aperçu, donc je sais que je vais pouvoir balancer mon compte rendu type "cheville normale"* » (Radiologue 11 – Extrait Observation Structure C2).

Le passage d'un usage en « second lecteur » à « premier lecteur » peut s'opérer progressivement à mesure que les médecins acquièrent une connaissance fine des performances et des limites du système, fondée sur son évaluation dans les situations réelles de travail.

Quel que soit le type d'usage, l'interprétation des résultats du système DF requiert l'expertise des radiologues, qui doivent les contextualiser à partir de la demande d'examen et de la situation clinique. La prise de connaissance de la demande d'examen et de la situation clinique est l'une des premières actions majeures dans cette activité d'interprétation. Or, contrairement aux médecins, le système analyse chaque image isolément et ne peut fournir un diagnostic global, ce qui limite sa portée.

Dans l'ensemble, le système DF est généralement décrit comme un « *filet de sécurité* » par les radiologues et les médecins urgentistes, surtout en cas de fatigue ou dans le travail de nuit. Il est considéré comme une protection contre les erreurs d'interprétation des radiographies. Le système DF semble également aider à orienter et à focaliser l'attention, en particulier dans la détection de fractures qui ne sont pas toujours facilement visibles. Et pour les urgentistes, il contribue à améliorer le flux des services d'urgence en offrant de prioriser les examens et d'accélérer la prise de décision (par exemple, libérer un patient en cas d'absence de lésions).

Malgré ces avantages perçus, plusieurs écueils liés à l'usage du système DF sont identifiés.

Une fiabilité limitée du système et un risque d'excès de confiance avec entrave à l'apprentissage dans le cas du système DF

Malgré une performance technique jugée globalement satisfaisante pour la plupart des médecins, le système DF présente plusieurs limites avec une fiabilité relative : production de faux positifs et fractures non détectées, informations parfois inutiles ou peu pertinentes (notamment lors des examens de contrôle de l'évolution des fractures, où le système détecte systématiquement une fracture qui est déjà connue par le radiologue), divergences de diagnostic générant un surcroit de vérification, voire des examens complémentaires comme le scanner, pour confirmer et « rendre la fracture visible ». De plus, ses explications limitées (zone détectée et niveau de certitude) ne suffisent pas toujours à confirmer une fracture invisible pour le radiologue.

Un autre enjeu important lié à l'usage du système DF concerne le risque de sur-confiance, lorsque les médecins se fient « aveuglément » à ses résultats sans les remettre en question. Cette dépendance peut conduire à des erreurs diagnostiques ou des ratés, d'autant que le système, dans sa version étudiée, ne détecte que les fractures, au risque de négliger d'autres pathologies visibles sur les images : « *Des fois, on fait tellement confiance (au système DF), on ne regarde plus les luxations* » (Radiologue 11). Conscients de ce risque, la plupart des radiologues insistent sur la nécessité de rester vigilants, de conserver la maitrise de l'interprétation et de ne pas tenir systématiquement pour « vrais » les résultats produits par le système.

Bien que perçu comme potentiellement utile pour les internes, fréquemment en première ligne dans l'interprétation des radiographies dans les hôpitaux et pour lesquels les radiologues experts seniors ne sont pas toujours présents ou disponibles, l'usage du système engendre des risques pour l'apprentissage de la radiologie. L'interprétation des radiographies, dont le volume ne cesse d'augmenter, demande beaucoup d'expertise que les internes risqueraient de ne pas suffisamment développer s'ils recourent en permanence au système. La question du maintien des compétences se pose également pour les radiologues seniors, certains manifestant un désintérêt croissant pour la radiographie, jugée

incertaine et « peu performante », au profit d'examens comme le scanner, considérés comme plus fiables et valorisants :

> « *Il (l'examen de radiographie) est très ingrat, c'est-à-dire qu'il demande beaucoup de connaissances, beaucoup d'expérience, beaucoup d'expertise, et pour autant on va quand même commettre des erreurs. (…) Et c'est vrai qu'en radio standard (…) on se pose toujours des questions, et du coup, on a souvent envie de dire : "Il faudrait quand même faire un scanner pour être sûr." (…) On se dit : "Je préfère être celui qui fait le scanner et qui trouve la pathologie que celui qui l'évoque peut-être et qui se trompe"* » (Radiologue 10).

Un dernier enjeu concerne le gain de temps attendu avec l'usage du système DF. Si son objectif initial était d'accélérer l'interprétation, ce bénéfice reste incertain : la détection de fractures invisibles ou de faux positifs entraine généralement un temps supplémentaire de vérification. Dès lors, certains médecins estiment que le système apporte davantage en termes de qualité et de sécurité qu'en termes d'efficacité temporelle. Il s'agit d'une problématique plus générale qui se retrouve avec d'autres systèmes d'IA, comme les SIA génératifs. La vérification des productions de ces systèmes nécessite parfois un temps considérable qui contrebalance le temps gagné sur d'autres tâches (voir par exemple, dans le domaine de la programmation, Simkute et al., 2024). Confrontés aux situations réelles de travail, l'usage des SIA et les bénéfices attendus comme le gain de temps ne vont pas de soi et peuvent même s'avérer difficiles à évaluer (Echajari et al., 2024).

Les résultats mettent en évidence que l'usage des SIA transforme le travail des radiologues (par exemple, adaptation au rythme de la dictée RV et surveillance du texte dicté pour la correction), parfois de manière contre-productive en augmentant la charge de travail. Cela témoigne du « *caractère inachevé, imparfait, voire contre-productif dans les situations de travail, de l'automatisation engendrée par des dispositifs d'IA* » (Gaglio et Mathieu-Fritz, 2024, p. 21).

Nos résultats mettent aussi en lumière que les SIA peuvent être pertinents dans certains contextes, par exemple en tant que « filet de sécurité » en conditions de travail dégradées, comme le montre également l'étude d'Olivier Ocquidant (2025) sur les services d'urgence. Cependant, l'usage des SIA soulève des enjeux d'opacité (Holmström et Hällgren, 2022 ; Zouinar, 2020), de sur-confiance (Parasuraman et Riley, 1997), de perte d'expertise (Lebovitz et al., 2023) et de faux sentiment de sécurité (Kohli et Jha, 2018) surtout si le système rencontre des problèmes de performance et de fiabilité.

3. CONCLUSION

Dans le contexte actuel, la question n'est plus de savoir si les SIA récents vont remplacer les radiologues, mais comment ils vont être progressivement intégrés dans les métiers de la radiologie. À rebours des discours monolithiques et techno-déterministes sur les conséquences de l'IA, cet article met en évidence des processus d'intégration et d'appropriation marqués par la variabilité des transformations liées à l'utilisation des SIA, selon les contextes d'usage, les choix organisationnels et les propriétés des systèmes. Il propose ainsi une lecture nuancée de leur introduction et de leurs conséquences en radiologie, soulignant la diversité des usages et la pluralité des effets observés, en cohérence avec les conclusions d'autres recherches récentes (Mathieu-Fritz et Trupia, 2024 ; Gaglio et Loute, 2023).

Nos résultats rappellent que l'évaluation des SIA doit aller au-delà de la seule mesure de performance technique dans des contextes contrôlés pour saisir leurs effets sur le travail et l'organisation des soins *in vivo*. Dans les cas étudiés, l'intégration des SIA n'apparait pas véritablement ou suffisamment pensée

en termes de gestion et de conduite des transformations des activités et de l'organisation du travail (Barcellini, 2022). On retrouve des écueils classiques de projets de conception (Barcellini, 2020), tels que la prédominance des considérations financières ou techniques – dans une logique « techno-solutionniste » – ainsi qu'une prise en compte insuffisante du travail dans toute sa complexité. C'est ce qui est observé dans les deux cas analysés. Or, ces dimensions sont cruciales, tant pour comprendre les transformations du travail liées aux SIA que pour introduire ceux-ci de manière « vertueuse » au sein des organisations de santé ou autres. Dans cette optique, la participation effective des travailleurs dans les processus de conception et de déploiement des SIA est essentielle (Darses et Reuzeau, 2004 ; Williams, 1983). C'est notamment dans ce sens qu'il nous semble nécessaire de développer une approche de l'IA centrée non seulement sur l'humain (Schmager et al., 2023 ; Xu, 2019), mais aussi sur le travail, et de prendre en compte le caractère contextuel des usages et des transformations induites (Hartswood et al., 2003 ; Holmström et Hällgren, 2022). Enfin, les risques liés aux usages de SIA comme la standardisation des pratiques, la sur-confiance, l'entrave potentielle à l'apprentissage des médecins, ou encore la perte d'expertise, doivent être pris en considération, voire dissipés, notamment pour continuer à garantir la fiabilité des analyses médicales, la qualité des soins et la sécurité des patients. Concernant la perte d'expertise, qui est un enjeu fondamental, l'aéronautique fournit un exemple intéressant de recherche de stratégies de préservation de compétences professionnelles. Dans ce secteur, un rapport du département américain des transports publié en 2016 constatait que les pilotes ne savaient plus bien voler en mode manuel depuis la forte automatisation du pilotage. Face à ce problème, une nouvelle règle de formation des pilotes a été mise en place par l'Administration fédérale américaine de l'aviation, exigeant davantage d'exercices de manœuvres de vol manuelles pour les pilotes. Une réflexion similaire sur ce risque parait essentielle en radiologie et au-delà pour le développement et la préservation de l'expertise professionnelle.

— RÉFÉRENCES

— Al-Aiad, A., Momani, A.K., Alnsour, Y. et Alsharo, M. (2020), « The Impact of Speech Recognition Systems on The Productivity and The Workflow in Radiology Departments: A Systematic Review », AMCIS 2020 TREOs, Salt Lake City, 12-16 aout.

— Anichini, G. et Geffroy, B. (2021), « L'intelligence artificielle à l'épreuve des savoirs tacites. Analyse des pratiques d'utilisation d'un outil d'aide à la détection en radiologie », in *Sciences sociales et santé*, vol. 39, n° 2, p. 43-69.

— Arena, L., Gaglio, G. et Vayre, J.-S. (2024), « Imbrication et "professionnalisation" des promesses sociotechniques. Le cas de l'IA en radiologie », in *Réseaux*, 248(6), 33-72.

— Balleyguier, C., Boyer, B., Athanasiou, A., Vanel, D. et Sigal, R. (2005), « Comprendre et utiliser le CAD (Aide informatisée au diagnostic) en mammographie », in *Journal de radiologie*, vol. 86, n° 1, p. 29-35.

— Barcellini, F. (2020), « Quelles conceptions de la coopération humains - robots collaboratifs ? », in *Activités*, vol. 17, n° 1.

— Barcellini, F. (2022), « The design of "future work" in industrial contexts », in Laroche, H., Bieder, C. et Villena-López, J. (dir.), *Managing Future Challenges for Safety: Demographic Change, Digitalisation and Complexity in the 2030s*, Berlin/Heidelberg, Springer, 124 p.

— Brunelle, F. et Brunelle, P. (2019), « Artificial intelligence and medical imaging: Definition, state of the art and perspectives », in *Bulletin de l'académie nationale de médecine*, vol. 203, n° 8-9, p. 683-687.

— Castellino, R.A. (2005), « Computer aided detection (CAD): an overview », in *Cancer Imaging*, vol. 5, n° 1, p. 17-19.

— Chartrand G., Cheng P. M., Vorontsov E., Drozdzal M., Turcotte S., Pal C. J., ... Tang A. (2017), Deep learning: A primer for radiologists, Radiographics, vol. 37, n° 7, p. 2113–2131.
— Chockley, K. et Emanuel, E. (2016), « The End of Radiology? Three Threats to the Future Practice of Radiology », in *Journal of the American College of Radiology*, vol. 13, n° 12, p. 1415-1420.
— Daniellou, F. et Rabardel, P. (2005), « Activity-oriented approaches to ergonomics: Some traditions and communities », in *Theoretical Issues in Ergonomics Science*, vol. 6, n° 5, p. 353-357.
— Darses, F. et Reuzeau, F. (2004), « Participation des utilisateurs à la conception des systèmes et dispositifs de travail », in Falzon, P. (dir.), *Ergonomie*, Paris, Presses universitaires de France « Hors collection », 680 p.
— Echajari, L., Jeanningros, H. et Lewkowicz, M. (2024), « Le codage de l'information médicale à l'épreuve de l'IA : performance et incertitude du codage en centre hospitalier », in *Réseaux*, 248(6), 153-191
— Fujita, H. (2020), « AI-based computer-aided diagnosis (AI-CAD): the latest review to read first », in *Radiological Physics and Technology*, vol. 13, n° 1, p. 6-19.
— Gaglio, G. et Loute, A. (2023), « L'émergence d'enjeux éthiques lors d'expérimentations de logiciels d'intelligence artificielle : le cas de la radiologie », in *Réseaux*, vol. 4, n° 240, p. 145-178.
— Gaglio, G. et Mathieu-Fritz, A. (2024), « IA, médecine et sciences sociales : une mise en perspective », in *Réseaux*, 248(6), 11-29.
— Gamkrelidze, T. (2022), « Des discours aux réalités de la conception, du déploiement et des usages des systèmes d'Intelligence Artificielle dans les situations de travail », thèse de doctorat en ergonomie, HESAM Université, https://theses.hal.science/tel-03959193v1/document
— Goodfellow, I., Bengio, Y. et Courville, A. (2016), *Deep Learning*, Cambridge, The MIT Press, 800 p.
— Hartswood, M., Procter, R., Rouncefield, M., Slack, R., Soutter, J. et Voss, A. (2003), « "Repairing" the Machine: A Case Study of the Evaluation of Computer-Aided Detection Tools in Breast Screening », in Kuutti, K., Karsten, E.H., Fitzpatrick, G., Dourish, P. et Schmidt, K. (dir.), ECSCW 2003, p. 375-394, Springer, 394 p.
— Henry, K.E., Kornfield, R., Sridharan, A., Linton, R.C., Groh, C., Wang, T., Wu, A., Mutlu, B. et Saria, S. (2022), « Human-machine teaming is key to AI adoption: clinicians' experiences with a deployed machine learning system », in *npj Digital Medicine*, vol. 5, n° 97.
— Holmström, J. et Hällgren, M. (2022), « AI management beyond the hype: exploring the co-constitution of AI and organizational context », in *AI & Society*, vol. 37, p. 1575-1585.
— Kempt, H., Heilinger, J.C. et Nagel, S.K. (2023), « "I'm afraid I can't let you do that, Doctor": meaningful disagreements with AI in medical contexts », in *AI & Society*, vol. 38, p. 1407-1414.
— King, B.F. (2018), « Artificial Intelligence and Radiology: What Will the Future Hold? », in *Journal of the American College of Radiology*, vol. 15, n° 3, p. 501-503.
— Kohli, A. et Jha, S. (2018), « Why CAD Failed in Mammography », in *Journal of the American College of Radiology*, vol. 15, n° 3PB, p. 535-537.
— Krishnaraj, A., Lee, J.K., Laws, S.A. et Crawford, T.J. (2010), « Voice recognition software: effect on radiology report turnaround time at an academic medical center », in *American Journal of Roentgenology*, vol. 195, n° 1, p. 194-197.
— Lebovitz, S., Lifshitz-Assaf, H. et Levina, N. (2023), « The no. 1 question to ask when evaluating AI tools », in *MIT Sloan Management Review*.
— Liew, C. (2018), « The future of radiology augmented with Artificial Intelligence: A strategy for success », in *European Journal of Radiology*, vol. 102, p. 152-156.
— Lincoln, C.M., Chatterjee, R. et Willis, M.H. (2019), « Augmented Radiology: Looking Over the Horizon », in *Radiology: Artificial Intelligence*, vol. 1, n° 1, p. 1-3.
— Machado, L., Vilgrain, V., Aubé, C., et Grégory, J. (2025), « La radiologie entre deux vagues : adoption de l'intelligence artificielle dans les hôpitaux universitaires français en 2025 », in *Journal d'imagerie diagnostique et interventionnelle*.

— Mathieu-Fritz, A., & Trupia, D. V. (2024), Le dermatologue, le mélanome et l'intelligence artificielle: Le (lent) développement des usages et les transformations de la réflexivité clinique. Réseaux, 248(6), 111-152.
— Mayo, R.C. et Leung, J. (2018), « Artificial intelligence and deep learning – Radiology's next frontier? », in *Clinical Imaging*, vol. 49, p. 87-88.
— Mignot, L. et Schultz, É. (2022), « Les innovations d'intelligence artificielle en radiologie à l'épreuve des régulations du système de santé », in *Réseaux*, n° 232-233, p. 65-97.
— Mignot, L. (2025), « Déléguer aux algorithmes ? Les enjeux de l'implémentation de l'intelligence artificielle en santé », in *Socio - La nouvelle revue des sciences sociales*, 20, 83-100.
— Moehrle, A. (2018), « "Radiology" Is Going Away… and That's Okay: Titles Change, A Profession Evolves », in *Journal of the American College of Radiology*, vol. 15, n° 3, p. 499-500.
— Morozov, E. (2013), *To Save Everything, Click Here. The Folly of Technological Solutionism*, New York, PublicAffairs, 432 p.
— Oakden-Rayner, L. (2019), « The rebirth of CAD: how is modern AI different from the CAD we know? », *Radiology: Artificial Intelligence*, vol. 1, n° 3, p. e180089.
— Ocquidant, O. (mai 2025), « Une place pour l'IA, mais l'IA à sa place », séminaire Santé et Big Data.
— Obermeyer, Z. et Emanuel, E.J. (2016), « Predicting the Future – Big Data, Machine Learning, and Clinical Medicine », in *The New England Journal of Medicine*, vol. 375, n° 13, p. 1216-1219.
— Oppenheimer, J., Luken, S., Hamm, B. et Niehues, S.M. (2023), « A Prospective Approach to Integration of AI Fracture Detection Software in Radiographs into Clinical Workflow », in *Life*, vol. 13, n° 1, p. 223.
— Paillé, P. (1994), « L'analyse par théorisation ancrée », in *Cahiers de recherche sociologique*, vol. 23, p. 147-181.
— Parasuraman, R. et Riley, V. (1997), « Humans and automation: Use, misuse, disuse, abuse », in *Human Factors*, vol. 39, n° 2, p. 230-253.
— Pezzullo J. A., Tung G. A., Rogg J. M., Davis L. M., Brody J. M., Mayo-Smith W. W. (2008), Voice recognition dictation: radiologist as transcriptionist, Journal of digital imaging, vol. 21, p. 384-389.
— Ramaswamy M. R., Chaljub G., Esch O., Fanning D. D., Vansonnenberg E. (2000), Continuous speech recognition in MR imaging reporting: advantages, disadvantages, and impact, American Journal of Roentgenology, vol. 174, n° 3, p. 617-622.
— Ranschaert, E. (2018), « Artificial intelligence in radiology: Hype or hope? », in *Journal of the Belgian Society of Radiology*, vol. 102, p. 1-2.
— Reverberi, C., Rigon, T., Solari, A., Hassan, C., Cherubini, P., Antonelli, G., Awadie, H., Bernhofer, S., Carballal, S., Dinis-Ribeiro, M., Fernández-Clotett, A., Esparrach, G.F., Gralnek, I., Higasa, Y., Hirabayashi, T., Hirai, T., Iwatate, M., Kawano, M., Mader, M., … Cherubini, A. (2022), « Experimental evidence of effective human-AI collaboration in medical decision-making », in *Scientific Reports*, vol. 12, n° 14952.
— Ringler, M.D., Goss, B.C. et Bartholmai, B.J. (2017), « Syntactic and semantic errors in radiology reports associated with speech recognition software », in *Health Informatics Journal*, vol. 23, n° 1, p. 3-13.
— Sahiner, B., Chan, H.P., Hadjiiski, L.M., Cascade, P.N., Kazerooni, E.A., Chughtai, A.R., … Attili, A. (2009), « Effect of CAD on radiologists' detection of lung nodules on thoracic CT scans: analysis of an observer performance study by nodule size », in *Academic Radiology*, vol. 16, n° 12, p. 1518-1530.
— Schmager, S., Pappas, I. et Vassilakopoulou, P. (2023), « Defining Human-centered AI: a comprehensive review of HCAI literature », The 15th Mediterranean Conference on Information Systems (MCIS) and the 6th Middle East et North Africa Conference on Digital Information Systems (MENACIS), Madrid (Espagne), 5 septembre.

— Shiraishi, J., Li, Q., Appelbaum, D. et Doi, K. (2011), « Computer-aided diagnosis and artificial intelligence in clinical imaging », in *Seminars in Nuclear Medicine*, vol. 41, n° 6, p. 449-462.
— Simkute, A., Tankelevitch, L., Kewenig, V., Scott, A.E., Sellen, A. et Rintel, S. (2024), « Ironies of Generative AI: Understanding and Mitigating Productivity Loss in Human-AI Interaction », in *International Journal of Human-Computer Interaction*, p. 1-22.
— Stoeklé, H.-C., Ivasilevitch, A., Mellot, F., et al. (2021), « Intelligence artificielle et apport du consentement éclairé dynamique en radiologie », in *Droit, santé et société*, vol. 2, p. 61-65.
— Tang, A., Tam, R., Cadrin-Chênevert, A., Guest, W., Chong, J., Barfett, J., Chepelev, L., Cairns, R., Mitchell, J.R., Cicero, M.D., Poudrette, M.G., Jaremko, J.L., Reinhold, C., Gallix, B., Gray, B., Geis, R., O'Connell, T., Babyn, P., Koff, D., … Shabana, W. (2018), « Canadian Association of Radiologists White Paper on Artificial Intelligence in Radiology », in *Canadian Association of Radiologists Journal*, vol. 69, n° 2, p. 120-135.
— Vallès-Peris, N. et Domènech, M. (2023), « Caring in the in-between: a proposal to introduce responsible AI and robotics to healthcare », in *AI & Society*, vol. 38, p. 1685-1695.
— White, K.S. (2005), « Speech recognition implementation in radiology », in *Pediatric Radiology*, vol. 35, p. 841-846.
— Williams, C. (1983), « A brief introduction to artificial intelligence », in IEEE (dir.), *Proceedings OCEANS'83*, p. 94-99.
— Xu, W. (2019), « Toward human-centered AI: a perspective from human-computer interaction », in *Interactions*, 26(4), 42-46.
— Zouinar, M. (2020), « Évolutions de l'Intelligence Artificielle et travail : quels enjeux pour l'activité humaine et la relation Humain-Machine ? », in *Activités*, vol. 17, n° 1.

DES IA EXPLICABLES POUR DES TECHNOLOGIES ACCEPTABLES ET SOUTENABLES POUR LE TRAVAIL : ENJEUX ET CONTOURS DE L'EXPLICATION DANS ET POUR L'ACTIVITÉ

Marc-Éric Bobillier Chaumon[1], **Ranya Bennani**[1 & 2], **Luca Mourgaud**[1, 3 & 4], **Myriam Fréjus**[2], **Denis Lemaître**[3], **Olivier Grisvard**[4]

1. Laboratoire CRTD – Cnam Paris – chaire de psychologie du travail
2. EDF Recherche et développement Paris-Saclay
3. École navale militaire – IRENAV – chaire NAIADE – Lanvéoc
4. Thales DMS – Brest

Cet article analyse les défis liés à l'intégration de l'intelligence artificielle (IA) dans les secteurs à haute fiabilité, tels que le nucléaire et la défense. Plutôt que de se focaliser sur la performance technique, les auteurs proposent une approche prospective et anthropocentrée, focalisée sur l'activité réelle des professionnels et les dynamiques socio-organisationnelles à l'œuvre. Le propos principal soutenu dans cet article est que la réussite de la transformation digitale repose sur la soutenabilité de l'IA et ses capacités d'explicabilité. Une IA soutenable est une IA qui fait sens pour les professionnels : préservant leur santé psychosociale et leur intégrité professionnelle, maintenant leur pouvoir d'agir, c'est-à-dire leur capacité à interpréter, décider et réguler leur activité face à la machine, et contribuant aussi à un développement favorable et durable de l'activité et du sujet. Cette réflexion prendra appui sur deux études de cas, issus du secteur nucléaire et de la défense, qui suivent la conception prospective de systèmes d'IA encore inexistants pour anticiper les besoins réels des futurs utilisateurs, notamment sur les questions d'explication et de transparence. Nos résultats permettent de définir cinq grands principes fondamentaux pour concevoir une IA explicable et soutenable qui assiste l'humain sans l'aliéner.

Mots-clés : intelligence artificielle, travail soutenable, transformation du travail, explication, confiance

INTRODUCTION : L'IA AU CŒUR DE LA TRANSFORMATION AU TRAVAIL

Le déploiement de l'intelligence artificielle (IA) dans les activités professionnelles ne se limite pas à une question d'intégration technique. Il interroge plus largement ses impacts psychosociaux, notamment sur la qualité de vie au travail, les conditions d'exercice et la santé des travailleurs, à travers la notion de soutenabilité technologique (Bobillier Chaumon, 2026). Les systèmes d'IA doivent faire sens pour les salariés, en étant appropriés et utilisables, mais aussi donner sens à l'activité en respectant les valeurs professionnelles et les règles de métier (Bobillier Chaumon, 2023), tout en garantissant un usage éthique et contrôlé (Beaudouin et Velkovska, 2023).

Cet article propose d'examiner cette soutenabilité à travers le prisme de l'explication, visant à rendre les systèmes d'IA plus transparents et compréhensibles. Leur diffusion rapide se heurte en effet à une limite majeure : leur opacité, qualifiée de « boite noire », qui renvoie à la difficulté de comprendre les processus algorithmiques sous-jacents (Rudin, 2019). Or, le niveau d'explication conditionne la confiance accordée à l'IA (Kizilcec, 2016), confiance qui peut varier entre méfiance, défiance ou surconfiance selon le degré de transparence perçu.

Deux approches se dessinent. La première, technocentrée, conçoit l'explication comme une propriété statique d'un modèle d'IA : si celui-ci peut justifier ses décisions par des raisonnements statistiques ou probabilistes, il est jugé explicable, indépendamment du contexte ou des profils utilisateurs. Or, le réel de l'activité est toujours plus complexe et incertain que ne le laissent penser ces formalisations. C'est pourquoi une seconde approche, anthropocentrée, mérite d'être explorée. L'idée d'une explication située s'inscrit dans un ensemble de travaux issus des sciences humaines et sociales qui considèrent l'explication non comme un contenu abstrait et universel, mais comme une activité contextualisée. Les sciences cognitives proposent également une définition de l'explication selon laquelle elle se focaliserait davantage sur les processus cognitifs amenant à rendre l'information compréhensible (Karsenty, 1996). Une autre perspective, celle de la clinique des usages (Bobillier Chaumon, 2023), conçoit l'explication comme un processus dynamique et co-construit avec l'utilisateur, ajusté à ses besoins, à son expérience et aux contingences de l'activité (Fréjus, 1999).

L'objectif de cet article est de mieux comprendre les contours et enjeux d'une IA explicable au travail, et d'en évaluer les implications sur l'activité et leur acceptation dans des milieux professionnels sensibles et à risques. Deux études de cas issues de thèses CIFRE en cours illustreront cette réflexion :

1. La conception prospective d'une IA explicative dans l'industrie nucléaire.
2. L'acceptation située d'un système d'IA de détection de menaces militaires.

1. L'IA DANS LES PRATIQUES DE TRAVAIL À RISQUES : ENTRE PROMESSES ET NOUVEAUX DÉFIS

1.1. Contexte et spécificités des environnements critiques

Les environnements professionnels dits « à risques » – aviation, nucléaire, santé, défense ou maritime – se caractérisent par une forte complexité organisationnelle, une incertitude permanente et une exigence critique de fiabilité (Hollnagel, 2017). Ces milieux relèvent des High Reliability Organizations (HRO), où la probabilité d'erreur doit être minimisée grâce à des pratiques robustes et une vigilance constante (Weick et Sutcliffe, 2007). Dans ces contextes, les technologies jouent historiquement un rôle de soutien central à l'activité humaine. L'intégration de l'intelligence artificielle représente toutefois un tournant majeur : elle est présentée par les chefs d'entreprise et concepteurs comme une technologie de rupture, susceptible d'optimiser le traitement de l'information, d'automatiser certaines tâches critiques et de renforcer la résilience organisationnelle (Cummings, 2021).

1.2. Les principales promesses de l'IA

Première promesse : l'IA, un levier d'efficacité opérationnelle. Sa capacité à analyser d'immenses volumes de données hétérogènes accélère les processus décisionnels et favorise l'anticipation d'événements critiques (Shneiderman, 2020). Deuxième promesse : le renforcement de la sécurité. En réduisant les risques d'erreur humaine et en maintenant une vigilance continue, l'IA agit comme un filet protecteur dans des secteurs sensibles comme l'aviation et le nucléaire (Amalberti, 2013). Troisième promesse : l'IA est un outil d'optimisation organisationnelle et économique, en permettant une meilleure allocation des ressources et une coordination accrue entre humains et systèmes techniques (Brynjolfsson et McAfee, 2017).

Ces promesses ne sont toutefois pas exemptes de tensions. En se concentrant sur une logique de performance technologique, elles risquent de sous-estimer la complexité du réel du travail et les besoins effectifs des salariés. L'IA, censée renforcer la fiabilité, peut alors se révéler contre-productive, voire génératrice de nouveaux risques. C'est précisément sur ces obstacles que va porter notre analyse.

— 2. LES OBSTACLES À L'ADOPTION DE L'IA DANS L'ACTIVITÉ

Les obstacles à l'adoption de l'IA dans l'activité professionnelle sont d'au moins quatre natures.

2.1. Niveaux de confiance et dépendance

L'automatisation ne supprime pas tous les risques : elle peut en produire de nouveaux. Le premier écueil tient à la surconfiance (*automation bias*), où les opérateurs accordent une confiance excessive aux recommandations algorithmiques, parfois au détriment de leur propre jugement (Parasuraman et Riley, 1997). À cela s'ajoute le paradoxe de l'automatisation partielle : en déchargeant les opérateurs de certaines actions directes et simples, l'IA déplace la charge cognitive vers des tâches de supervision plus complexes. L'IA entraine une rareté des interventions humaines, qui se traduit par une perte progressive de compétences et une difficulté à reprendre la main en cas de défaillance ; ce que Bainbridge (1983) nommait l'« ironie de l'automatisation ». Ce phénomène affecte aussi le sentiment d'efficacité personnelle (Bandura, 1977, repris dans Compeau et al., 1999, dans le domaine de l'informatique) : en déléguant trop à la machine, les salariés doutent de leurs propres capacités à assurer la tâche attendue. Cette dépendance croissante favorise une passivité cognitive, réduisant l'engagement actif et la vigilance des professionnels. La défiance est alors moins contre le système qu'envers ses propres capacités et ressources.

2.2. Opacité, éthique et responsabilité

La deuxième nature d'obstacle tient à l'opacité des systèmes. Les modèles d'IA, notamment ceux fondés sur l'apprentissage profond, fonctionnent souvent comme des « boites noires » dont les logiques décisionnelles échappent aux utilisateurs (Burrell, 2016). L'absence d'explication constitue un frein majeur à leur appropriation et à leur confiance (Doshi-Velez et Kim, 2017). Par ailleurs, les algorithmes sont loin d'être neutres : en reproduisant les biais des données d'entrainement, les systèmes risquent d'induire des discriminations ou des erreurs systémiques (O'Neil, 2016). Dans des environnements sensibles, de tels biais peuvent avoir des conséquences graves, allant de la mauvaise identification d'un signal à des décisions opérationnelles inadaptées. Enfin se pose la question cruciale de la responsabilité : qui doit être tenu comptable en cas d'erreur grave ? L'opérateur, l'organisation ou le concepteur (Cihon et al., 2021) ? Ces enjeux deviennent critiques dans les domaines militaires ou médicaux, où les décisions engagent directement la sécurité des individus.

2.3. L'IA comme outil de prescription

L'IA ne se contente pas de soutenir l'activité : elle prescrit des façons de faire et des manières d'être au travail. Elle rationalise l'action en modélisant le travail comme un processus mesurable et optimisable, parfois au détriment de l'expérience professionnelle. Dans certains entrepôts logistiques, les IA d'assistance aux préparateurs de commandes fixent en temps réel les cadences et tolérances, transformant l'opérateur en simple exécutant de prescriptions automatisées[1]. Cette logique panoptique accroit la contrôlabilité des individus, tout en réduisant leur marge de manœuvre. Mais en cherchant à contrôler une activité de plus en plus hybride, dématérialisée et fragmentée, les organisations risquent de réduire l'IA à un instrument de normalisation plutôt que d'en faire un levier d'innovation – ce qui pourrait expliquer l'échec massif de près de 95 % des projets de déploiement, selon une étude du Media Lab de l'université du MIT (MIT Media Lab, 2025).

Au-delà de l'action, l'IA normalise aussi la subjectivité, c'est-à-dire les façons de s'investir, de ressentir et de se projeter dans le travail. Les interactions avec des chatbots ou des systèmes génératifs exigent un langage simplifié et des requêtes standardisées (Zamfirescu-Pereira et al., 2023). Loin d'élever la machine au niveau de l'humain, ces pratiques tendent à abaisser l'humain au niveau des contraintes fonctionnelles de la machine. En outre, cette logique d'alignement s'étend à l'engagement subjectif attendu du salarié. Parce que l'IA est censée « augmenter » l'humain, l'organisation en attend davantage : autonomie, réactivité, adaptabilité, résilience. Mais en déléguant des tâches (de bas niveau) à l'IA, il se retrouve confronté aux activités les plus complexes, intensifiant sa charge cognitive et psychique.

2.4. Entre mètis et IA : les formes invisibles de l'intelligence humaine au travail

Une erreur fréquente consiste à surestimer la capacité des systèmes à se substituer à l'intelligence humaine. Face aux aléas de l'activité et à la complexité des contextes professionnels, il est illusoire de penser pouvoir se passer de l'intelligence humaine, qui rend précisément possibles l'expression et l'actualisation de ces formes artificielles d'intelligence. La *mètis* humaine est indispensable pour que ces technologies fonctionnent. L'exemple de Tesla l'illustre : l'automatisation excessive de la production de la Model 3 a provoqué d'importants blocages, conduisant Elon Musk à reconnaitre, sur sa plateforme X, que « *les humains sont sous-estimés* », et à réintroduire de la flexibilité humaine dans les processus (Wilson et Daugherty, 2018). Cet épisode révèle que la flexibilité et l'ingéniosité humaines restent indispensables pour faire face à la variabilité et à la complexité du travail, qui échappe toujours à la modélisation.

En définitive, ces divers constats (non exhaustifs) rappellent que l'intégration de l'IA dans les environnements critiques ne peut se limiter à une logique de performance technologique. Elle doit s'inscrire dans une approche systémique, attentive aux dimensions humaines, psychosociales et organisationnelles du travail (Bobillier Chaumon, 2016). L'IA s'impose comme un acteur de la régulation du travail : elle redéfinit à la fois les modalités de l'action et les attentes envers les individus, façonnant un nouvel idéal du travailleur, redéfinissant les façons de travailler (les pratiques, les compétences et les responsabilités des acteurs impliqués). L'enjeu est dès lors de concevoir des IA adaptées aux contextes réels et construites avec les collectifs de travail, plutôt que d'exiger des individus qu'ils se plient aux logiques des machines. C'est notamment dans cette perspective que les enjeux épistémologiques et pratiques sur l'explication de l'IA se positionnent.

1 Soper, S. (2021), « Fired by bot at Amazon: "It's you against the machine" », *The Minnesota Star Tribune*. https://www.startribune.com/fired-based-on-algorithms/600072977/

— 3. QU'EST-CE QU'UNE IA EXPLICABLE POUR LE TRAVAIL ?

L'explicabilité est souvent envisagée comme une propriété statique des systèmes d'IA. Or, dans une activité vivante et incertaine, elle doit être comprise comme un processus construit dans un contexte spatio-temporel et organisationnel donné. L'explication devient alors un construit dynamique, et non un donné figé.

3.1. De l'explicabilité à l'explication des IA dans l'activité

a) De l'explicabilité…

L'explicabilité (*explainable artificial intelligence*, ou XAI) est définie comme la capacité à rendre intelligibles à la fois les processus techniques d'un système d'IA et les décisions humaines associées (Gornet et Maxwell, 2023). Cette approche, centrée sur la transparence algorithmique, est pertinente dans le champ de la science des données mais demeure insuffisante pour saisir la complexité de l'interaction humain-IA. En effet, l'utilisateur n'est pas qu'un décideur rationnel : il est situé dans une histoire, une culture professionnelle, un collectif de travail et des contraintes organisationnelles dynamiques. Pour être réellement opérante, la conceptualisation de l'explicabilité doit donc intégrer ces dimensions contextuelles.

b) … à l'explication

Dépasser l'explicabilité technique conduit à une approche anthropocentrée : l'explication. Issue du latin *explicare* (« déployer, dérouler, développer »), elle vise à rendre compréhensible une situation par un interlocuteur, selon ses besoins. Adadi et Barrada (2018) identifient quatre fonctions de l'explication d'une IA : justifier, contrôler, améliorer et découvrir. Miller (2019) insiste sur son caractère social et interactif, rejoignant Fréjus (1999) : une explication n'existe qu'à partir du moment où elle est reçue comme telle par l'interlocuteur.

Vuarin et Steyer (2023) distinguent deux fonctions complémentaires : (1) expliquer pour comprendre (intelligibilité technique), cherchant à décortiquer le fonctionnement interne des modèles d'IA, et (2) expliquer pour rendre des comptes (*accountability*), qui vise à justifier les décisions du système d'IA. Cette double visée souligne un enjeu central : qui est responsable de l'explication et qui répond aux éventuels dommages ? Toutefois, ces approches négligent encore un point essentiel : l'organisation et l'activité dans lesquelles s'insère l'IA. Une explication pertinente ne peut être réduite à une simple transparence technique ni à des standards normatifs. Apparait alors un défi majeur : définir tant la nature des explications que leur forme, pour qu'elles puissent être utiles et porteuses de sens dans le contexte socio-organisationnel et opérationnel visé.

Du point de vue de la forme, les travaux en ergonomie et en IHM (interactions humain-machine) montrent que les explications doivent s'adapter au contexte, aux publics et à leurs besoins informationnels. Leur pertinence repose notamment sur des modalités interactionnelles (types de signaux visuels ou sonores ; Gornet et Maxwell, 2023) et sur la qualité du dialogue avec la machine (Kirsch, 2017) qui facilitent l'attention et la co-construction de l'explication. Dans ce cadre, la *flexibilité* est décisive en ce qu'elle propose une certaine « plasticité » qui permet à l'utilisateur d'explorer, de corriger ou de modifier les explications produites, afin de renforcer sa capacité d'agir. Ce qui nécessite d'intégrer au plus tôt les individus et la connaissance de leurs activités dans la conception et d'adapter les formes à leurs besoins et leur profil.

Du point de vue de la nature ou du type des explications, ces dernières doivent être sélectives et pertinentes plutôt qu'exhaustives, en cohérence avec les modes de raisonnement humain souvent

contrastifs[2], orientées vers l'action et adaptées aux contextes cognitifs, sociaux, historiques et organisationnels des acteurs. Pour une IA qui se veut utile et pertinente, il est donc crucial de comprendre ces logiques, afin qu'elle puisse générer des explications à la fois concises et adaptées à la façon dont nous traitons l'information en contexte. Les approches contrefactuelles – qui indiquent ce qu'il faudrait changer pour obtenir un autre résultat – apparaissent aussi plus opérantes. L'exemple d'Alice (tiré de Verma et al., 2024), dont la demande de prêt est refusée par un algorithme, illustre bien la différence : l'explication causale (« *votre solvabilité est trop basse* ») est insuffisante, alors qu'une explication contrefactuelle (« *que faire pour que ma demande soit acceptée la prochaine fois ?* » ou « *que va-t-il se passer si… ?* ») devient un véritable levier d'action : « *Si votre revenu augmente de X, votre demande sera acceptée.* »

3.2. Explication et soutenabilité de l'IA

Les explications d'un système d'IA ne peuvent être dissociées des activités professionnelles ni des processus cognitifs mobilisés par les travailleurs. L'explication est une condition préalable à la confiance, car elle permet de construire un sens partagé et de rendre le monde plus prévisible (Marzano, 2010).

Par ailleurs, des explications adaptées au contexte contribuent au développement de l'activité et des compétences : elles soutiennent l'apprentissage, allègent la charge de travail et améliorent l'efficacité. Surtout, leur co-construction avec les salariés est essentielle. Intégrer les besoins des professionnels et la spécificité de leurs métiers offre la possibilité de définir les conditions d'une IA acceptable, c'est-à-dire un outil qui soutient la qualité du travail, respecte les gestes de métier et favorise le développement du pouvoir d'agir. À ces conditions, l'explication devient un facteur-clé de soutenabilité : elle permet à l'IA de s'inscrire durablement dans l'activité humaine, en servant l'opérateur plutôt qu'en le contraignant.

Nous présenterons deux études de cas issues de recherches en cours pour illustrer ces enjeux :

1. La conception prospective d'une IA explicative dans une industrie à risques (sureté nucléaire) ;
2. L'analyse des conditions d'acceptation située d'un système d'IA de détection de menaces militaires.

— 4. RETOURS D'EXPÉRIENCE ET ÉTUDES DE CAS SUR LE DÉPLOIEMENT DE SOLUTIONS IA DANS DES SITUATIONS À RISQUES

Afin d'éclairer les enjeux d'explication de l'intelligence artificielle dans des environnements critiques, nous présentons deux recherches-actions issues de thèses CIFRE lancées en 2024. Ces travaux visent à documenter les besoins en matière d'explication, de confiance et de contrôle, et à orienter la conception de systèmes d'IA soutenables et acceptables.

4.1. Outil d'aide au diagnostic en centrale nucléaire

En partenariat avec un centre de production nucléaire, une thèse CIFRE étudie la conception d'un système d'aide au diagnostic par machine learning[3] pour anticiper un événement rare mais critique : l'arrivée de « colmatants » obstruant les circuits de refroidissement et pouvant provoquer l'arrêt d'un réacteur. Ce qu'on appelle le « risque SAR[4] » est lié à une concentration anormale d'ions sodium susceptible de perturber les circuits de refroidissement. La rareté et la complexité de cet événement justifient un dispositif d'assistance technologique au diagnostic.

2 Miller (2019) rappelle en effet que l'individu recherche davantage l'explication d'un événement non connu, incertain ou incohérent.
3 L'apprentissage automatique (machine learning) vise à donner aux machines la capacité d'« apprendre » à partir de données, via des modèles mathématiques.
4 Lié à une concentration anormale d'ions sodium affectant la perméabilité des sols et les installations de filtration.

Une première étape d'analyse terrain des activités montre la diversité des profils d'utilisateurs et la complexité organisationnelle de la gestion du risque, qui mobilise plusieurs métiers aux logiques complémentaires (ingénieurs « source froide », responsables de la surveillance de la source froide, et ingénieurs « tranche en marche », responsables de la coordination des activités sur le site, ainsi que d'autres métiers comme les métiers de maintenance du site). L'activité de supervision de la survenue du risque nécessite également une articulation fine entre expertise locale et coordination collective.

> « Ça, c'est la réunion du mardi exactement, où on fait le point avec l'analyse de l'ingénieur source froide, et on se projette sur ce qui va se passer dans les jours à venir, la semaine prochaine, les actus, on se projette qu'il peut y avoir des modifs avec les impacts vers les métiers, pour faire des nettoyages en fonction de ce qu'on se projette, de ce qui peut arriver, et de retravailler tout le planning, avec les ITEM notamment. (...) Ça, c'est quand tout va bien et que c'est calme, et quand on monte des tâches là où ça commence à se compliquer un peu plus, les réunions deviennent plus fréquentes. » (Ingénieur tranche en marche, ITEM)

Grâce à cette première phase d'analyse, nous pouvons définir en quoi l'outil doit s'intégrer dans un écosystème organisationnel et technologique préexistant, situé et mouvant. Elle constitue un socle nécessaire mais insuffisant : une approche projective est désormais requise pour anticiper les transformations du travail liées au futur outil. Cette projection est la condition pour répondre aux exigences spécifiques de granularité et de flexibilité des explications relevées dans nos premiers résultats. Pour cela, un dispositif méthodologique mixant diverses approches a été déployé :

— Des « entretiens resituants » pour accéder aux expériences de travail (Cahour, 2016) ;
— Des méthodes de simulations participatives (Van Belleghem et Barcellini, 2011) afin de projeter les utilisateurs dans des scénarios futurs et tester des choix technico-organisationnels ;
— Le tri de cartes (Fastrez et al., 2009) pour structurer les catégories mentales et préciser l'architecture de l'information et la granularité des explications.

Les résultats préliminaires documentent trois dimensions-clés.

— L'explication : elle sera sollicitée en cas d'incohérences perçues ou de changement de niveau de risque. « *Le fait de connaitre les seuils. Parce que là, je me dis qu'il y a des couleurs en fonction des seuils. Qu'est-ce qui fait les changements de seuil, en fait ? Qu'est-ce qui dit le faible ? Pourquoi il est faible ? C'est parce qu'on a regardé cette donnée-là* » (ITEM). Le degré de détail attendu varie selon l'expertise : les novices privilégient la simplicité (« *Mais par contre, je mets justement ce petit encart. Donc oui, il y a un intérêt d'avoir un truc détail, synthétique, j'aurais dit. Je ne sais pas, synthèse concise. C'est pour les non-initiés* » (ITEM)), les experts demandent des justifications approfondies (« *Moi, je fonctionne comme ça, ça me fait prendre du temps. J'aime bien tout regarder. (…) Parce que moi, je vais assez dans ces détails-là, des fois trop* » (ITEM)).

— La confiance : elle apparait fragile, en raison du manque de données d'entrainement et par des expériences négatives avec d'autres outils. Elle varie selon le niveau de risque : plus celui-ci est élevé, plus les opérateurs se montrent prudents et exigent des garanties. « *Pourquoi elle projette comme ça ? Si elle ne sait pas pourquoi, on aura du mal à donner de la confiance* » (ING SF). L'absence d'explication claire peut renforcer la méfiance – « *Si l'IA nous dit qu'elle ne sait pas pourquoi, on aura du mal à lui faire confiance* » (ING SF) –, tandis qu'une justification partielle assumant ses limites incite deux avis différents : (1) un refus d'utilisation de l'outil, une considération de l'outil comme non fonctionnel (« *Je l'éteins, je le débranche. Il ne me sert à rien. (…) Non, non, mais s'il ne peut*

pas se prononcer, on fait sans lui, déjà » (ING SF)), (2) une acceptation de l'outil à condition d'avoir une explication et une volonté de compléter par l'opérateur les informations manquantes (« *S'il ne sait pas se prononcer parce qu'il manque une donnée, ce serait bien de savoir pourquoi il ne peut pas dire. S'il ne peut pas parce qu'il manque une donnée, et que toi, tu as la donnée et que tu l'implémentes dedans* » (ITEM)).

— La flexibilité : la possibilité d'ajuster les paramètres (seuils de vérification, nature des données et types de données prises en compte) est jugée essentielle pour maintenir le contrôle face à l'évolution du contexte. « *Par contre, ce qui serait bien, typiquement, les enquêtes (outil de surveillance) que je t'ai montrées, c'est moi qui les ai faites à la main, j'y mets ce que je veux et je sais ce qu'il y a dedans. (...) Vu que cet outil permet de récupérer les données, il faudrait qu'on ait la main, (...) que je puisse modifier comme dans Nova, mes échelles* » (ITEM).

Dans ce premier cas, il apparait donc que l'acceptabilité de l'outil dépendrait de sa capacité à fournir des explications différenciées et pas forcément systématiques, à instaurer une confiance dynamique et à garantir un contrôle ajustable par les opérateurs.

4.2. Prise de décision en contexte militaire

Les environnements militaires se caractérisent par une forte complexité technique, une incertitude permanente et des contraintes temporelles extrêmes. Dans les environnements militaires (frégates FREMM, avions de patrouille ATL2), la décision repose sur l'intégration de signaux multiples (radars, guerre électronique, renseignements externes) et s'inscrit dans des chaines hiérarchiques où chaque choix peut avoir des implications stratégiques majeures (Endsley, 1995 ; Hollnagel, 2017). L'introduction de systèmes d'IA d'aide à la décision vise alors à soutenir la performance des opérateurs en filtrant l'information, en détectant les anomalies et en classifiant les signaux. Mais l'enjeu central reste l'articulation entre l'apport de l'IA et l'expertise humaine dans des situations critiques, où l'erreur est lourde de conséquences. Comprendre cette articulation est essentiel pour concevoir des systèmes à la fois performants, sûrs et réellement exploitables/acceptables par les opérateurs dans des environnements opérationnels complexes et contraints.

L'étude se focalise sur des opérateurs radars, responsables de la détection et de la classification des signaux. Trois défis d'adoption apparaissent déjà dans les premières analyses effectuées :

1. **Équilibre assistance/autonomie :** l'IA doit soutenir sans réduire ni le jugement humain ni la capacité à interpréter des signaux ambigus.
2. **Gestion collective :** l'efficacité d'une recommandation dépend de sa place dans la chaine hiérarchique militaire et des coordinations inter-niveaux (au sein du collectif de travail très important et structuré).
3. **Transformation du travail :** l'automatisation déplace l'effort de classification vers la supervision, modifiant les compétences requises et posant la question du maintien des savoir-faire critiques.

Une série d'entretiens auprès d'opérationnels (actuellement en poste)(n=14) ainsi que des sessions d'observations sur simulateur de vol ont été effectués. Les résultats de ces travaux montrent que les besoins d'explication apparaissent fortement situés, dépendant du poste occupé, du niveau d'expérience des opérateurs et des contraintes temporelles de la situation opérationnelle. Les opérateurs privilégient des explications synthétiques, ciblées et immédiatement exploitables, compatibles avec des situations de forte pression temporelle : « *Il faut que l'information soit claire et directe : est-ce que je dois agir maintenant ou pas ? Les détails viennent après* » (opérateur radar). Des explications trop détaillées ou trop techniques sont perçues comme perturbantes. Une inadéquation entre le format des explications et l'activité réelle peut générer une surcharge cognitive, de la confusion et une perte

de confiance envers le système : « *Si l'outil fournit des infos qui ne correspondent pas à ce qu'on voit, et même parfois à ce qu'on peut ressentir, on commence à douter de la fiabilité de tout le système* » (opérateur radar). Pour finir, l'explication est conçue comme un support à la coopération, devant être partageable au sein de l'équipage afin de soutenir la construction collective de la décision et la coordination inter-niveaux : « *L'information n'est utile que si elle circule. Si chacun garde son raisonnement pour lui, on ne peut pas coordonner les actions efficacement. Il est important que l'équipage comprenne la situation et agisse en conséquence* » (coordinateur tactique).

Ainsi, l'IA ne se limite pas à fournir une réponse correcte : elle doit contribuer à renforcer la confiance et la coopération entre les opérateurs.

D'autres analyses de terrain viendront compléter ces analyses – via des méthodes de restitution indirecte, compte tenu de la difficulté d'accéder à ces milieux de travail confinés et confidentiels (secret défense).

4.3. Principes généraux d'explication en environnements critiques

La mise en parallèle des deux études révèle des convergences et des spécificités contextuelles.

— Dans les environnements militaires, soumis à des contraintes temporelles extrêmes et des signaux multiples, les opérateurs privilégient des explications synthétiques, immédiatement exploitables et partageables, pour une décision collective traçable.
— Dans le cadre nucléaire, les utilisateurs sollicitent surtout des explications « à la demande », déclenchées par une incohérence ou un changement de niveau de risque. Elles doivent être graduées (simples pour les novices, détaillées pour les experts) et accompagnées de marges de flexibilité (ajustement des seuils, complétion des données) pour renforcer la confiance.

De ces premiers résultats, cinq principes généraux d'explication émergent :

1. **Adaptativité multi-niveaux :** ajuster le degré de détail selon le profil et la situation, avec possibilité de passer d'un résumé opérationnel à une justification approfondie.
2. **Pertinence contextuelle et temporelle :** synchroniser l'explication avec les fenêtres décisionnelles et déclencher l'information uniquement en cas d'anomalie significative.
3. **Explication coopérative :** rendre les décisions compréhensibles et partageables pour renforcer la confiance collective et la coordination.
4. **Transparence sur l'incertitude :** signaler clairement les limites et manques de données, afin de permettre une vérification complémentaire et de réduire l'incertitude.
5. **Flexibilité et contrôle utilisateur :** offrir des options d'ajustement et préserver la maitrise humaine.

Ces principes, encore en cours de consolidation, nous semblent constituer des repères essentiels pour concevoir des outils d'IA soutenables et réellement appropriables en contexte critique.

— CONCLUSION

Les travaux sur l'explication de l'IA tendent encore trop souvent à en faire un état figé, réduit à la transparence des modèles. Or, comme nous l'avons soutenu dans cet article, l'explication doit plutôt être envisagée comme un processus dynamique, évolutif et contextualisé, capable de s'adapter aux besoins différenciés des utilisateurs. Elle ne se limite pas à rendre un modèle intelligible : elle ouvre la voie à un dialogue continu entre système et utilisateurs, guidé par les cinq principes énoncés plus haut.

Ainsi, l'explication de l'IA doit être perçue comme une composante vivante du couplage humain-machine. Dans les environnements critiques, l'enjeu dépasse la simple « boite noire » algorithmique : il s'agit de concevoir une interaction explicative située, sensible aux rôles, aux temporalités et aux dynamiques socio-organisationnelles et socioprofessionnelles du métier, afin que le système contribue réellement à l'efficacité, à la performance et surtout à la qualité/fiabilité de l'activité. L'objectif n'est pas de rationaliser l'usage de l'IA, mais bien de permettre aux professionnels de conserver – voire de renforcer – un rôle central, actif et significatif dans leur métier. Pour cela, il convient de cultiver un sentiment d'utilité, de sens et d'engagement, en garantissant que l'IA ne se substitue pas à l'humain, mais l'accompagne.

En ce sens, la confiance et l'acceptation de l'IA ne dépendent pas uniquement de la qualité des explications fournies, mais aussi de la capacité des professionnels à se sentir, grâce à l'appui de l'IA :

— **Auteurs de leur travail :** le système doit donner à l'utilisateur le sentiment d'être pour quelque chose dans ce qui se passe, d'avoir l'impression d'être à l'origine des choses qui le concernent dans l'exercice de son travail, et surtout de ne pas se sentir comme un simple spectateur ou exécutant de conduites qui se réaliseraient sans lui, voire contre lui.

— **Acteurs de leur travail :** l'utilisateur doit se sentir à la fois légitime dans ses actions et légitimé par l'outil, capable de faire autorité dans son travail ; se sentir comptable des actions entreprises et disposer d'un réel « droit d'agir » dans l'activité, avec un degré de contrôle sur les activités déployées.

— **Agents de transformation :** l'IA doit renforcer la capacité de l'utilisateur à influer sur son travail et ses conditions de réalisation, plutôt que de le contraindre à les subir. Son pouvoir d'agir doit dès lors être maintenu, voire élargi, lui offrant la possibilité de réinventer ses pratiques et d'innover dans son métier, facteur-clé de santé au travail.

Penser la soutenabilité de l'IA, c'est donc reconnaitre que les professionnels ne sont pas de simples utilisateurs passifs, mais des co-constructeurs du sens et des usages. Cela implique de prendre en compte leur expérience de l'activité, leur expertise du terrain, comme des leviers incontournables d'une conception éthique, durable et réellement appropriable de l'IA au travail.

BIBLIOGRAPHIE

— Adadi, A., Berrada, M. (2018), Peeking inside the black-box: a survey on explainable artificial intelligence, *IEEE Access 6, 52138–52160.*
— Amalberti, R. (2013), *Piloter la sécurité*, Paris, Springer, 141 p.
— Bainbridge, L. (1983), Ironies of Automation, *Automatica*, n° 19, 775-779.
— Bandura, A. (1977), Social Learning Theory. New York: General Learning Press
— Beaudouin, V. et Velkovska, J. (2023), « Enquêter sur l'"éthique de l'IA" », in *Réseaux*, n° 240(4), 927.
— Bobillier Chaumon, M.É. (2016), Acceptation située des TIC dans et par l'activité : Premiers étayages pour une clinique de l'usage, *Psychologie du Travail et des Organisations*, 22(1), 4-21.
— Bobillier Chaumon, M.-É. (2023), *Psychologie du travail digitalisé, Nouvelles formes du travail et c*linique des usages, Malakoff, Dunod. 208 p.
— Bobillier Chaumon, M.-É. (à paraitre en 2026), « De la soutenabilité des technologies à la vulnérabilité du/au travail : une approche par la clinique des usages », in Dalmaso, C., Debout-Cosme, F. et al. (dir.), *Le Travail vulnérable*, Toulouse, Octares.
— Brynjolfsson, E. et McAfee, A. (2017), « The Business of Artificial Intelligence », in *Harvard Business Review*, 7, 3-11.

— Burrell, J. (2016), « How the machine "thinks": Understanding opacity in machine learning algorithms », in *Big Data & Society*, 3(1).
— Cahour, B. (2016), « Approches ergonomiques de la mobilité durable », actes du 51e Congrès de la SELF, Marseille, 21-23/09/16.
— Cihon, P., Schuett, J. et Baum, S.D. (2021), « Corporate Governance of Artificial Intelligence in the Public Interest », in *Information*, 12(7), 275.
— Compeau, D., Higgins, C.A. et Huff, S. (1999), « Social Cognitive Theory and Individual Reactions to Computing Technology: A Longitudinal Study », in *MIS Quarterly*, 23(2), 145158.
— Cummings, M.L. (2021), « Artificial Intelligence and the Future of Warfare », Chatham House.
— Doshi-Velez, F. et Kim, B. (2017), « Towards a rigorous science of interpretable machine learning ».
— Endsley, M.R. (1995), « Toward a theory of situation awareness in dynamic systems », in *Human Factors*, 37(1), 32–64.
— Fastrez, P., Campion, B. et Collard, A.S. (2009), « Le tri de cartes », in *Document numérique*, 12(2), 23-45.
— Fréjus, M. (1999), « Analyser l'activité d'explication pour concevoir en terme d'aide : application à la formation et à la négociation commerciale », thèse de doctorat, Université Paris 5 René Descartes.
— Gornet, M. et Maxwell, W. (6 juillet 2023), « L'IA explicable appliquée à la détection de ceintures et de téléphones au volant », https://hal.science/hal-04158889
— Hollnagel, E. (2017), *Safety-II in practice: Developing the resilience potentials*, Londres, Routledge, 144 p.
— Karsenty, L. (1996), « Une définition psychologique de l'explication », in *Intellectica, revue de l'Association pour la recherche cognitive*, 23(2), 327345.
— Kirsch, A. (2017), « Explain to whom? Putting the User in the Center of Explainable AI », https://hal.science/hal-01845135
— Kizilcec, R.F. (2016), « How Much Information? », The 2016 CHI Conference, 23902395.
— Marzano, M. (2010), *Le Contrat de défiance*, Paris, Grasset, 320 p.
— MIT Media Lab, Project NANDA (juillet 2025), « The GenAI Divide: State of AI in Business 2025 », https://mindmapai.app/mind-mapping/state-of-ai-in-business-2025-the-genai-divide
— Miller, T. (2019), « Explanation in artificial intelligence: Insights from the social sciences », in *Artificial intelligence*, 267, 1-38.
— O'Neil, C. (2016), *Weapons of Math Destruction, How Big Data Increases Inequality and Threatens Democracy*, New York, Broadway Books, 275 p.
— Parasuraman, R. et Riley, V. (1997), « Humans and Automation: Use, Misuse, Disuse, Abuse », in *Human Factors, The Journal of the Human Factors and Ergonomics Society*, 39(2), 230253.
— Rudin, C. (2019), « Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead », in *Nature Machine Intelligence*, 1(5), 206215.
— Shneiderman, B. (2020), « Human-Centered Artificial Intelligence: Three Fresh Ideas », AIS Electronic Library (AISeL).
— Van Belleghem, L. et Barcellini, F. (2011), Démarche ergonomique de conduite de projet, cours « ERG 220 : Ergonomie et conception : modèles et outils pour l'action ».
— Verma, A., Kumar, Y. et Kohli, R. (2021), « Study of AI Techniques in Quality Educations: Challenges and Recent Progress », in *SN Computer Science*, 2(4), 10.1007/s42979-021-00635-3.
— Vuarin, L. et Steyer, V. (2023), « Le principe d'explicabilité de l'IA et son application dans les organisations », in *Réseaux*, n° 240(4), 179210.
— Weick, K.E. et Sutcliffe, K.M. (2007), *Managing the Unexpected: Resilient Performance in an Age of Uncertainty (2nd ed.)*, San Francisco, Jossey-Bass, 208 p.
— Wilson, H.J. et Daugherty, P.R. (2018), *Human + Machine: Reimagining Work in the Age of AI*, Brighton, Harvard Business Review Press, 264 p.
— Zamfirescu-Pereira, J., Wong, R.Y., Hartmann, B. et Yang, Q. (2023), « Why Johnny Can't Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts », in *CHI'23*, 121.

{Travailler les IA : enquêter, concevoir, discuter, réguler}

73 PARTICIPER À L'IA AUTREMENT : RETROUVER PRISE SUR UN OBJET FUYANT PAR L'ENQUÊTE COLLECTIVE

Pauline Gourlet

87 CONCEVOIR DES SYSTÈMES D'IA ATTENTIFS AUX CONDITIONS DE TRAVAIL : RETOUR RÉFLEXIF SUR UNE DÉMARCHE PARTICIPATIVE ET INTERDISCIPLINAIRE

Maria Sol Romero Goldar, Zacharie Hellouin de Cenival

96 POURQUOI LE DIALOGUE SOCIAL TECHNOLOGIQUE NE PEUT-IL PAS FAIRE DE MIRACLES À L'HEURE DES IAS ?

Vincent Mandinaud, Ludovic Bugand, Nicolas Vispi, Grégory Plançon

105 INTELLIGENCE ARTIFICIELLE ET PRÉVENTION DES RISQUES PSYCHOSOCIAUX : UNE DÉMARCHE RÉFLEXIVE DE NANTES MÉTROPOLE

Marie Bernard, Priscilla Capovilla

112 DIALOGUER SUR L'IA ? RETOURS SUR UNE EXPÉRIMENTATION ORGANISATIONNELLE COLLABORATIVE : LE PROJET DIAL-IA

Odile Chagny, Christophe Teissier

PARTICIPER À L'IA AUTREMENT : RETROUVER PRISE SUR UN OBJET FUYANT PAR L'ENQUÊTE COLLECTIVE

Pauline Gourlet
(Laboratoire i3, Télécom Paris, IPP)

Les débats sur l'intelligence artificielle sont largement structurés par des récits qui tendent à réifier les technologies, à invisibiliser les transformations concrètes du travail et à cantonner la participation à des dispositifs éloignés des situations réelles. En mobilisant les travaux sur l'économie de la promesse et la notion d'« objet fuyant » d'Yrjö Engeström, cet article analyse trois formes d'« éclipse » – technique, sociale et participative – qui façonnent les développements de l'IA. Il présente ensuite les enseignements d'une enquête collective menée dans le cadre du projet international *Shaping AI,* réunissant une pluralité de praticiens issus de secteurs variés. Fondée sur l'analyse des activités de travail, cette démarche met en lumière les problèmes concrets auxquels l'IA donne prise et ouvre des espaces de problématisation collective. L'article montre ainsi que les activités situées constituent un levier essentiel pour renouveler les formes de participation, construire des capacités d'action collectives et réinscrire les choix technologiques dans un débat démocratique.

Mots-clés : intelligence artificielle, participation, innovation sociotechnique, objet fuyant, enquête collective

« There is more regulation on selling a sandwich to the public than to develop potentially lethal technology[1] », commente Connor Leahy, chercheur en intelligence artificielle (IA) et entrepreneur, dans une interview donnée à CNN en 2023. Cette déclaration, devenue virale, témoigne d'un sentiment répandu d'« absence de prise » quant aux développements de l'intelligence artificielle. Alors même que se multiplient les appels à une participation accrue du public, à une vigilance éthique et à un encadrement politique, l'IA semble demeurer insaisissable, sorte d'objet fuyant, qui se dérobe aux nombreuses tentatives de définition, de stabilisation et de régulation.

Dans le prolongement des travaux sur l'économie de la promesse (Joly, 2015), cet article analyse les développements de l'IA au travers de trois phénomènes d'éclipse jouant simultanément sur les plans

1 « Il y a plus de réglementations sur la vente d'un sandwich au public que sur le développement d'une technologie potentiellement mortelle » (notre traduction). Source : https://edition.cnn.com/videos/tv/2023/05/02/amanpour-leahy-ai.cnn

de la technique, du social et de la participation. Ce jeu de mise en (in)visibilité dessine un horizon d'attente construit sur l'effacement des activités et des situations concrètes par lesquelles l'« IA » se réalise. Une telle perspective invite à prendre au sérieux le concept d'*objet fuyant* en se référant aux travaux d'Yrjö Engeström, qui met au centre des processus d'innovation les interactions entre des systèmes d'activités hétérogènes et voit dans leurs opérations concrètes un potentiel de participation sous-considéré. L'enjeu n'est donc pas tant de dénoncer les effets d'invisibilisation produits par les discours des faiseurs de promesses, que de redonner une légitimité aux activités situées – et notamment aux activités de travail – comme lieux d'enquête et de co-production du social et du technique.

Après avoir décrit ces jeux de mise en visibilité et d'occultation à l'œuvre dans les discours sur l'IA, je présente la méthode et les résultats d'une enquête collective qui s'est tenue dans le cadre d'un projet de recherche internationale *Shaping AI*. Je discute la capacité d'un tel dispositif, conçu pour re-problématiser l'« IA » à partir d'une pluralité d'activités situées, à générer des prises susceptibles d'influencer les développements technologiques futurs.

1. TROIS ÉCLIPSES AU SERVICE DE LA PROMESSE

L'intelligence artificielle fait figure d'innovation de rupture ; certains parlent même d'une quatrième révolution industrielle (Schwab, l'UNESCO). Disruptive et en continuelle évolution, elle exigerait la révision des grilles d'analyse et des procédés classiques de description, de régulation et d'intégration des innovations techniques. Cette exceptionnalité de l'« IA », loin d'être démontrée, participe d'un mode de gouvernement de l'innovation que Joly (2015) décrit comme une économie de la promesse. Les très nombreux récits qui narrent ses développements spectaculaires, plus performatifs que descriptifs (Gourlet et al., 2024), contribuent dans une large mesure à invisibiliser les processus d'infrastructuration qui s'opèrent bel et bien en arrière-plan. C'est au regard de ce mode de gouvernement que l'on peut comprendre les « éclipses » qui régissent les jeux d'ombre et de lumière déterminants dans les développements de l'IA.

1.1. L'éclipse technique : réifier une force autonome et puissante

Un présupposé largement répandu dans les discours sur l'intelligence artificielle consiste à considérer que le terme « IA » renvoie à un objet technique stable, clairement identifiable et doté d'une définition partagée. Cette fausse évidence repose sur une opération de réification et de stabilisation, qui « fait tenir » l'IA comme une entité technique autonome et cohérente. Celle-ci ne reflète pas la réalité historique de ses développements mais agit au service d'un récit civilisationnel : promesse de progrès pour beaucoup, menace existentielle pour certains (Gourlet et al., 2024).

Le contenu de la promesse importe moins que l'instauration d'une relation entre les problèmes des uns et les solutions des autres (Joly, 2015) ; dans le cas de l'IA, l'indifférence au contenu concret de la promesse est particulièrement manifeste. Depuis plus de 70 ans, la « révolution » de l'IA est annoncée comme imminente, alors même qu'elle repose sur des entités techniques qui, au cours de cette histoire longue, ont largement évolué jusqu'à ne plus avoir grand-chose de commun (Cardon et al., 2018)[2]. Cette plasticité est précisément ce qui permet à une promesse socio-technique de circuler et de rallier sans cesse de nouveaux acteurs, dans la perspective d'opérer une fermeture du futur afin que les innovateurs puissent mobiliser des ressources et stabiliser leur environnement (van Lente, 1993).

2 Les développements de ce qu'on a nommé « IA » depuis les années 50 ne recouvrent pas les mêmes réalités techniques : celles-ci ont évolué, à la fois en termes d'approches et de modélisations mathématiques (IA connexioniste – IA symbolique) et en termes de réalités matérielles, que l'on pense à l'évolution de la puissance de calcul, des infrastructures et à la miniaturisation des puces (de l'IBM « 704 » à cartes perforées aux centres de données *hyperscale*).

Ainsi, la mise en scène de l'IA, à la fois comme objet stable et comme horizon inéluctable, remplit pleinement sa fonction de colonisateur du présent (Joly, 2015 ; Giddens, 1998) : en créant un état de nécessité irréversible, il devient alors difficile d'imaginer des futurs dans lesquels l'IA ne serait pas.

Cette dynamique est amplifiée par les discours sur la puissance de l'IA, qui l'érigent en force autonome, susceptible d'échapper à tout contrôle humain. Ces récits, au fondement du courant « éthiciste », portés par les plus grands « experts » mondiaux[3] et relayés massivement par les médias, donnent du crédit à la réification de l'IA, jusqu'à naturaliser son existence. Or, comme le rappelle Lucy Suchman (2023), il est étonnant que ce « signifiant flottant » soit si rarement remis en cause. Dans une perspective à la fois ontologique et épistémologique, elle s'interroge : à qui profite la stabilité apparente de cette figure « IA » ? Pour elle, la tâche critique des sciences sociales consiste précisément à ouvrir le champ de la controverse pour déstabiliser cette représentation, en montrer les effets et les rapports de pouvoir qu'elle dissimule. Autrement dit, au lieu de prendre l'IA pour un objet déjà constitué, il s'agit d'examiner comment cette catégorie s'installe, se maintient et se justifie dans les pratiques, les discours et les infrastructures qui la réalisent.

1.2. L'éclipse des rapports sociaux : *l'infrastructuration* silencieuse du présent

Au fur et à mesure de l'intégration des technologies de calcul dans le quotidien, leurs effets sur la société sont questionnés, et avec eux les intérêts qu'ils servent. Le philosophe des techniques David Grunkel prévient : « *These large corporations are in effect running a very large-scale experiment on all of humanity*[4]. » Pourtant, ces expérimentations sociales à grande échelle se déploient dans un régime de faible visibilité. Les dispositifs numériques se banalisent dans les environnements de travail, les pratiques professionnelles et les routines quotidiennes, inscrivant dans la durée des formes spécifiques de capture, de calcul et d'évaluation. Marres (2025) montre que, loin de rendre les transformations socio-techniques disponibles à l'examen public, les expérimentations numériques contemporaines tendent à les rendre implicites et à naturaliser les conditions de leur fonctionnement. Le régime d'expérimentation associé aux développements numériques récents transforme les environnements où ils se déploient en infrastructures de test, configurées pour la collecte et la gestion des données. Ces processus discrets d'infrastructuration ne sont pas neutres : ils réorganisent les rapports sociaux dans une continuité entre production et usage et changent la grammaire des responsabilités, en exigeant des « bêta-testeurs » qui s'ignorent une mise en conformité avec les conditions expérimentales.

Dans les arènes publiques, on reste attaché à la séparation entre production et usages qui prédomine encore dans le monde de la conception. Des risques sont anticipés dans une perspective de régulation, et les appels à la morale des ingénieurs de l'industrie numérique[5] se multiplient. Les yeux rivés sur le futur promis, cadré par les faiseurs de promesses, on perd de vue ce qui, au présent, est en train de s'infrastructurer durablement.

Lorsque ces transformations donnent lieu à des analyses critiques, ce sont souvent les dynamiques macro-structurelles qui sont mises en avant – « ère de la smartness » (Halpern et Mitchell, 2023), « technoféodalisme » (Varoufakis, 2024), ou encore « nouvelle discipline du travail » (Pasquinelli, 2023). En restant à distance des activités par lesquelles l'IA se réalise concrètement, ces analyses passent à

3 En 2023, une lettre ouverte signée par plusieurs figures proéminentes de l'IA – Yoshua Bengio, Elon Musk, Steve Wozniak, Yuval Noah Harari, ou encore Raja Chatila et Laurence Devillers pour la France – appelait à « mettre en pause » les grandes expérimentations d'IA, le temps de « prendre du recul » sur cette technologie menaçante pour la vie.

4 « Ces grandes multinationales mènent en réalité des expérimentations à grande échelle sur l'ensemble de l'humanité » (notre traduction) Source : https://www.theguardian.com/technology/2025/sep/09/ai-chatbot-love-relationships

5 On en voit un exemple dans un article co-écrit par Luccioni et Bengio en 2019, qui en appellent à la morale des concepteurs de systèmes d'IA pour guider les choix de conception et éviter des risques majeurs pour « l'humanité » (Luccioni et Bengio, 2019). Dans sa version française, on trouve des initiatives comme le « Serment d'Hippocrate pour Data Scientist ».

côté des multiples performativités situées et des problèmes qui se posent concrètement au sein des pratiques. Cet éclairage apparait pourtant déterminant pour renouveler l'étude politique et sociale de l'IA, dans une perspective de réappropriation et de partage de la décision quant aux développements technologiques.

1.3. L'éclipse du public : la participation limitée

L'impératif contemporain de « participation » tente de prendre en charge cet enjeu de pluralisation dans la gouvernance des technologies. Deux formes de participation dominent (Gourlet et al., 2024) : d'une part, le co-design, qui cherche à intégrer les considérations éthiques, juridiques et sociales dans les expérimentations et solutions techniques, et d'autre part, la délibération publique, invitant la société civile à discuter et anticiper des risques. Ces deux approches réaffirment la frontière entre conception et usage, entre production et régulation, et en reproduisent les écueils. Du côté de la conception, la focalisation sur les solutions techniques donne peu de latitude pour discuter des effets sociaux des innovations. De l'autre, un partage des rôles entre experts et citoyens perpétue une hiérarchie des savoirs qui amplifie le point de vue des producteurs et experts de la technique tout en affaiblissant celle des autres et en minimisant les expériences vécues.

Ces deux approches négligent le fait que l'objet « IA » se forme d'abord au sein d'activités qui s'en *soucient* – et que celles-ci ne sauraient se limiter à la production d'un côté et à la régulation de l'autre. La question devient alors : de quelles activités l'IA est-elle l'objet ? Comme le remarque l'anthropologue Nick Seaver dans son ethnographie d'une entreprise de recommandation, l'IA « semble n'être nulle part en particulier[6] » (2017). Lors de son enquête, toutes les personnes de l'entreprise ont situé leurs activités à sa périphérie. L'IA apparait alors comme un « objet fuyant » qui se réalise dans la multiplicité des systèmes d'activités et relie des mondes hétérogènes sans se laisser saisir par aucun. Suivant Engeström (2006), de tels objets ont le potentiel de s'étendre et de s'amplifier jusqu'à atteindre une échelle d'influence mondiale, générateurs d'effets vastes et inattendus, sans que quiconque n'exerce un contrôle exclusif. Ces objets ressemblent à des « monstres », dit-il (ibid.) : ils semblent dotés d'une vie propre, menaçant notre sécurité et notre stabilité de multiples façons, suscitant controverses, oppositions et débats.

C'est à une vision du politique ramenée dans le quotidien des activités situées de travailleurs qu'invite la notion d' « objet fuyant » d'Engeström (2006). En effet, en envisageant d'abord cet objet comme ce qui organise une diversité d'activités, ce concept nous permet de tenir ensemble à la fois l'horizon que se donnent les acteurs et sa traduction dans des schèmes d'actions et des opérations propres à chaque activité. Méthodologiquement, saisir un objet fuyant demande d'enquêter depuis les *publics* de l'IA (Dewey, 1927), c'est-à-dire avec les personnes soucieuses des conditions de sa constitution et de ses effets au sein de leurs activités. Dans cette perspective, la participation cesse d'être un idéal normatif pour devenir une pratique d'enquête collective, qui part des activités de travail et redistributive du pouvoir d'agir. Se doter de tels moyens d'enquête vise à retrouver des prises pour influencer les développements technologiques et permet d'éviter les problèmes qui génèrent de l'emprise en leur opposant des problèmes construits depuis des activités réelles.

6 « Même les personnes avec les rôles les plus "algorithmiques" de l'entreprise, qui travaillaient sur les infrastructures de *machine learning* ou sur la personnalisation des playlists, **situaient l'algorithme juste en dehors du champ de leur travail, quelque part dans le code de l'entreprise** » (notre traduction et notre soulignement). Source : Seaver, 2017.

2. MISE À L'ÉPREUVE : LES ACTIVITÉS DE TRAVAIL COMME LIEU D'ENQUÊTE ET DE PARTICIPATION

En 2021, nous avons démarré un travail de recherche sur les problèmes publics de l'IA au médialab de Sciences Po, dans le cadre d'un projet international, *Shaping AI*[7]. Méthodologiquement, nous avons tenté de mettre à l'épreuve cette notion d'« objet fuyant » à travers un protocole original d'enquête collective qui élargit le champ des activités ayant pour objet l'IA. L'objectif était d'étudier les frictions dans et entre systèmes d'activités pour problématiser collectivement l'IA à partir des activités de travail. Pour éviter de reproduire la séparation artificielle entre « producteurs » et « usagers » et ne pas limiter *a priori* les activités *soucieuses* de l'« IA », nous avons cherché à pluraliser les secteurs d'activités, en plus de solliciter des co-enquêteur·ice·s aux parcours, âges, statuts et positionnements variés. Journalistes, juristes, artistes, chercheurs, mathématiciens, cadres de la fonction publique, mais également start-uppeurs, syndicalistes et activistes : une trentaine de « praticien·ne·s de l'IA » se sont au final engagés dans un processus réflexif au long cours[8], avec pour objectif de donner corps à « l'objet fuyant IA ». Nous tirons ici quelques leçons de cette mise à l'épreuve en nous attachant à montrer les décalages dont est porteur un tel dispositif d'enquête et en discutant les ouvertures qu'il a permis de générer.

2.1. L'enquête collective : problématiser l'IA à partir des *soucis* de praticien·ne·s

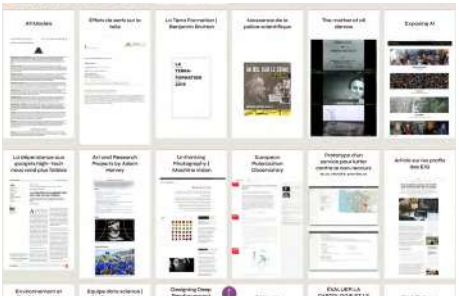
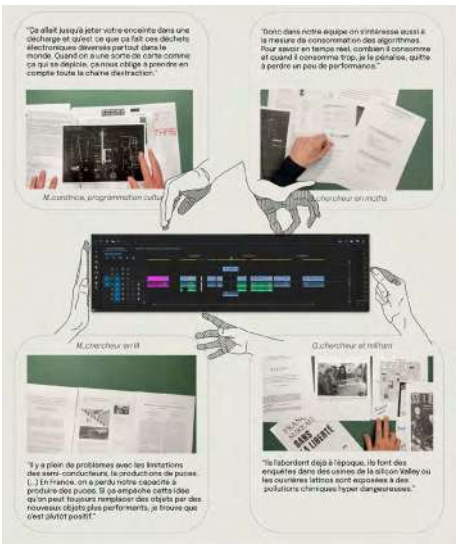
Mêlant une démarche de design participatif (Ricci, 2019) et de clinique de l'activité (Clot et al., 2001), le dispositif a été conçu de manière à pluraliser, expliciter et faire résonner les différents points de vue. Avec chacun·e individuellement d'abord, puis collectivement, nous avons cherché à recueillir des descriptions épaisses des activités, tout en prenant soin d'en déplier les enjeux, les jeux de frontière, les circulations et les tensions à l'interface de plusieurs sphères d'activités, comme en témoignent les deux extraits d'entretien ci-dessous. Non pas des chroniques d'activités qui donneraient à voir des interactions avec des systèmes techniques (classiques dans les analyses du travail), ni des recueils de représentations ou de sentiments (comme peut le faire la sociologie), ces descriptions matérialisent des situations de travail où l'« IA » s'instancie dans une forme singulière.

> « En 2017, on contacte une journaliste du *Monde* qui vient d'écrire un article sur un projet d'observatoire *big data* de la tranquillité publique à Marseille. En lisant le projet, ça semblait évident que c'est du croisement de *big data* avec une idée de police prédictive. (…) ça correspond à une logique qu'on observait déjà dans d'autres marchés ailleurs dans le monde. Ça nous choquait un peu de voir les aspects sur la surveillance et les risques en termes de libertés publiques complètement éludés dans cet article. (…) Suite à ça, on veut en savoir plus, première demande CADA. »
> *Entretien avec F., chercheur en sciences politiques et activiste.*

> « Là, c'est les contacts. En étant chercheur, on a des contacts. Par exemple avec les étudiants, on encadre des mémoires de Master, donc on a des contacts avec des entreprises. C'est prendre son téléphone et raconter une histoire. Et là, j'ai vraiment été surpris. Je m'en doutais un peu mais la parole d'un chercheur comme ça, vers 2015, spécialiste des algorithmes de *machine learning*, vu le boom qu'il y avait de l'IA, on est très écouté. Et les gens nous suivent. »
> *Entretien avec S., chercheur en IA.*

Concrètement, l'enquête collective s'est déroulée en cinq étapes :

7 Pour plus d'informations, voir le site du projet : shapingai.org
8 L'enquête s'est déroulée pendant neuf mois, entre 2021 et 2022. Au total, 29 personnes ont participé à l'ensemble du processus décrit dans le tableau 1.

ÉTAPES	DESCRIPTION DU DISPOSITIF	ILLUSTRATIONS DU DISPOSITIF	DONNÉES PRODUITES
1. Entretiens biographiques	Première rencontre individuelle : entretien ouvert (2 h) sur les expériences vécues dans lesquelles l'IA se concrétise et explicitation de tensions ou enjeux.		(29 entretiens individuels de 2 h réalisés d'octobre à mars) Enregistrements audio et retranscriptions
2. Constitution d'une archive documentaire	Matérialisation des éléments mentionnés dans les biographies : recherche de traces tangibles, principalement trouvés sur Internet. (exemples : actes de colloque, articles de presse, manifestes militants, textes réglementaires, échanges sur un forum en ligne, encore contenus de cours, librairie open source)		1 230 documents (en moyenne, 45 par personne)
3. Explicitation/ formulation des soucis à partir de l'archive	Deuxième rencontre (individuelle) : session de formulation et d'explicitation des soucis (± 2 h). Tri et sélection des documents (étalés sur un mur), puis explication/élaboration sur une table montée d'une caméra.		29 vidéos
4. Composition des *soucis* collectifs	Traitement et montage des vidéos : analyse thématique des soucis individuels en vue de composer des « soucis collectifs ». Les manipulations et les gestes associés aux documents sont autant de ressources pour l'analyse que les commentaires oraux.		19 vidéos (entre 12 et 15 minutes chacune) + constitution des dossiers (documents réunis par souci)

5. Discussion collective : problématisations de l'objet fuyant IA et esquisse de prises communes	Workshop à la Cité des sciences et de l'industrie (22 juin 2022). Les co-enquêteur·ice·s sont réunis pendant deux jours pour discuter et amender les 19 soucis analysés. Discussion générale sur les effets de l'enquête et du dispositif.		6 vidéos

Tableau 1. Les étapes du dispositif d'enquête collective de *Shaping AI,* conçu par l'équipe du médialab de Sciences Po.

Cette démarche, empirique et collective, a d'abord permis de « dé-faire l'IA » – de lui donner de la consistance en tant qu'objet fuyant – en en resituant la constitution dans des histoires, des cours d'action et des gestes concrets qui obligent à revisiter à nouveaux frais les cadrages dominants. D'autre part, l'analyse a fait émerger un espace de problèmes alternatifs, construits à partir de « soucis[9] » décrits au sein de plusieurs activités situées. Comme le montrent les verbatims de la figure 1, la problématique des « biais » des IA apparait sous deux lumières bien différentes selon qu'elle émerge dans une activité dont il est l'objet (ici, éviter les biais de détection et les artefacts dans une imagerie cérébrale) ou qu'elle est mobilisée comme élément de neutralisation de la critique dans un processus de déploiement technologique mené sans concertation (là, dans une administration publique).

Ainsi, 19 *soucis* rendent compte de problèmes situés contribuant à la constitution de l'IA[10]. Sa rematérialisation, dans les documents et les explicitations des co-enquêteur·rice·s, donne à voir des registres d'action propres à certains milieux, mettant en lumière des stratégies de conservation et de circulation déterminantes dans la carrière de certaines notions et dans la fabrication de « prises », comme en témoigne l'exemple de la figure 1.

9 Le terme de souci a été choisi pour son ambivalence (« être préoccupé par » et « être soucieux de ») qui met en avant la manière dont les praticien·ne·s rencontrent, construisent et prennent soin de problèmes émergents au sein de leur activité.
10 Pour des raisons de concision, je renvoie ici le lecteur vers les résultats de cette analyse et vers l'ensemble des sources de l'enquête : https://medialab.github.io/ShapingAI

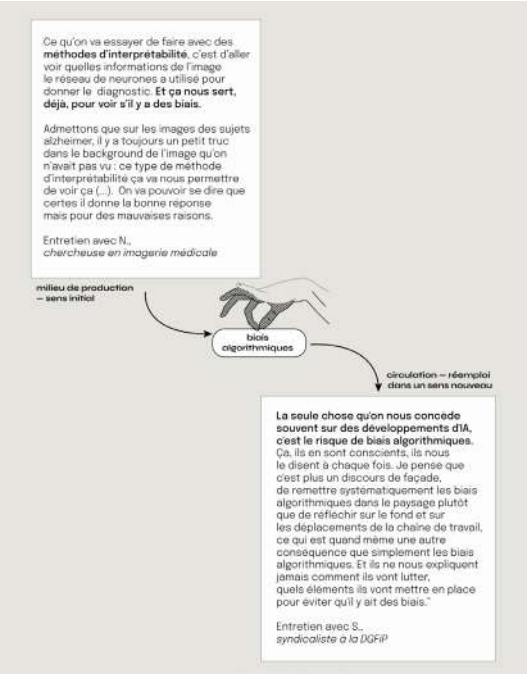


Figure 1. Exemple de re-prise : saisie ici d'abord à partir d'une activité de production de modèles d'IA pour la détection de tumeurs, puis reprise avec un sens nouveau au sein d'une autre activité de négociation du déploiement d'un système au sein d'une administration. Il y a donc formation de « prise », opérée par le glissement opportun d'un souci vers un autre milieu qui transforme le sens initial : du souci des biais qui porte sur la « bonne représentation du réel » au souci de l'adhésion dans le cadre du déploiement d'un système d'intelligence artificielle (SIA) dans l'administration.

2.2. Problématiser collectivement l'IA et construire des prises

Une réflexion collective a été menée (étape 5) relativement à cet « espace des problèmes de l'IA » et aux décalages qu'il permet d'opérer. Six groupes de travail se sont formés et ont proposé six problématisations, sous forme de vidéos : 1) La représentation du réel ; 2) La légitimité démocratique des choix techniques ; 3) Les capacités d'action avec les calculs ; 4) La performativité des récits ; 5) L'écologie de l'IA ; 6) La transformation numérique de l'action publique. Bien que synthétiques, ces vidéos rendent bien compte de la manière dont les systèmes d'activités des uns et des autres entrent en interaction dans la constitution de l'IA et génèrent, ou non, des prises sur ses devenirs. Cette section présente trois des six problématisations. Pour chacune, un résumé de la problématisation est suivi d'un tableau qui donne à voir les soucis exprimés, depuis quelles pratiques, et avec les prises qui pourraient permettre de les prendre en charge.

a. Les capacités d'action avec les calculs

Cette problématisation est restituée à travers un passage en revue commenté des types de rapports entre « calculs, calculés et calculateurs » : quelles opérations de calcul sont effectuées par les systèmes d'IA, que calculent-ils à propos de quoi et de qui ? Par qui sont-ils opérés, à partir de quelles données et à quelles fins ? Des cas concrets sont documentés pour donner à voir des configurations variées, allant de l'exclusion (comme dans l'exemple des calculs de la CAF à partir des travaux de l'association Changer de Cap, cf. tableau ci-dessous) à l'intégration des calculés (comme dans l'exemple des travaux de l'association Inter-hop, cf. tableau ci-dessous). Sans qu'un point de vue normatif se dégage – sinon l'attachement à la capacité d'agir des individus et des collectifs –, les co-enquêteur·rice·s appellent à l'émergence d'espaces de discussions et de négociations collectives ouvertes depuis les expériences d'agencements concrets. Ils opposent ainsi à une forme de « participation publique », des formes de participation située dans une négociation des rapports entre calculs, calculés et calculateurs.

« La participation, il faut donc la comprendre en cycle, dans un mouvement qui doit être entretenu. On sera calculé, il faut prendre part à ce cycle de manière itérative. »

PROBLÉMATISATION	PRATICIEN·NE·S	SOUCIS	PRISES
« La question qu'il nous semblait utile de traiter dans cette idée de participation publique, c'était : **comment négocier avec le calcul et ses troubles ?** »	Changer de Cap (asso)	Étudier le fonctionnement et les conséquences des algo. de la CAF	Enquête, appel à témoignages
	Memo Akten (artiste)	Jouer et rendre visible la part d'interprétation dans les calculs des réseaux de neurones	Expérimentations artistiques
	Agence du num. en santé	Acceptabilité sociale	Ateliers de pédagogie, formations
	Inter-hop (asso)	Promouvoir l'usage de logiciels libres dans la santé et données auto-gérées	Conception de chaines socio-techniques alternatives (hébergements, librairie, modèles, etc.), animation de réseau
	CNIL (Linc)	Interface entre approches instituées et émergentes	Publication de rapports
	Représentants du personnel / syndicats	Négocier avec les employeurs lorsqu'un système d'IA est introduit dans l'entreprise	Outils du dialogue social
	Commission consultative des droits de l'homme	Implication des utilisateurs finaux dans le choix du recours à un système d'IA, puis exercer une vigilance continue	Avis
	Juriste	Mener des études d'impact sur les droits fondamentaux	Publication professionnelle
	Designer-chercheur	Vocabulaire normalisateur (« utilisateur ») qui définit des rôles et responsabilités *a priori*	Expérimentations artistiques pour inspirer et générer de nouvelles approches institutionnelles

b. La performativité des récits

Partant du constat que les développements de l'IA sont toujours agencés dans des récits, les co-enquêteur·ice·s ont fait le choix d'adopter une posture réflexive pour déceler comment eux-mêmes sont pris dans la fabrique et la circulation de certains récits de l'IA. Ce qui ressort de leurs témoignages est l'intérêt mais aussi la difficulté de cet exercice réflexif. En effet, la confrontation avec d'autres implicites « *conceptuels, axiologiques et praxiques, qui forment système avec les objets de l'action* » (Béguin, 2004), permet de faire réapparaitre le « monde » dans lequel les pratiques s'inscrivent, comme en témoigne cette journaliste :

> « Par exemple, je voulais illustrer avec cet article que j'ai écrit en 2017. Là, par exemple, c'est un document que m'a sorti l'équipe du médialab. [...] il y a une citation de DeepMind directement où ils disent : "maintenant DeepMind va se concentrer sur d'autres défis." Et parmi ces défis, il y a : "guérir des maladies, réduire notre consommation d'énergie ou inventer une nouvelle matière révolutionnaire. Car si les IA peuvent faire progresser les connaissances et développer de nouvelles stratégies dans d'autres domaines que le Go, leurs découvertes pourraient être capitales." Cette citation est assez significative. Et ça, c'est quelque chose qu'on a bcp retrouvé dans mon début de carrière. [...] Je pense que ça c'est vraiment le récit auquel j'ai participé : l'IA comme source de puissance. »

D'autre part, ce geste réflexif invite à poursuivre l'enquête sur les liens entre dimensions sociales et psychologiques.

PROBLÉMATISATION	PRATICIEN·NE·S	SOUCIS	PRISES
« Comment a commencé notre histoire avec l'IA, càd comment l'IA est intervenue dans nos parcours professionnels ? Et par là, **comment nous-mêmes, dans nos activités professionnelles, on participe aux récits, au pluriel, de l'IA ?** »	Le groupe	Identifier les récits qui accompagnent l'IA et leur performativité	Décortiquer l'article fondateur de Turing (1950)
	Chercheur en maths	Étudier les opérations et les effets des algorithmes	Développement d'algorithmes, Participation à la communauté de recherche en IA
	Journaliste	Traiter des sujets de la tech pour une large audience	Diffusion de la parole des leaders/ producteurs de la tech
	Artiste	Rendre visibles les failles, défaillances et mensonges qui se cachent derrière les promesses et la hype de l'IA	Exposition d'une fresque historique du travail humain dans l'automatisation

c. La transformation numérique de l'action publique

Deux co-enquêteurs, en franc désaccord, ont décidé ici de partir d'un souci commun, celui du « bon » fonctionnement de l'administration, pour exposer leurs divergences. Le rythme des arguments est rapide et dessine progressivement des frontières nettes, entre une vision dans laquelle l'État doit assurer le rôle de garant des libertés individuelles (et notamment la liberté d'entreprendre) dans une forme d'intervention minimale, et une vision d'un État social et interventionniste, garant des services publics et des intérêts collectifs qui priment sur les intérêts privés. Si les co-enquêteurs partagent certaines préoccupations

quant aux développements de l'IA (surveillance, nécessité de contre-pouvoirs efficaces), l'un voit dans l'IA « des outils techniques » efficaces pour mieux gérer une action publique défaillante (notamment grâce à l'introduction de nouveaux acteurs privés), quand l'autre y lit un mot d'ordre techniciste visant à privatiser et déshumaniser les services publics. Leur dialogue permet d'éclairer une nette asymétrie dans les prises disponibles, l'un disposant d'un arsenal juridique et de possibilités d'expérimentations concrètes, lorsque l'autre dispose quasi exclusivement de jeux d'influence sur l'opinion publique.

> « - Moi, mon espoir, c'est que le développement de l'IA oblige les administrations à se conformer aux règles de l'open data puisqu'elles existent, qu'elles sont mal appliquées et qu'elles sont au service du plus grand nombre. Donc tout le monde y trouve son compte, administrations comme administrés. Moi je pense beaucoup à l'open data comme une externalité positive des développements de l'IA.
> - L'idée n'est pas d'avoir une position dogmatique, mais la question c'est : "démocratiquement, qu'est-ce qu'on en fait ?" Et pour moi, le vrai sujet de l'open data, il est là. Est-ce que ces données doivent être confiées à des intérêts privés ? »

PROBLÉMATISATION	PRATICIEN·NE·S	SOUCIS	PRISES
« **L'IA peut-elle vraiment améliorer l'efficacité de l'action publique ?** C'est précisément le point de notre discussion »	Syndicaliste + juriste	Augmentation des capacités de l'État à surveiller la population (data mining sur les réseaux sociaux, contrôles CAF, interconnexion des données entre administrations, etc.)	Médiatisation des affaires
	Syndicaliste (administration)	Non-contrôle des expérimentations technologiques	Recours juridiques Pression politique
	Syndicaliste (administration)	Absence de preuve d'efficacité ou de gains, construire un contre-discours à la promesse	Documentation des transformations du travail réel
	Syndicaliste (administration)	Dégradation des services publics et numérisation comme correctif	×
	Juriste d'entreprise	Renforcement des droits des citoyens par le numérique	Open data
	Syndicaliste (administration)	Normativité des modèles utilisés pour faire des projections (ex. leximpact)	×
	Juriste d'entreprise	Collaboration entre public et privé, notamment sur le partage des données	Communs numériques Droit souple (chartes, etc.)

CONCLUSION - VERS UNE *RE-PRISE* DES FUTURS TECHNOLOGIQUES ?

Ce dispositif collectif de recherche a permis de suspendre, au sein de l'enquête, les jeux d'éclipse pour braquer le projecteur sur des performances situées qui affectent le développement de l'IA. Les descriptions fines d'activités de travail variées, quel que soit leur degré de proximité *a priori* avec les dimensions techniques ou les problèmes supposés de l'IA, ont offert la possibilité de resituer la constitution de cet objet fuyant dans des rapports sociaux, avec leur historicité et leur matérialité. L'IA change ainsi de nature : d'une entité technique puissante, elle devient le résultat d'un processus de sédimentation, un travail de frontières au long cours entre différentes pratiques, construites par leurs soucis, leurs ressources et leurs jeux de prises.

En créant ce déplacement, l'espace de l'enquête fait une *différence* importante par rapport aux nombreuses autres arènes de délibération sur l'IA, *différence* que les co-enquêteur·rice·s ont soulignée et appréciée. Les matériaux recueillis et discutés (tant l'archive que les vidéos) ont donné à voir une palette de gestes, de manières de faire et d'être, qui sont autant d'attitudes de *participation* possibles. De plus, des capacités à sentir, à se rendre attentifs, à se mettre à d'autres places, à être « idiots » (Stengers, 2007) et à assembler des fragments du puzzle qui paraissaient sans rapport ont pu être expérimentées. Au-delà d'une lecture renouvelée de ce qui est en jeu dans l'ici et maintenant des différentes situations, travailler depuis et entre les systèmes d'activités ouvre une brèche dans la colonisation du présent mise en œuvre par l'économie de la promesse, en rendant tangible les interdépendances au sein des chaines socio-techniques. De nouvelles prises peuvent ainsi s'imaginer pour prendre soin de ces nouvelles solidarités, selon des alliances à géométries inédites.

Ce dispositif d'enquête tente ainsi de renouveler les formes de participation en articulant trois lignes de temps : le temps de l'enquête, le temps des pratiques et le temps de la trajectoire de développement. Ce parti pris permet d'éviter deux écueils classiques : le premier est celui du simulacre, qui fait de la participation un mécanisme de fabrique d'acceptabilité (Angeli Aguiton, 2014). Le second concerne le fait d'examiner les problèmes sur le plan de l'ajustement en situation, qui a tendance à diviser plutôt que collectiviser la charge de résolution des problèmes et à renforcer un sentiment d'impuissance ou de geste minuscule. Comme le notait déjà Engeström (2015) : « *We need intermediate runaway objects which are less spectacular and more inviting*[11] »

Par cette enquête se sont élaborés des problèmes-frontières qui neutralisent la pensée par échelle et permettent de dessiner de nouvelles prises. Néanmoins, il ressort de ce processus et des problématisations qui ont émergé que le rapport de force, largement favorable aux faiseurs de promesses, alimente en continu les éclipses de la technique, du social et de la participation. Malgré la pluralité des prises déployées, celles-ci peinent à avoir des effets concrets sur les développements en cours. Mais comme nous y incite Chateauraynaud (2022), rendre visible la politique des promesses technologiques est une première étape nécessaire pour neutraliser leur emprise.

> « En explorant ce qui permet à des coalitions d'acteurs de maintenir un ordre social foncièrement injuste et toxique, malgré des crises majeures, il faut en effet viser les procédés d'emprise par lesquels des entités en tiennent d'autres sous leur contrôle dans un savant mélange de manipulation, d'exploitation des asymétries, de jeu avec les contraintes organisationnelles et de servitude volontaire » (Chateauraynaud, 2022).

11 « Nous avons besoin d'objets fuyants intermédiaires qui sont moins spectaculaires et plus engageants » (notre traduction).

BIBLIOGRAPHIE

— Angeli Aguiton, S. (2014), « La démocratie des chimères : Gouvernement des risques et des critiques de la biologie synthétique, en France et aux États-Unis » (thèse de doctorat, Paris, Institut d'études politiques). https://www.theses.fr/2014IEPP0055

— Béguin, P. (2004), « Monde, version des mondes et monde commun », in *Bulletin de psychologie*, 57(1), 4548.

— Cardon, D., Cointet, J.-P. et Mazières, A. (2018), « La revanche des neurones », in *Réseaux*, n° 211(5), 173220. https://doi.org/10.3917/res.211.0173

— Chateauraynaud, F. et Debaz, J. (2017), *Aux bords de l'irréversible. Sociologie pragmatique des transformations*, Paris, Éditions Pétra, 648 p.

— Chateauraynaud, F. (2022), « Des expériences ordinaires aux processus critiques non-linéaires. Le pragmatisme sociologique face aux ruptures contemporaines », in *PRAGMATA*, 2022-5 : 18-92.

— Clot, Y., Faïta, D., Fernandez, G. et Scheller, L. (2001), « Entretiens en autoconfrontation croisée : une méthode en clinique de l'activité », in *Éducation permanente*, 146(1), 1725.

— Dewey, J. (1927), *The Public and Its Problems*, New York, Henry Holt and Company, 224 p.

— Engeström, Y. (2006), « Development, movement and agency: Breaking away into mycorrhizae activities », in *Building activity theory in practice: Toward the next generation*, 1, 143.

— Engeström, Y. (2008), « THE FUTURE OF ACTIVITY THEORY: A ROUGH DRAFT* », keynote lecture presented at the ISCAR Conference in San Diego, Sept. 8-13, 2008. This is an expanded and edited version of a chapter with the same title, to appear in 2009 in the book *Learning and Expanding with Activity Theory*, edited by Annalisa Sannino, Harry Daniels and Kris Gutierrez (Cambridge, Cambridge University Press).

— Giddens, A. (1998), « Risk society: The context of British politics », in Franklin, J. (ed.), *The Politics of Risk Society* (p. 23), Cambridge, Polity Press, 152 p.

— Gourlet, P., Ricci, D. et Crépel, M. (2024), « Reclaiming artificial intelligence accounts », in *Big Data & Society,* 2024, 11 (2), https://hal.science/hal-04584823v1

— Halpern, O. et Mitchell, R. (2023), *The Smartness Mandate*, Cambridge, The MIT Press, 336 p. https://mitpress.mit.edu/9780262544511/the-smartness-mandate/

— Joly, P.-B. (2015), « Le régime des promesses technoscientifiques », in Audétat, M. (dir.), *Pourquoi tant de promesses ?* (p. 31-48), Paris, Éditions Hermann, 318 p.

— Marres, N. et Stark, D. (2020), « Put to the test: For a new sociology of testing », in *The British Journal of Sociology*, 71(3), 423443. https://doi.org/10.1111/1468-4446.12746

— Marres, N. (2025), « How Tech Trials Put Society to the Test », in *Theory, Culture & Society,* 0(0). https://doi.org/10.1177/02632764251342911

— Pasquinelli, M. (2023), *The Eye of the* Master, New York, Verso, 272 p. https://www.versobooks.com/products/735-the-eye-of-the-master

— Ricci, D. (2019), « Tensing the Present: An Annotated Anthology of Design Techniques to Inquire into Public Issues », in *Diseña*, 2019, 14, p. 68-99. https://sciencespo.hal.science/hal-02056093v1

— Seaver, N. (2017), « Algorithms as culture: Some tactics for the ethnography of algorithmic systems », in *Big Data & Society,* 4(2), 2053951717738104. https://doi.org/10.1177/2053951717738104

— Sloane, M., Moss, E., Awomolo, O. et Forlano, L. (2020), « Participation is not a Design Fix for Machine Learning », arXiv: 2007.02423 [cs]. http://arxiv.org/abs/2007.02423

— Stengers, I. (2007), « La proposition cosmopolitique », in Lolive, J. et Soubeyran, O. (dir.), *L'Émergence des cosmopolitiques* (p. 47), Paris, La Découverte, 382 p.

— Suchman, L. (2023), « The uncontroversial 'thingness' of AI », in *Big Data & Society*, 10(2), 20539517231206794. https://doi.org/10.1177/20539517231206794

— Turing, A. (1950), « Computing Machinery and Intelligence », in *Mind*, LIX (236): 433–460.

— Van Lente, H. (1993), « Promising Technology: The Dynamics of Expectations in Technological Development », Enschede, Department of Philosophy of Science and Technology, University of Twente.
— Varoufakis, Y. (2024), *Technofeudalism.* , Londres, Penguin Books, 304 p.

CONCEVOIR DES SYSTÈMES D'IA ATTENTIFS AUX CONDITIONS DE TRAVAIL : RETOUR RÉFLEXIF SUR UNE DÉMARCHE PARTICIPATIVE ET INTERDISCIPLINAIRE

Maria Sol Romero Goldar
(Université de Toulouse Jean-Jaurès, LISST-CNRS)
Zacharie Hellouin de Cenival
(Université de Toulouse, INSA, LAAS-CNRS)

L'essor des systèmes d'intelligence artificielle (SIA) dans les milieux professionnels repose souvent sur une logique d'optimisation et de performance, au risque de négliger leurs effets sur l'organisation du travail et la qualité de vie des salarié·es. Cet article interroge les conditions d'une conception plus attentive au travail réel à partir d'une recherche participative et interdisciplinaire réunissant salarié·es, sociologues et chercheurs en informatique et réseaux. L'analyse met en lumière les tensions liées aux temporalités, aux imaginaires et aux hiérarchies disciplinaires, ainsi que les leviers permettant de les dépasser. Il montre que l'interdisciplinarité, loin d'être une simple addition de savoirs, constitue un processus exigeant de négociations et de compromis, ouvrant la voie à la conception de SIA plus attentifs au travail réel et à la qualité de vie au travail.

Mots-clés : interdisciplinarité, conditions de travail, intelligence artificielle, conception participative

INTRODUCTION

La diffusion rapide des systèmes d'intelligence artificielle (SIA) transforme profondément les milieux professionnels. Selon l'enquête de l'Insee sur les technologies de l'information et de la communication dans les entreprises, 10 % des entreprises françaises déclaraient en 2024 utiliser au moins une technologie d'IA, notamment pour l'analyse du langage (44 %), l'exploitation de données (41 %) et l'aide à la décision (33 %). Et si l'on en croit l'enquête Ipsos « L'IA au travail en 2025 », la moitié des salarié·es estime que l'IA a déjà modifié ou modifiera prochainement leur activité. Ces données témoignent d'une adoption significative des SIA, qui transforme les pratiques professionnelles et l'organisation du travail.

Souvent conçus dans une logique d'efficacité et de performance, ces dispositifs ne sont pas neutres (Akrich, 1987). Ils traduisent des visions spécifiques du travail et matérialisent des rapports de pouvoir, tout en étant développés à distance des situations concrètes. Leurs effets sont ainsi difficilement anticipés, pouvant conduire à des usages détournés des finalités initialement visées (Zouinar, 2020 ; Theviot, 2023) ou à des effets délétères sur la qualité de vie au travail (Bobillier Chaumon, 2023 ; Borel, 2024). Dès lors, une question centrale se pose : comment concevoir des technologies fondées sur l'intelligence artificielle qui soient attentives aux conditions réelles de travail ?

Une piste consiste à intégrer, dès la conception, des disciplines spécialisées dans la qualité de vie et les conditions de travail (QVCT), dans le cadre de démarches interdisciplinaires. Toutefois, celles-ci ne reposent sur aucune méthode stabilisée et se construisent dans l'action, à travers des formes de coopération permettant d'articuler des savoirs hétérogènes.

Cet article propose un retour réflexif sur les conditions de possibilité de cette articulation. À partir de notre expérience de jeunes chercheur·euses engagé·es dans un projet participatif interdisciplinaire, nous analysons les modalités concrètes de coopération, les ajustements réalisés et les tensions rencontrées, afin de discuter leur contribution à la conception de technologies plus attentives aux conditions de travail[1].

1. LE PROJET DE RECHERCHE « VERTUOSE »

Notre expérience s'inscrit dans le projet de recherche Vertuose (« *Vers une infrastructure de communication vectrice de pratiques de communication médiées soutenables* »). Il vise à co-concevoir un dispositif sociotechnique destiné à réduire la part communicationnelle de l'activité perçue comme problématique, en articulant les savoirs des chercheur·euses et l'expertise des acteur·rices concerné·es.

1.1. Une recherche participative et interdisciplinaire

Le projet réunit depuis trois ans des sociologues, des chercheurs en informatique et réseaux (IR) et une quinzaine de participant·es d'une organisation œuvrant à l'amélioration des conditions de travail dans les entreprises. Cette organisation a accepté de devenir partenaire d'une recherche combinant enquête ethnographique et ateliers de co-conception. Par ailleurs, un comité de pilotage associant chercheuses et salarié·es de cette organisation garantit une participation effective aux décisions.

La quasi-totalité du collectif (15 sur 16) s'est engagée volontairement dans ce processus de trois ans, voyant dans ce projet une occasion de réfléchir sur la charge communicationnelle dans un contexte de restructuration. Composé de chargé·es de mission, de la responsable hiérarchique et des assistantes

1 Les propos présentés dans cet article résultent des discussions et des réflexions issues de notre expérience en tant que doctorant·es. Les autres membres de l'équipe n'ont ni relu ni validé ce contenu, dont l'entière responsabilité incombe à l'autrice et à l'auteur.

de gestion, ce collectif offre une diversité de perspectives et d'expériences essentielle pour appréhender les enjeux communicationnels. La personne n'ayant pas participé a estimé le sujet peu pertinent au regard de ses priorités professionnelles. En outre, la direction des systèmes d'information (DSI), située à un niveau organisationnel plus large et non intégrée au collectif, initialement favorable, s'est retirée en cours de projet, limitant le déploiement de certaines dimensions du dispositif et entrainant un recentrage sur des solutions plus locales, à l'échelle du collectif.

La démarche repose sur l'idée que la mise à profit des opportunités techniques ne peut être pertinente qu'en prenant en compte l'activité réelle, les préférences, les règles collectives et l'écologie de travail (Datchary, 2017). L'investigation et l'expérimentation se déploient ainsi sur un plan à la fois technique et socio-organisationnel.

1.2. Co-concevoir un dispositif sociotechnique

Sur le plan technique, l'objectif est de concevoir un système permettant aux participant·es d'agir sur l'infrastructure de communication, afin d'en adapter le fonctionnement à leurs besoins. Le contrôle de cette infrastructure, qui comprend à la fois les terminaux (ordinateurs, smartphones) et le réseau, doit être suffisamment flexible pour évoluer avec les pratiques, tout en restant accessible à des personnes non expertes. Dans un contexte de multiplication des leviers d'action sur les infrastructures de communication, nous faisons l'hypothèse que l'on dispose d'une infrastructure programmable, c'est-à-dire qu'elle intègre des technologies avancées telles que :

— le réseau défini par le logiciel, approche qui centralise et automatise la gestion des flux de communication via des contrôleurs logiciels,
— la virtualisation des fonctions réseau, qui autorise le déploiement dynamique de services réseau sous forme de logiciels, sans dépendre d'un matériel dédié,
— et des outils d'intermédiation permettant d'agir sur les communications technologiquement médiées, tels que des agents logiciels ou des extensions pour navigateurs et clients e-mail.

Concrètement, le système technique formalise les demandes relatives aux communications dans un format interprétable par machine. Ces règles peuvent par exemple consister à bloquer les e-mails après 18 h, rediriger les messages urgents vers un autre canal et différer certains messages pendant des visioconférences.

Afin de réaliser ces demandes, nous proposons une architecture modulaire : différents modules (contrôleur réseau, agents logiciels, services de messagerie) représentent les entités permettant d'agir sur les communications. À l'aide d'algorithmes issus du domaine de la planification automatique (Ghallab et al., 2025), le système sélectionne les actions les plus pertinentes en réponse à des demandes variées, tout en prenant en compte un environnement hétérogène et dynamique. À ce titre, le système présente des propriétés caractéristiques de l'intelligence artificielle, notamment la perception de son environnement, la prise de décision autonome et la planification d'actions adaptatives. La nature itérative du processus de raffinement et son caractère explicite facilitent l'interprétation du comportement du système, en cohérence avec l'objectif d'« automatisation sous contrôle » qui structure le projet.

Sur le plan socio-organisationnel, la recherche s'intéresse aux pratiques effectives des salarié·es et aux règles collectives et institutionnelles qui les encadrent. L'intervention consiste alors à construire des espaces de discussion autour de la communication au travail. Celles-ci visent à accompagner les salarié·es dans la régulation de leurs usages et, dans un second temps, à encadrer les actions du dispositif sociotechnique. Nous avons nommé ces espaces « ateliers de délibération collective » (ADC), que nous présentons dans la suite de cette section.

2. IDENTIFIER LES BESOINS ET ANCRER LA CONCEPTION DANS LE TRAVAIL RÉEL

Afin d'instaurer un cadre d'écoute propice à l'expression des expériences et tensions liées aux pratiques communicationnelles, nous avons organisé quatre ADC, articulés à des phases d'observation et d'entretiens. Ce dispositif a permis d'ancrer la réflexion dans le travail réel, et de faire émerger des problématisations directement issues des situations de travail (Figure 1).

Le premier atelier a abouti à l'identification des principales difficultés rencontrées : multiplication des canaux de communication, surcharge informationnelle, difficultés de priorisation des sollicitations et des tâches, ainsi que besoins accrus de coordination dans un contexte de généralisation du télétravail (Benedetto-Meyer, 2023). Une liste d'une vingtaine de problèmes a ensuite été élaborée puis soumise à un vote afin de dégager les priorités du collectif. Parmi les enjeux les plus votés figurent les dysfonctionnements techniques qui font que, comme l'explique une salariée, « *nous (les salarié·es) rentrons en réunion agacés* », l'absence de régulation collective de la charge de travail, ainsi que la multiplication des procédures et des logiciels. Cette liste a été discutée au sein de l'équipe de recherche afin d'identifier les problématiques communes pouvant être traitées dans le cadre du projet.

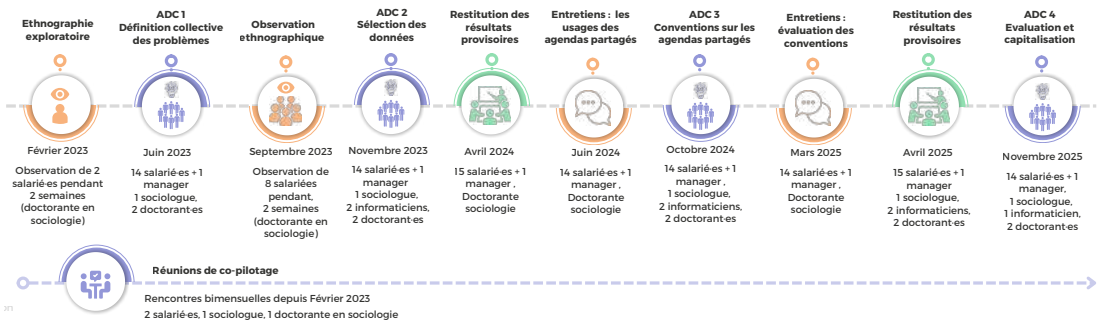


Figure 1 : Étapes du projet de recherche VERTUOSE

Dans une volonté de respecter la dimension démocratique de la démarche, entendue comme l'ouverture des choix techniques aux personnes concernées, le second atelier portait sur les conditions de partage des données et l'identification de fonctionnalités potentielles du système, à partir d'exercices de *design fiction*. Dix types de données partageables ont été définis, avec différents niveaux d'anonymisation selon les destinataires (Figure 2). Les discussions ont également mis en lumière le fait que les fonctionnalités recherchées ne visaient pas principalement la délégation ou l'automatisation de certaines tâches, mais plutôt des fonctionnalités d'assistance comme la visualisation, la priorisation et la mise en forme des informations.

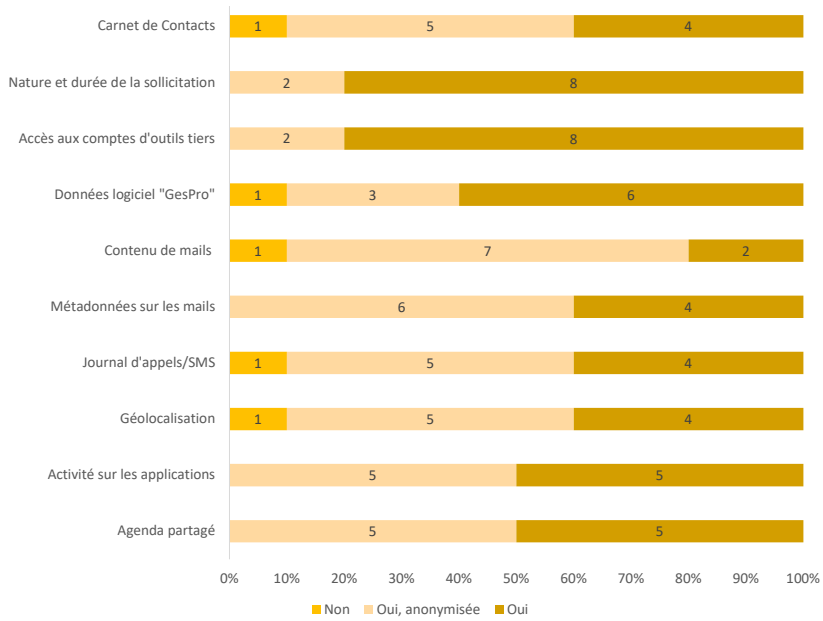


Figure 2 : Volontés de partage exprimées en fonction du type de données concernées

Le troisième atelier s'est concentré sur les usages des calendriers numériques partagés, outil central pour l'organisation du travail individuel et la coordination collective, mais également générateur de surcharge et de tensions. Il a conduit à l'élaboration de conventions d'usage visant à faciliter la compréhension mutuelle et la coordination, tout en préservant les pratiques individuelles existantes. Ces conventions ont alors été expérimentées et évaluées collectivement au regard de leur utilité, de leur impact sur la charge de travail et de leurs effets sur la coordination. Le quatrième atelier, enfin, été consacré à l'évaluation globale du projet et à la capitalisation de l'expérience, tant pour le collectif que pour de futures interventions en entreprise.

La démarche a impliqué, pour l'équipe de recherche, un relâchement du contrôle sur la trajectoire du projet et sur le processus de conception. Elle a supposé d'accepter des écarts par rapport à la trajectoire initialement prévue et de réaliser des ajustements.

Cela s'est traduit par un double déplacement de notre positionnement. D'une part, les participant·es ont été reconnu·es comme expert·es de leurs situations, et non comme de simples « utilisateurs finaux » d'un dispositif. Cette reconnaissance remet en cause une posture de surplomb du·de la chercheur·se supposé·e, de type « tout comprendre, tout expliquer et tout résoudre » (Bühlera et al., 2006). D'autre part, nous avons cherché à limiter les asymétries de pouvoir entre disciplines, en reconnaissant tant la sociologie que l'informatique comme de véritables partenaires dans le processus de conception et dans les prises de décision qui y sont associées. Les données ethnographiques et les accords issus des ateliers participatifs ont joué ici un rôle structurant, en révélant les écarts entre les hypothèses initiales et les besoins exprimés par les salarié·es, et offrant la possibilité d'ajuster le dispositif.

Alors que l'automatisation et la délégation du traitement des communications constituaient initialement le cœur du système, les besoins exprimés et les usages observés ont mis l'accent sur des fonctionnalités d'assistance. Compte tenu des contraintes temporelles, des limitations liées au terrain et des exigences disciplinaires des thèses, un nouveau compromis a été trouvé. La contribution technique s'est concentrée sur le développement d'un cadre de travail unifié autorisant l'expression de demandes variées au sein d'une architecture générique, incluant des fonctionnalités d'assistance au traitement des communications, telles que l'étiquetage des messages ou l'interopérabilité entre systèmes, comme la mise à jour du statut de disponibilité dans le calendrier en fonction des communications échangées.

Le dispositif sociotechnique proposé s'inscrit ainsi dans un travail participatif et interdisciplinaire, impliquant une articulation progressive des savoirs et des réorganisations partielles des cadres théoriques, dont la dynamique est loin d'être linéaire. Comme l'ont montré d'autres expériences (Deffontaines et Hubert, 2004 ; Lewkowicz, 2015 ; Louvel, 2015), ce type de démarche comporte de nombreux défis liés à l'articulation de disciplines aux orientations méthodologiques, théoriques et épistémologiques diverses.

3. LE TRAVAIL INTERDISCIPLINAIRE : LES DIVERGENCES ET LES APPRENTISSAGES

Afin d'illustrer les négociations permettant cette articulation, nous revenons, dans les sections suivantes, sur deux types de divergences disciplinaires rencontrées : les divergences de vocabulaire et les divergences de représentations, avant d'en dégager des conditions favorables au travail interdisciplinaire.

3.1. Divergences de vocabulaire

Le travail en co-présence a progressivement mis en évidence des écarts significatifs dans l'usage de certains termes. Afin de limiter les malentendus et de soutenir la coopération, nous avons élaboré un lexique partagé (Figure 3). Prenons l'exemple de la notion de « transparence ». Pour les sociologues, elle renvoie à la possibilité de rendre visibles et discutables les règles, les processus et les décisions qui structurent un système, afin d'ouvrir un espace de compréhension et de délibération. Pour les ingénieurs, cette notion qualifie la propriété d'un système dont l'utilisation par un individu ou un autre système n'expose pas la complexité interne. Cette divergence illustre la manière dont un même terme peut recouvrir des approches très différentes et révéler des enjeux centraux pour la conception et la gouvernance des systèmes sociotechniques : rendre visibles et compréhensibles les processus automatisés ou, au contraire, qualifier des mécanismes techniques opérant sans être perceptibles.

Terme	Informatique	Sociologie
Simuler / Tester / Emuler	**Simuler** : (1) Modéliser la structure et le comportement du système étudié, (2) Programmer ce comportement, (3) Exécuter **Tester** : Changer les paramètres de la simulation et la "jouer" **Emuler** : Faire fonctionner le système réel dans un environnement contrôlé	Simuler: changer un paramètre et voir ce qui se passe. Tester: faire fonctionner pour voir si cela marche. Emuler: ?
Monitoring / Collecte de données	Activités et outils permettant d'observer, de mesurer et d'analyser en continu l'état d'un système informatique (applications, serveurs, réseaux, bases de données, etc).	Monitoring: collecter données/ surveiller utiliser ces données pour contrôler le comportement de quelqu'un notamment d'un point de vue hiérarchique
Protocole	**Protocole** : Algorithme décrivant comment une activité doit se dérouler. **Protocole de communication** : algorithme suivi par les entités communicantes pour "dialoguer" **Protocole de Synchronisation** : algorithme suivi par des programmes pour réguler l'accès aux ressources partagées.	protocole = procédure dans un cas identifié
Collectif de travail		toute équipe de travail n'est pas un collectif
Transparence	La transparence d'un système -> les mécanismes internes /techniques sont **masqués** afin que l'utilisation soit simple, intuitive, indépendamment de la complexité sous-jacente.	La visibilité des règles, des décisions, des processus et des résultats. La lisibilité/ possibilité de comprendre pour les acteurs concernés

Figure 3 : Extrait du lexique commun construit par l'équipe de recherche
Lecture : vert pour définitions convergentes, rouge pour définitions divergentes, jaune pour divergences nécessitant une mise en accord sur une définition commune

Le dialogue et la compréhension entre chercheur·ses issu·es de champs disciplinaires éloignés sont reconnus comme difficiles (Jollivet et Legay, 2005), car la négociation d'un vocabulaire commun ne se réduit pas à un simple accord sémantique. Elle suppose une négociation continue entre visions épistémiques, disciplinaires et politiques hétérogènes, influençant à la fois les orientations du projet, les choix de conception et, *in fine*, les modalités d'appropriation ou de contestation du système par les salarié·es.

3.2. Divergences de représentations : les imaginaires disciplinaires

Au-delà des divergences terminologiques, les chercheur·ses mobilisent des représentations implicites de ce que font les autres disciplines et de leurs modes de travail. Ces « imaginaires disciplinaires » (Soubeyran, 1997) orientent les attentes réciproques, influençant la répartition des rôles et les formes de coopération.

Dans notre projet, la sociologie était perçue par l'équipe informatique comme la discipline capable de faciliter l'accès au terrain, en associant les participant·es à l'expérimentation et en offrant la possibilité d'identifier les besoins auxquels le système technique devrait répondre. Elle devait produire des analyses ethnographiques et un état de la littérature pour définir la « désirabilité » du système, ainsi qu'un modèle générique des communications professionnelles et des critères mesurables de la surcharge communicationnelle, transposables dans un dispositif technique. Autrement dit, l'informatique attendait de la sociologie qu'elle fournisse l'équivalent d'un cahier des charges.

Inversement, l'informatique était perçue par les sociologues comme orientée vers la production rapide d'un système opérationnel, intégrant des outils de captation de l'activité (« *monitoring* »), dont les données pourraient être confrontées aux données ethnographiques. Cette représentation reposait sur une vision de l'informatique comme discipline de production d'artefacts et de solutions rapides, sans distinction claire entre « ingénierie » et « recherche en informatique ». Ces décalages entre ce qui est attendu et ce que peut faire chacune des disciplines ont parfois engendré des postures défensives : tantôt la sociologie se protégeait du risque d'être réduite au rôle d'« auxiliaire » de la conception, tantôt l'informatique refusait d'endosser celui de « simple prestataire technique ».

Une hypothèse permettant d'éclairer ces tensions tient à la différence entre les logiques de totalisation propres à l'ethnographie et à la modélisation informatique. D'un côté, une logique inductive, attentive à la singularité des situations, pour qui il est difficile de fixer des conclusions rapidement. De l'autre, une logique hypothético-déductive orientée vers la modélisation et la production de connaissances généralisables. Ces écarts rejoignent les constats formulés dans les domaines de la Human-Computer Interaction (HCI) et du Computer-Supported Cooperative Work (CSCW), qui soulignent la difficulté d'articuler données ethnographiques et spécifications techniques, puisque le décalage entre ce qu'il serait souhaitable d'intégrer dans le système et ce qui est techniquement faisable demeure souvent difficile à combler (Hughes et al., 1992 ; Ackerman, 2000).

3.3. Défis et apprentissages de la collaboration interdisciplinaire

Comme nous venons de le montrer, l'interdisciplinarité ne se réduit pas à une addition de compétences. Elle suppose une négociation continue entre plusieurs logiques d'appréhension du réel. Si l'on considère que les conditions d'exercice du projet de conception font pleinement partie de l'activité de conception (Triclot, 2024), garantir des conditions matérielles favorables à la collaboration interdisciplinaire apparait indispensable. Ainsi, la mise en place d'espaces-temps de co-présence (des réunions mensuelles et des séminaires annuels en résidence) s'est révélée déterminante pour favoriser l'acculturation entre disciplines et instaurer la confiance au sein de l'équipe. Ces moments de travail

ont permis le développement d'une culture partagée et une véritable articulation de savoirs, facilitant la construction d'une vision commune du système.

Pour ce faire, nous avons dû réviser les attentes initiales et élaborer des compromis. Plutôt que de viser un prototype complet ou la mise en place d'un système de collecte de données automatisé en temps réel, l'équipe a privilégié une collecte rétrospective de données existantes. Concrètement, cela a consisté à récupérer et analyser un an d'historique de données issues des e-mails et des agendas. Cette réorientation met en évidence le fait que l'interdisciplinarité relève d'un processus de négociation continue des cadres, des temporalités et des objectifs, permettant de maintenir un horizon commun : concevoir des dispositifs ajustés à l'activité, aux besoins et aux valeurs des publics concernés.

Enfin, notre expérience insiste sur le rôle structurant des cadres institutionnels. Malgré les injonctions au décloisonnement des disciplines, les pratiques scientifiques restent encadrées par des normes disciplinaires d'évaluation et de validation. L'interdisciplinarité, en particulier au niveau doctoral, peut apparaitre comme un « exercice de style » (Bühlera et al., 2006) contraint par des logiques disciplinaires persistantes. Ces tensions se traduisent notamment par des temporalités de recherche distinctes et une valorisation inégale des apports interdisciplinaires. La coopération interdisciplinaire suppose alors de coordonner ces rythmes et de négocier des jalons communs, afin de rendre compatibles exigences scientifiques et contraintes du projet, malgré des logiques de production scientifique hétérogènes.

— CONCLUSION : PARTICIPATION ET DIALOGUE INTERDISCIPLINAIRE COMME LEVIERS POUR DES SIA ATTENTIFS AUX CONDITIONS DE TRAVAIL

La conception et le déploiement des systèmes d'intelligence artificielle mobilisent une pluralité d'acteurs, de savoirs et de visions du travail. Les modalités concrètes de coopération entre concepteur·trices, décideur·ses et utilisateur·rices, en particulier la place accordée au dialogue et au partage du pouvoir décisionnel, participent directement à la configuration des dispositifs. Loin d'être neutre, toute opération de conception engage en effet une composition de normes et de valeurs liées au travail.

Les résultats présentés montrent que l'interdisciplinarité ne peut être réduite à une simple juxtaposition de compétences : elle constitue un processus situé, fait de négociations, d'ajustements et de compromis entre cadres théoriques, méthodes et temporalités. Cette dynamique fait écho aux interactions observées dans les processus plus larges de conception des SIA, où la construction d'accords entre acteurs hétérogènes demeure un enjeu central. L'expérience analysée souligne l'importance d'instituer des espaces-temps de coopération, de reconnaitre la valeur des savoirs situés et « profanes », ainsi que de maintenir ouvertes les décisions le plus longtemps possible afin d'intégrer, de manière itérative, les enseignements et les besoins du terrain.

Dans un contexte de diffusion rapide des SIA dans les organisations, leur légitimité ne saurait reposer sur la seule performance technique. Elle dépend aussi de leur capacité à s'inscrire dans les pratiques professionnelles, à soutenir les collectifs de travail et à s'articuler aux dynamiques organisationnelles existantes. Les démarches participatives et interdisciplinaires apparaissent, à ce titre, comme des conditions essentielles pour concevoir des systèmes attentifs au travail réel, susceptibles de contribuer effectivement à l'amélioration des conditions de travail.

RÉFÉRENCES BIBLIOGRAPHIQUES

— Ackerman, M.S. (2000), « The Intellectual Challenge of CSCW: The Gap Between Social Requirements and Technical Feasibility », in *Human-Computer Interaction*, 15(2-3), 179-203. https://doi.org/10.1207/S15327051HCI1523_5

— Akrich, M. (1987), « Comment décrire les objets techniques ? », in *Techniques et culture*, (9), 49. https://shs.hal.science/halshs-00005830

— Benedetto-Meyer, M. (2023), « Que deviennent les collectifs en télétravail ? Une analyse par l'usage des outils collaboratifs », in *SociologieS*. https://doi.org/10.4000/sociologies.20615

— Bobillier Chaumon, M.-É. (2023), *Psychologie du travail digitalisé. Nouvelles formes de travail et clinique des usages*, Malakoff, Dunod, 208 p.

— Borel, S. (2024), « Étude des impacts de l'IA sur le travail. Rapport d'enquête LaborIA Explorer ». https://www.laboria.ai/wp-content/uploads/2024/05/Rapport-denquete-LaborIA-Explorer.pdf

— Bühlera, È.A., Cavaillé, F. et Gambino, M. (2006), « Le jeune chercheur et l'interdisciplinarité en sciences sociales. Des pratiques remises en question », in *Natures Sciences Sociétés*, 14(4), 392-398.

— Datchary, C. (2017), « Individu, collectif, organisation : les trois échelles de prise en charge de la multi-activité au travail », in Bidet, A., Datchary, C. et Gaglio, G. (dir.), *Quand travailler, c'est s'organiser. La multi-activité à l'ère numérique*, Paris, Presses des Mines, 222 p.

— Deffontaines, J.-P. et Hubert, B. (2004), « Dossier Interdisciplinarité. Un regard sur l'interdisciplinarité à l'Inra. Point de vue de deux chercheurs du département Sciences pour l'action et le développement (SAD) », in *Natures Sciences Sociétés*, 12(2), 186-190. https://doi.org/10.1051/nss:2004025

— Ghallab, M., Nau, D. et Traverso, P. (2025), *Acting, Planning, and Learning*, Cambridge, Cambridge University Press, 632 p.

— Hughes, J.A., Randall, D. et Shapiro, D. (1992), « From ethnographic record to system design: Some experiences from the field », in *Computer Supported Cooperative Work (CSCW)*, 1(3), 123-141. https://doi.org/10.1007/BF00752435

— Jollivet, M. et Legay, J.-M. (2005), « Dossier Interdisciplinarité. Canevas pour une réflexion sur une interdisciplinarité entre sciences de la nature et sciences sociales », in *Natures Sciences Sociétés*, 13(2), 184-188. https://doi.org/10.1051/nss:2005030

— Lewkowicz, M. (2015), « Vers une démarche interdisciplinaire d'ingénierie des pratiques collectives », in *Revue des sciences et technologies de l'information - Série ISI : Ingénierie des systèmes d'information*, 20(6), 121. https://doi.org/10.3166/isi.20.6.121-140

— Louvel, S. (2015), « Ce que l'interdisciplinarité fait aux disciplines. Une enquête sur la nanomédecine en France et en Californie », in *Revue française de sociologie*, 56(1), 75-103. https://doi.org/10.3917/rfs.561.0075

— Soubeyran, O. (1997), *Imaginaire, science et discipline*, Paris, L'Harmattan, 480 p. https://shs.hal.science/halshs-00382793

— Théviot, A. (2023), *Gouverner par les données ? Pour une sociologie politique du numérique*. Lyon, ENS éditions, 288 p.

— Triclot, M. (dir., avec le groupement d'intérêt scientifique Unité des technologies et des sciences de l'homme)(2024), *Prendre soin des milieux. Manuel de conception technologique*, Paris, Éditions Matériologiques, 458 p.

— Zouinar, M. (2020), « Évolutions de l'Intelligence Artificielle : quels enjeux pour l'activité humaine et la relation Humain-Machine au travail ? », in *Activités*, (17-1), Art. 17-1. https://doi.org/10.4000/activites.4941

POURQUOI LE DIALOGUE SOCIAL TECHNOLOGIQUE NE PEUT-IL PAS FAIRE DE MIRACLES À L'HEURE DES IAS ?

RETOUR SUR UNE SÉRIE DE PROJETS PORTÉS, SOUTENUS OU ACCOMPAGNÉS PAR L'ANACT

Vincent Mandinaud, Ludovic Bugand, Nicolas Vispi, Grégory Plançon
(Anact)

Face au développement de l'intelligence artificielle, le dialogue social est souvent présenté comme un levier essentiel pour accompagner les transformations du travail liées à son déploiement. À partir d'une revue de littérature et de plusieurs projets portés, soutenus ou accompagnés par l'Anact, cet article interroge les conditions et les limites de cette ambition. Il montre que les systèmes d'IA relèvent d'agencements sociotechniques qui, sous l'emprise de la « religion industrielle », tendent à invisibiliser le travail, à naturaliser les choix techniques et à restreindre les espaces de délibération. Les expérimentations analysées soulignent toutefois qu'en repartant des situations de travail et du travail réel, le dialogue social technologique peut contribuer à rendre ces technologies discutables, à développer une exploration collective des transformations possibles et à rouvrir des marges d'action. L'article conclut cependant que cette démocratisation des choix technologiques ne peut reposer sur le seul dialogue social, mais suppose des institutions capables d'organiser durablement la discussion entre expertise technique, expérience du travail et représentation des intérêts, afin de rendre les trajectoires technologiques véritablement discutables.

Mots-clés : intelligence artificielle, dialogue social, travail, démocratie technique

« Le dégoût sous toutes ses formes est une des misères les plus précieuses qui soient données à l'homme comme échelle pour monter. »
Simone Weil, *La pesanteur et la grâce*, 1947

INTRODUCTION

Comment expliquer que l'intelligence artificielle (IA) apparaisse simultanément comme une technologie incontournable, un objet difficilement discutable et une question pourtant renvoyée au dialogue social ? Notre hypothèse est que cette contradiction résulte d'un verrouillage socio-technique propre à la religion industrielle, qui tend à sacraliser la technique, à subordonner le travail et à assigner au dialogue social une fonction d'accompagnement de choix déjà largement arrêtés. Dès lors, la question n'est pas seulement de savoir comment introduire les IA dans les organisations, mais comment mettre en discussion les choix techniques, politiques et organisationnels qu'elles incorporent. À partir de la littérature académique et de plusieurs projets portés, soutenus ou accompagnés par l'Anact, cet article montre d'abord comment la religion industrielle tend à rendre les technologies indiscutables. Il examine ensuite comment le retour au travail réel peut constituer une première opération de désacralisation des IA. Enfin, il interroge les conditions dans lesquelles le dialogue social peut devenir une véritable intervention démocratique sur les choix technologiques, tout en montrant pourquoi il ne peut, à lui seul, faire de miracles.

1. LE CULTE DE L'IA AU TRAVAIL : UN AGENCEMENT SOCIO-TECHNIQUE INDISCUTABLE ?

L'intelligence artificielle est souvent présentée comme une rupture historique. Les discours qui l'accompagnent promettent une transformation profonde de la production, des services, de la décision, voire de l'intelligence humaine. Pourtant, ces promesses prolongent des imaginaires anciens. Bien avant l'IA, la machine à vapeur, l'électricité, l'automatisation ou Internet furent investis d'une mission quasi salvatrice : résoudre les contradictions économiques et sociales par le progrès technique. L'intelligence artificielle ne constitue donc pas une parenthèse de l'histoire industrielle ; elle en représente l'expression contemporaine.

Autrement dit, l'IA apparait comme l'épigone de la religion industrielle (Musso, 2017). Elle ne se manifeste pas seulement par de nouveaux produits et infrastructures, mais aussi comme une cosmologie politique, une superstructure, attribuant à la technologie une capacité quasi providentielle à organiser, voire gouverner, l'économie et la société. Les technologies deviennent les médiatrices privilégiées du progrès, tandis que l'innovation tend à se confondre avec le bien commun. L'IA hérite pleinement de cette croyance : elle concentre désormais les promesses de croissance, d'efficacité, de réindustrialisation et de souveraineté.

Cette religiosité produit un premier effet politique : elle transforme progressivement les technologies en évidences. Plus une innovation est présentée comme inévitable, moins elle parait discutable. Les débats portent alors moins sur les finalités que sur les modalités de déploiement et d'accompagnement du changement. Le progrès cesse d'être un objet de délibération pour devenir une trajectoire à laquelle il faudrait s'adapter. Le verrouillage socio-technique commence lorsque l'avenir prend la forme d'une nécessité plutôt que d'une possibilité.

Cette illusion est pourtant largement déconstruite par les travaux contemporains en sociologie des sciences et des techniques. Comme le rappellent Benbouzid et ses collègues (2022), l'IA ne désigne pas une réalité homogène mais une pluralité de dispositifs, de pratiques et de situations articulant des dimensions culturelles, disciplinaires, socio-politiques, industrielles et juridiques. Cardon et ses collègues (2018) montrent que ces systèmes ne peuvent être compris qu'en reconstituant les chaines de médiations techniques, organisationnelles et sociales qui les composent. L'IA n'est jamais un objet isolé ; elle est toujours le produit d'un assemblage, y compris entrepreneurial et financier (Boullier, 2019).

Les systèmes d'IA ne sont donc pas de simples logiciels alimentés par des données. Ils constituent, pour reprendre Michel Serres (1982), des quasi-objets qui redistribuent simultanément les relations entre humains, institutions et techniques. Gourlet (2024) prolonge cette idée en montrant que ces systèmes organisent nos rapports au monde à travers des infrastructures de données, des modèles computationnels et des procédures de calcul qui orientent progressivement nos perceptions, nos décisions et nos coordinations. Leur pouvoir tient moins à leur intelligence supposée qu'à leur capacité à reconfigurer les conditions de l'action collective.

L'un des paradoxes majeurs de cette religiosité réside dans l'effacement du travail qui rend ces technologies possibles. L'autonomie apparente des systèmes masque une immense infrastructure humaine. Casilli (2019) a montré que derrière les performances de l'IA se déploie une division internationale du travail composée d'annotateurs de données, de modérateurs de contenus, de travailleurs des plateformes, d'ingénieurs, de mainteneurs ou de logisticiens. On pourrait y ajouter les activités extractives et industrielles qui soutiennent ces infrastructures. Ce travail est moins invisible qu'invisibilisé. Plus la machine parait autonome, plus les activités humaines qui la rendent possible disparaissent du récit technologique.

Cette invisibilisation participe du fonctionnement même de la religion industrielle. Celle-ci attribue à la technique des qualités qui proviennent en réalité d'organisations complexes, de savoir-faire professionnels et de coopérations multiples. Le miracle technologique naturalise ainsi les choix politiques et dématérialise le travail humain. L'IA apparait comme une intelligence autonome alors qu'elle constitue avant tout un vaste dispositif reliant connaissances, infrastructures, ressources matérielles, financières et travail humain.

Evgeny Morozov (2013) invite précisément à se méfier de cette naturalisation qu'il qualifie de solutionnisme technologique. Les problèmes sociaux sont alors reformulés comme des problèmes d'information ou d'optimisation susceptibles d'être résolus par de meilleurs calculs. Il ne s'agirait plus de transformer les organisations ou les rapports sociaux, mais d'améliorer les technologies. L'intelligence artificielle devient ainsi moins une technologie qu'une manière de penser le gouvernement des sociétés, des entreprises ou des administrations.

Cette évolution explique pourquoi les systèmes d'IA au travail ne sauraient être considérés comme spontanément dignes de confiance. Comme le rappelle Ferguson, Mandinaud et Ponce Del Castillo (2026)[1], toute technologie incorpore des rapports de pouvoir, des arbitrages et des conflits de valeurs. Les données ne sont jamais neutres ; les indicateurs privilégient certaines conceptions de la performance ; les modèles d'apprentissage traduisent des choix économiques, juridiques ou politiques. La confiance ne peut donc résulter de la seule performance technique. Elle suppose des institutions capables de rendre discutables les finalités qui président à la conception, au fonctionnement, au déploiement et aux usages de ces systèmes.

C'est précisément à ce point que le dialogue social technologique prend tout son sens. Son enjeu n'est pas d'accompagner une innovation présentée comme inévitable, mais de rouvrir des trajectoires que la religion industrielle tend à considérer comme déjà écrites. Il ne s'agit pas seulement de discuter les usages de l'intelligence artificielle, mais les choix techniques eux-mêmes, qui sont toujours des choix politiques. Encore faut-il que les acteurs disposent d'espaces et de temps leur permettant de revenir à ce que ces technologies transforment réellement : le travail et le pouvoir. C'est précisément ce que les expérimentations conduites, accompagnées ou soutenues par l'Anact se sont efforcées d'explorer.

1 Voir l'article dans ce dossier

2. DÉSACRALISER LES IAS PAR L'EXPÉRIENCE DU TRAVAIL : JOUER POUR APPRENDRE À DÉJOUER LES MACHINES, LES MACHINERIES ET LES MACHINATIONS

Si la religion industrielle tend à sacraliser la technique, elle transforme simultanément le travail en simple variable d'ajustement des innovations. Les promesses contemporaines de l'IA mettent l'accent sur la productivité, la compétitivité ou les gains d'efficacité, mais beaucoup plus rarement sur le travail comme expérience vécue. Les travailleurs apparaissent surtout comme des utilisateurs ou des bénéficiaires. L'activité réelle, ses compromis, ses savoir-faire et ses conflits s'effacent derrière les performances attendues des dispositifs techniques. Le travail cesse alors d'être un point de départ pour devenir une conséquence des choix technologiques.

Cette inversion est précisément ce que Simone Weil, dans *La pesanteur et la grâce* (1947), dénonçait déjà au cœur de la civilisation industrielle. La mystique du travail qu'elle appelle de ses vœux ne consiste pas à glorifier le travail, mais à critiquer une organisation industrielle qui prive les individus de toute prise sur leur activité. Pour la philosophe, le travail possède une dimension politique et spirituelle parce qu'il met les êtres humains en relation avec le réel, avec les autres et avec eux-mêmes. Et peut-être plus encore lorsqu'ils font l'expérience du dégout du travail, sans pourtant y céder. Lorsque cette expérience disparait derrière les procédures, les prescriptions ou les automatismes, ce n'est pas seulement la productivité qui est en jeu, mais la possibilité même d'habiter son travail.

Cette intuition éclaire les débats contemporains sur l'IA. Le solutionnisme technologique tend à considérer les situations de travail comme des problèmes de calcul à optimiser, comme si la complexité des arbitrages quotidiens, des coopérations ou des dilemmes professionnels pouvait être intégralement traduite en variables et en données. Ce qui échappe à cette formalisation tend alors à perdre de sa légitimité. Plus les systèmes numériques gagnent en sophistication, plus le risque grandit de confondre ce qui est calculable avec ce qui importe réellement aux yeux de celles et ceux qui travaillent.

L'intérêt du jeu sérieux développé par l'Institut syndical européen[2], puis adapté par l'Anact (*L'IA au travail, discutons-en !*), réside précisément dans sa capacité à renverser cette logique. Contrairement à de nombreuses formations consacrées à l'IA, il ne cherche pas d'abord à expliquer le fonctionnement des modèles génératifs ou des algorithmes. Son ambition est de réapprendre à parler du travail avant de parler de la technologie, ou plutôt de parler des technologies d'IA à partir du travail. Cette inversion constitue sa principale originalité. Il ne demande pas aux participants ce qu'ils pensent de l'IA, mais ce que l'IA fait ou pourrait faire au travail, et ce que le travail pourrait faire à l'IA.

Le dispositif repose sur une idée simple : les controverses deviennent fécondes lorsqu'elles cessent de porter sur les performances supposées des technologies pour revenir aux situations concrètes dans lesquelles elles s'inscrivent. Les participants ne discutent plus d'une IA abstraite, mais de décisions de recrutement, de diagnostics médicaux, de gestion des ressources humaines, de maintenance industrielle ou de relation de service médiés par des systèmes d'IA. L'objet du débat n'est plus l'algorithme lui-même, mais l'activité qu'il transforme, les conditions de cette transformation et ses conséquences.

Les échanges montrent rapidement que les effets d'un même système d'IA varient selon les métiers, les organisations, les collectifs de travail et les modalités de déploiement. Une technologie susceptible d'alléger certaines tâches administratives peut simultanément réduire les possibilités d'apprentissage des nouveaux arrivants. Un outil d'aide à la décision peut renforcer la coopération entre professionnels ou, au contraire, appauvrir leurs capacités d'interprétation. Les catégories de « bénéfices » ou

2 https://www.etui.org/ai-game

de « risques » apparaissent alors insuffisantes. Les effets des IA résident moins dans les technologies elles-mêmes que dans les configurations socio-techniques auxquelles elles participent.

Le jeu permet ainsi de rendre perceptible ce que Ponce Del Castillo (2023) désigne comme « les multiples visages d'une technologie dénuée de visage ». Derrière l'expression « intelligence artificielle » apparaissent une diversité de situations, de rapports de pouvoir, d'intérêts et de compromis, au fil des rôles endossés par les participants : dirigeants, salariés, représentants du personnel, fournisseurs de technologies ou autorités publiques. À travers le jeu, les participants découvrent autant une pluralité technologique qu'une pluralité d'organisations possibles du travail. Le jeu produit ainsi moins une littératie numérique *stricto sensu* qu'une capacité d'enquête sur le façonnage croisé des mondes sociaux, des technologies et de leurs usages.

Cette démarche rejoint plusieurs expérimentations soutenues par l'Anact autour du dialogue social technologique. Qu'il s'agisse du projet DIAL-IA[3], des méthodes inspirées de Simondon comme la méthode de l'objet technique[4], des démarches ISIDOR ou APIDOR[5], ou encore de l'outil DIAG[6], toutes procèdent d'un même mouvement : déplacer la discussion des objets techniques vers les activités et les organisations qu'ils transforment, afin de reprendre du pouvoir d'agir (Clot, 2008). L'enjeu n'est pas seulement de comprendre la machine, mais de rendre visibles les médiations reliant les infrastructures numériques aux métiers, aux compétences, à la santé et à la qualité du travail. Les technologies cessent alors d'apparaitre comme des objets autonomes pour redevenir des composantes d'un écosystème socio-technique où les enjeux du travail retrouvent toute leur place.

L'un des principaux enseignements tirés de ces expérimentations tient au rôle des médiations proposées par la démarche ludo-pédagogique. Les cartes, les rôles, les scénarios, les défis et les consignes constituent autant de points d'appui permettant aux participants de confronter des rationalités différentes, de construire progressivement un langage commun et de réinvestir le travail comme objet de discussion. Le dialogue ne repose plus sur un échange d'opinions à propos des technologies, mais sur une exploration collective du travail réel et des transformations possibles. Il apparait ainsi moins comme une procédure formelle de concertation que comme un dispositif de traduction entre des univers professionnels et des logiques institutionnelles qui communiquent difficilement.

Ces expériences montrent également que le dialogue produit d'abord des transformations cognitives. Les représentations initiales, souvent polarisées entre enthousiasme technophile et inquiétude technophobe, évoluent vers une compréhension plus située des problèmes. Les participants découvrent que les controverses portent autant sur les finalités poursuivies que sur les capacités des systèmes. Ce déplacement réintroduit de la contingence là où dominait un sentiment d'inévitabilité et rappelle que les choix techniques demeurent des choix collectifs.

Ces capacités de dialogue ne doivent toutefois pas être surestimées. Si ces démarches permettent de reconstruire une intelligence collective du travail, elles se déploient dans un espace sécurisé, sur des bases fictionnelles et avec des participants volontaires. Elles rendent les technologies plus discutables, mais rarement négociables dans leurs orientations fondamentales en réalité. Car dans les entreprises, ces discussions restent peu nombreuses et interviennent fréquemment lorsque les trajectoires industrielles sont déjà largement engagées.

3 https://dial-ia.fr/
4 https://www.octares.com/accueil/307-la-methode-de-lobjet-technique-reprendre-la-main-sur-lactivite.html
5 https://methodes-dialogue-social-numerique.eu/
6 https://sites.google.com/matrice.io/dispositif-diag/accueil

C'est ici qu'apparait la limite structurelle du dialogue social technologique tel qu'il est mis en scène dans le jeu : il ouvre des espaces de délibération dans un univers où les principales décisions ont souvent été prises ailleurs et auparavant. Cette dissociation entre apprentissage démocratique et pouvoir effectif conduit finalement à interroger la portée politique du dialogue social face au développement actuel de l'IA dans les entreprises. À moins de considérer que les erreurs d'hier et d'aujourd'hui puissent encore nourrir les transformations de demain.

3. LE DIALOGUE SOCIAL TECHNOLOGIQUE : UNE INTERVENTION DÉMOCRATIQUE DANS LA RELIGION INDUSTRIELLE ?

Les expérimentations conduites autour du dialogue social technologique conduisent à un paradoxe. Elles montrent, d'une part, que les acteurs du travail sont capables de développer une lecture critique des systèmes d'IA lorsqu'ils disposent d'espaces de discussion, de médiations et de temps d'enquête ; et, d'autre part, que cette intelligence collective se heurte rapidement aux rapports de pouvoir qui structurent les trajectoires d'innovation. Le problème ne réside donc pas seulement dans un déficit de dialogue, mais aussi dans un verrouillage socio-technique qui cantonne la discussion à des orientations déjà largement stabilisées.

Ce verrouillage ne désigne pas uniquement la difficulté à modifier une technologie une fois déployée. Il renvoie à un ensemble de mécanismes économiques, institutionnels, cognitifs et symboliques qui rendent certaines trajectoires plus probables que d'autres. Investissements industriels, infrastructures numériques, plateformes, modèles économiques, stratégies de propriété intellectuelle, politiques publiques et imaginaires du progrès contribuent à refermer progressivement le champ des possibles. Lorsque les représentants des salariés sont consultés, les systèmes sont souvent déjà achetés, les architectures arrêtées, les données collectées et les calendriers de déploiement engagés[7]. Le dialogue porte alors davantage sur les modalités d'appropriation que sur les choix ayant conduit à ces dispositifs.

Ce constat rejoint les analyses de Winner (2022) sur le somnambulisme technologique. Les sociétés industrielles débattent volontiers des innovations, mais beaucoup plus rarement des conditions qui rendent certaines innovations pensables et d'autres impensables. Winner parle de « mythinformatique » pour désigner la croyance selon laquelle les technologies numériques posséderaient une rationalité propre dispensant d'interroger leurs finalités. La religion industrielle trouve ici une nouvelle expression : ce n'est plus Dieu qui écrit l'histoire, mais l'innovation.

Cette mythification explique en partie la fascination contemporaine pour l'IA. Celle-ci tire moins profit de ses seules performances que de l'autorité symbolique que lui confère sa sophistication. Plus les systèmes deviennent complexes et opaques, plus ils semblent relever d'un domaine réservé aux experts. La discussion démocratique tend alors à se réduire à des questions de conformité réglementaire, de cybersécurité ou de transparence algorithmique, laissant dans l'ombre les arbitrages économiques, sociaux et politiques qui président à leur conception. La boite noire n'est donc pas seulement algorithmique ; elle est aussi institutionnelle.

Or, comme le rappelle Feenberg (2004), les technologies ne constituent jamais des réalités closes. Elles demeurent susceptibles d'interventions démocratiques, c'est-à-dire de réorientations issues des usages, des conflits, des apprentissages ou des compromis construits par les acteurs concernés. Les

7 Rappelons cependant que sur ce point, le droit du travail et la jurisprudence sont clairs : les IRP doivent être informées/consultées avant toute forme de déploiement, faute de quoi le projet est suspendu, et l'entreprise sous astreinte financière jusqu'à la consultation.

expérimentations de dialogue social technologique illustrent cette possibilité. Elles montrent que les travailleurs n'interrogent pas seulement les performances techniques des systèmes, mais aussi leurs effets sur les métiers, les responsabilités, les collectifs, la santé, les compétences et la qualité du travail. En remettant l'activité réelle au centre de la discussion, ils rouvrent partiellement des choix que le discours technologique présente comme déjà tranchés.

Cette capacité d'intervention demeure toutefois profondément asymétrique. Les ressources consacrées au développement de l'IA se comptent aujourd'hui en centaines de milliards d'euros. À l'inverse, les moyens investis dans la démocratie technique restent marginaux. En France, les expérimentations recensées par l'Anact témoignent d'une réelle créativité méthodologique, mais reposent sur des équipes limitées, des financements ponctuels et des initiatives dispersées. L'asymétrie ne se limite donc pas aux rapports de pouvoir dans l'entreprise : elle tient également aux investissements consentis en faveur de technologies autoritaires ou de technologies démocratiques, selon la distinction proposée par Mumford (2021). Les premières bénéficient d'une puissance industrielle considérable ; les secondes reposent encore largement sur des expérimentations locales et éparses.

Il serait dès lors illusoire d'attendre du dialogue social en entreprise qu'il compense, à lui seul, les déséquilibres produits par l'économie politique contemporaine de l'IA. Bien qu'il puisse y contribuer, il ne peut remplacer une politique industrielle, de recherche, de formation ou du travail. Il ne saurait davantage corriger les effets de la concentration des infrastructures numériques, de la dépendance aux grandes plateformes ou de la division internationale du travail qui sous-tend le développement des systèmes d'IA. Lui attribuer cette mission reviendrait à lui conférer une fonction quasi miraculeuse.

Cette limite ne constitue pourtant pas un échec. Elle invite à reconsidérer ce que l'on entend par démocratisation des technologies. Trop souvent, celle-ci est assimilée à la diffusion massive des outils. Or, des millions de salariés peuvent utiliser quotidiennement des systèmes d'intelligence artificielle sans jamais participer aux décisions qui président à leur conception, à leur paramétrage ou à leur déploiement. La massification des usages ne saurait donc tenir lieu de démocratie technique.

L'enjeu est ailleurs. Les expérimentations analysées montrent que la démocratisation suppose le développement conjoint d'une culture technique, d'une culture du travail et d'une culture du dialogue. La première permet de comprendre les possibilités et les limites des systèmes d'IA. La seconde rappelle que les technologies n'ont de sens qu'au regard des activités humaines et des organisations qu'elles soutiennent ou transforment. La troisième crée les institutions capables de confronter durablement savoirs experts, savoirs professionnels, intérêts et valeurs. Ces trois dimensions sont indissociables. Sans culture technique, le débat démocratique reste prisonnier des mythes ; sans culture du travail, il perd son ancrage dans le réel ; sans culture du dialogue, il demeure sans prise sur les décisions. En l'absence de leur association, les systèmes d'IA au travail ne sauraient être des vecteurs de progrès économique et social.

C'est bien ce qu'avaient compris les partenaires sociaux européens en négociant l'accord-cadre intersectoriel de 2020 sur la transformation numérique au travail. Ils ont cherché à promouvoir une méthodologie de conception conjointe de cette transformation afin de favoriser la saisie des opportunités tout en réduisant les risques au plus près des réalités de terrain. Malgré des initiatives comme DIAL-IA, mises en valeur lors du Sommet mondial pour l'action sur l'intelligence artificielle de Paris en 2025[8], la traduction de ces orientations ne s'est pas (encore ?) actualisée en France à l'échelle interprofessionnelle. C'est pourtant en articulant de façon cohérente l'action à différentes échelles qu'il est, en réalité, possible de progresser.

8 https://www.youtube.com/watch?v=RSW8wgFaRcg

CONCLUSION

Cette perspective conduit à revenir sur ce que désigne réellement l'expression « intelligence artificielle ». Le premier enseignement des travaux en sociologie des sciences et des techniques est qu'il n'existe pas une intelligence artificielle, mais des intelligences artificielles, incarnées dans des systèmes, des milieux et des trajectoires variés (Boullier, 2025). Leur unité tient moins à une réalité technique homogène qu'au récit qui les accompagne. Parler d'IA au singulier masque cette diversité et entretient une confusion qui complique leur compréhension et leur discussion.

Le deuxième enseignement est que ces systèmes ne constituent pas des objets radicalement nouveaux. Ils prolongent des dynamiques anciennes de calcul, d'automatisation, de rationalisation, de standardisation, d'individualisation du travail et de gouvernance par les nombres (Supiot, 2015). Les questions qu'ils soulèvent – répartition du pouvoir, invisibilisation du travail, confiance, contrôle, autonomie professionnelle, justice, responsabilité ou démocratie technique – traversent l'histoire industrielle. Les IA ne créent pas ces problèmes ; elles les reconfigurent à une échelle et avec une intensité nouvelles.

Le troisième enseignement est peut-être le plus important. Les IA ne sont ni des objets uniques ni des réalités autonomes. Elles désignent une diversité de modèles, de données, d'infrastructures, de chaines logistiques, de ressources matérielles et financières, de travailleurs invisibilisés, d'organisations et de situations d'usage. Elles sont toujours des assemblages socio-techniques. Les gouverner ne consiste donc pas seulement à réguler des algorithmes, mais aussi à transformer les rapports sociaux qui les rendent possibles et qu'ils contribuent à reconfigurer.

Le déverrouillage socio-technique, loin de constituer uniquement un horizon, doit ouvrir un champ d'expérimentation. L'enjeu n'est ni de célébrer l'intelligence artificielle, ni de la rejeter, ni de sanctifier le travail face au progrès technologique. Il est de redonner aux collectifs de travail la capacité d'instituer les transformations socio-techniques à partir de leur expérience du travail, plutôt que de les subir comme des fatalités. Si le dialogue social technologique ne peut pas faire de miracles, il peut toutefois contribuer à l'essentiel : repartir du travail et rappeler que les technologies d'IA demeurent des constructions sociales susceptibles d'être discutées, orientées, agencées, façonnées et gouvernées afin d'assurer une prévention des risques plus ajustée et une répartition plus équilibrée des gains qui leur sont associés, à condition que les institutions démocratiques s'en donnent réellement les moyens (Acemoglu et Johnson, 2024).

BIBLIOGRAPHIE

— Acemoglu, D. et Johnson, S. (2024), *Pouvoir et progrès. Technologie et prospérité, notre combat millénaire*, Paris, Pearson Éditions, 632 p.
— Benbouzid, B., Meneceur, Y. et Smuha, N. (2022), « Quatre nuances de régulation de l'intelligence artificielle. Une cartographie des conflits de définition », in *Réseaux*, n° 232-233, p. 29-64.
— Boullier. D. (2019), *Sociologie du numérique*, Malakoff, Armand Colin Éditions, 380 p.
— Boullier. D. (2025), *Déshumanités numériques*, Malakoff, Armand Colin Éditions, 224 p.
— Cardon, D., Cointet, J.-P. et Mazières, A. (2018), « La revanche des neurones : l'invention des machines inductives et la controverse de l'intelligence artificielle », in *Réseaux*, n° 211, p. 173-220.
— Casilli, A. (2019), *En attendant les robots. Enquête sur le travail du clic*, Paris, Éditions du Seuil, 400 p.
— Clot. Y. (2008), *Travail et pouvoir d'agir*, Paris, Presses universitaires de France, 324 p.
— Feenberg, A. (2004), *(Re)penser la technique. Vers une technologie démocratique*, La Découverte.

— Ferguson, Y., Mandinaud, V., Ponce Del Castillo, A., (2026), « L'IA générative au travail : quelles transformations à l'horizon ? », La Revue des Conditions de travail, numéro 18, Anact.
— Gourlet, P., Ricci, D., Crépel, M., Meunier, A. et Barret-Bertelloni, M. (2023), « Prises, méprises, emprises, déprises, reprises… Les formes de participation au développement de l'IA en France », Web-publication, Médialab, Sciences Po Paris.
— Gourlet, P., Ricci, D. et Crépel, M. (2024), « Reclaiming artificial intelligence accounts », in Big Data & Society, 2024, 11 (2), https://hal.science/hal-04584823v1
— Morozov, E. (2013), *Pour tout résoudre cliquez ici. L'aberration du solutionnisme technologique*, Limoges, FYP Éditions, 358 p.
— Mumford, L. (2021), *Technique autoritaire et technique démocratique*, Lodève, La Lenteur Éditions, 112 p.
— Musso, P. (2017), *La Religion industrielle. Monastère, manufacture, usine. Une généalogie de l'entreprise*, Fayard.
— Ponce Del Castillo, A. (2023), *IA : les multiples visages d'une technologie sans visage. Outil pour faciliter la cartographie de l'IA, renforcer la pensée critique et soutenir les négociations relatives au déploiement de systèmes d'IA*, Bruxelles, ETUI Éditions, 43 p.
— Serres, M. (1982), *Genèse*, Paris, Éditions Grasset, 228 p.
— Supiot, A. (2015), *La gouvernance par les nombres. Cours au Collège de France (2012-2014)*, Paris, Éditions Fayard, 512 p.
— Weil, S. 1947, *La pesanteur et la grâce*, Plon.
— Winner, L. (2022), *La baleine et le réacteur. À la recherche de limites au temps de la haute technologie*, Herblay, Éditions Libre, 332 p.

INTELLIGENCE ARTIFICIELLE ET PRÉVENTION DES RISQUES PSYCHOSOCIAUX : UNE DÉMARCHE RÉFLEXIVE DE NANTES MÉTROPOLE

Marie Bernard (Nantes Métropole)
Priscilla Capovilla (Ex Nantes Métrople)

Nantes Métropole encadre l'intégration de l'intelligence artificielle grâce à une boussole de l'IA qui évalue chaque cas d'usage de l'IA selon des critères éthiques, en lien avec les organisations syndicales. L'un de ces critères vise à protéger la santé et les conditions de travail de ses agents. La collectivité a instauré un observatoire sur cinq ans pour mieux objectiver dans quelle mesure l'IA peut être un levier d'efficacité professionnelle ou, à l'inverse, une source d'isolement et d'intensification des tâches. En impliquant des agents experts, un doctorant CIFRE et les partenaires sociaux, l'organisation cherche à anticiper les mutations des métiers sur le long terme.

Mots-clés : prévention des RPS, intelligence artificielle, anticipation, dialogue social

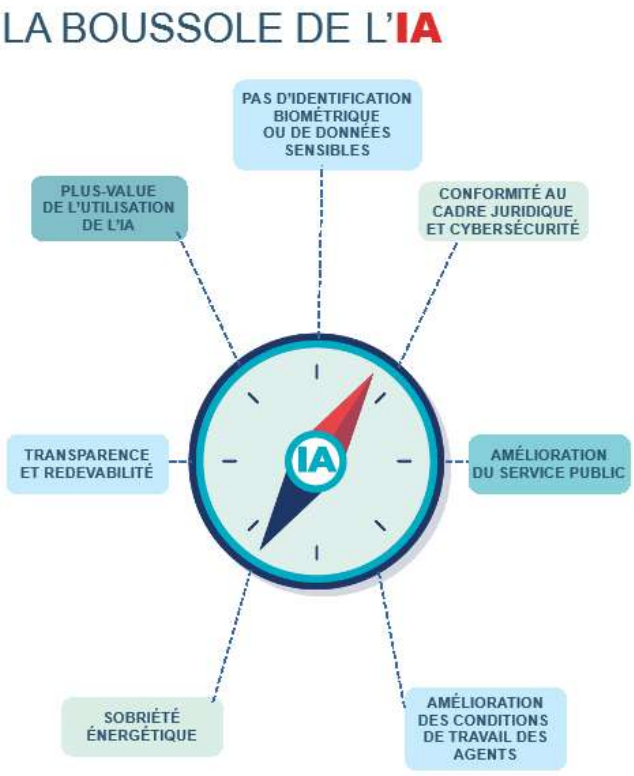
À Nantes Métropole, un observatoire des impacts de l'intelligence artificielle sur les conditions de travail a été mis en place en novembre 2024. Cet observatoire mesure dans le temps de quelle manière l'intelligence artificielle influence à la fois le travail, l'environnement de travail, et l'agente ou l'agent dans ses pratiques et ses interactions professionnelles. S'appuyant sur une démarche de recherche à long terme et dans un dialogue régulier avec les partenaires sociaux, plusieurs experts internes se sont prêtés à un exercice d'anticipation visant à identifier les impacts potentiels de l'IA sur les risques psychosociaux. L'objectif de cet article est de retracer les étapes et contenus de cette démarche.

1. LA BOUSSOLE DE L'IA, UN OUTIL POUR RÉGULER L'UTILISATION DE L'INTELLIGENCE ARTIFICIELLE

En mars 2024, Nantes Métropole a publié une doctrine sur l'intelligence artificielle, travaillée avec un collectif d'élus. Cette doctrine s'inscrit dans la continuité de l'action volontariste de la collectivité en matière de données (Charte métropolitaine de la donnée et de l'IA, parue en 2019 et mise à jour en 2025) et de numérique responsable (Stratégie numérique responsable, parue en 2024). Elle est un cadre d'exigences et de confiance pour le recours à l'IA au service des politiques publiques de la collectivité et pour les acteurs agissant pour son compte.

Cette doctrine est mise en pratique à travers un outil : la boussole de l'IA, reposant sur sept critères d'évaluation.

1. Absence d'identification biométrique et de collecte de données sensibles qui pourraient caractériser un ou des individus susceptibles de générer des biais discriminatoires (notamment caractérisation physique, comportementale, vestimentaire)
2. Conformité au cadre juridique et à la politique de cybersécurité
3. Contribution à l'amélioration du service public
4. **Contribution à l'amélioration des conditions de travail des agents publics**
5. Respect des engagements en matière de sobriété énergétique
6. Exigence de transparence et redevabilité
7. Évaluation du bénéfice par rapport à une solution alternative sans intelligence artificielle



La collectivité s'engage à développer uniquement des usages de l'IA qui s'appuient sur des bases juridiques existantes. Par ailleurs, elle s'interdit, et interdit aux acteurs publics et privés agissant pour son compte, toute utilisation d'algorithmes pour des prises de décision individuelles entièrement automatisées.

L'évaluation, outillée de la boussole de l'IA, est réalisée collectivement par un groupe-projet composé des services supports et services utilisateurs. Elle fait ensuite l'objet d'une validation politique et d'un partage aux organisations syndicales. Cet outil est nécessairement appelé à évoluer dans le temps, en fonction des pratiques et avancées technologiques. Il a vocation à se décliner dans les marchés publics lors de l'achat ou du renouvellement d'outils numériques.

— 2. CONTRIBUER À L'AMÉLIORATION DES CONDITIONS DE TRAVAIL, UN ESSENTIEL DANS LE CHOIX DES IA DÉPLOYÉES AU SEIN DE LA COLLECTIVITÉ

Avant l'introduction d'un outil embarquant de l'IA, la collectivité se pose plusieurs questions dans le but d'en évaluer l'opportunité du point de vue des conditions de travail (efficacité et qualité du travail, amélioration du service rendu, diminution de l'usure professionnelle et des tâches pénibles, etc.). Ainsi, il est vérifié que les besoins des agentes et agents ont été recueillis et pris en compte dans le processus de conception, d'acquisition et de déploiement du nouvel outil. Cette vérification permet de s'assurer que l'outil est bien au service des agentes et agents utilisateurs, qu'elles et ils ont conscience d'avoir recours à de l'IA et en comprennent *a minima* le fonctionnement pour éviter les biais éventuels.

La collectivité s'assure également de l'information et de la mise en discussion avec les représentants du personnel des enjeux en matière de conditions de travail en amont de l'introduction du nouvel outil.

Toujours dans cette phase amont, un certain nombre de questions sont posées afin de prendre en compte concrètement les enjeux en matière de risques psychosociaux liés à la mise en place de l'outil :

— Le projet a-t-il une action favorable sur la régulation de la charge de travail ? Permet-il de diminuer le nombre de tâches ou d'interruptions ?
— Le projet permet-il aux agentes et agents de maintenir ou de développer leur expertise et leurs compétences ? L'outil autorise-t-il une diversification des tâches ou le développement de nouvelles compétences ?
— Le projet favorise-t-il les relations interpersonnelles ? Permet-il de favoriser les interactions ou de limiter l'isolement ?
— Le projet est-il source d'autonomie et de responsabilisation pour les agentes et agents ? Sont-ils maitres du processus de contrôle et de décision ?
— *In fine*, le projet permet-il de préserver ou d'améliorer la santé des agents ? A-t-il une action favorable sur l'usure professionnelle, physique ou psychique ?

Plusieurs outils sont aujourd'hui mis en œuvre dans la collectivité, à la suite de leur évaluation réalisée avec la boussole de l'IA. Par exemple, des capteurs de fréquentation sont installés dans un musée, remplaçant ainsi le comptage manuel opéré par les agentes et agents – une tâche répétitive et fastidieuse demandant une attention et une vigilance importantes. Ce comptage est essentiel en matière de sécurité incendie. L'utilisation d'un outil avec de l'IA offre aux agentes et agents de se concentrer sur d'autres tâches que celle du comptage, comme l'accueil et l'orientation des visiteurs. Un autre exemple d'outil mis en place concerne la mise en circulation d'un véhicule équipé d'une caméra intelligente permettant de repérer l'état des éléments de voirie (panneaux de signalisation, marquage au sol, corbeilles, candélabres, etc.). Ce véhicule sillonne progressivement l'ensemble du territoire de la métropole. Il aide à la décision stratégique relative à l'espace public, et guide l'action opérationnelle au plus près des besoins (organisation des tournées d'entretien et de réparation de ces éléments de voirie et rationalisation des déplacements des agentes et agents, anticipation des demandes d'interventions des usagers, etc.). Un dernier exemple d'outil est l'utilisation d'une IA générative autorisant la recherche, la

rédaction et la synthèse de délibérations du conseil de la Métropole. Cet outil a pour objectif de faire gagner du temps aux agentes et agents rédacteurs et de faciliter la tâche de celles et ceux chargés de la relecture et du contrôle des délibérations.

Il est difficile de prévoir à quelle échéance pourront être observés les effets concrets, attendus ou imprévus, de tels outils requérant de l'IA, sur la réalisation du travail, sur les compétences des agentes et agents, sur les relations entre collègues ou encore sur l'organisation du travail. La collectivité prend donc le parti d'une démarche sur le long terme et se donne cinq ans pour mesurer les risques et les opportunités de l'IA à l'égard des conditions de travail de ses agentes et agents et de l'évolution de leurs métiers.

Ce travail d'observation et de mesure de l'utilisation réelle de l'IA repose sur des outils tels que des questionnaires, de l'observation ou encore des focus groups, et associe différents experts au sein de la collectivité : spécialistes des questions de santé et de sécurité au travail (médecin du travail, psychosociologue, préventeur), des usages numériques, de l'emploi et de la formation, du management et de l'organisation, de la communication interne, etc. Il associe également le monde de la recherche, à travers le recrutement d'un doctorant en contrat CIFRE (convention industrielle de formation par la recherche) de 2026 à 2028. Il s'agit de garantir la validité scientifique des observations et conclusions sur les impacts de l'IA, ainsi qu'une mise en perspective vis-à-vis d'observations à plus grande échelle ou dans d'autres secteurs d'activité.

Cette démarche repose en outre sur un partage régulier avec les organisations syndicales. L'échange avec les représentants du personnel inclut la revue des outils dont la collectivité se dote, mais aussi des observations et retours au fil de l'eau quant aux impacts de ces utilisations. Il s'adosse enfin à des temps d'acculturation et de formation pour le partage d'une culture commune.

En matière d'IA, un agent averti en vaut deux !

Pour mener à bien leur travail, les membres du groupe de travail et les représentants des organisations syndicales bénéficient régulièrement de temps d'acculturation et de formation. Deux temps de formation communs au premier semestre de 2024 ont marqué le lancement des échanges autour des conditions de travail : un premier temps d'acculturation ayant permis de considérer les enjeux techniques et humains de cette technologie ; un deuxième temps d'approfondissement des enjeux en matière de ressources humaines.

Par ailleurs, des temps d'acculturation et de sensibilisation se déploient au sein de la collectivité. Des séminaires, sur sollicitation des directions, offrent la possibilité de partager le cadre posé par la collectivité et d'aborder les préoccupations des encadrants et agents en matière d'IA (comment l'IA va-t-elle transformer le travail ? Est-ce que l'IA peut permettre d'être plus efficace ? Du point de vue de l'évaluation du travail, faut-il et si oui comment, distinguer le travail réalisé par l'agent avec ou sans IA ?).

Des webinaires trimestriels sont organisés afin de partager l'état d'avancement de la démarche et certains sujets de questionnement.

Concernant l'IA générative, devant l'augmentation du volume d'utilisation, des ateliers regroupant des agents et des encadrants (dont des directeurs) ont été organisés dans le but de repérer les pratiques et de tester des outils (l'analyse de documents, l'aide à la rédaction, la transcription d'échanges, la production de présentations, etc.). L'idée étant, à l'issue des ateliers, de proposer des recommandations d'outils.

L'association des organisations syndicales s'appuie sur une mise en dialogue social régulière des enjeux de conditions de travail au sein de la collectivité. Ces échanges internes prolongent, dans une symétrie des attentions, la démarche d'écoute et de dialogue menée par Nantes Métropole sur son territoire autour des sujets de l'IA.

Nantes Débat de l'IA :

Pendant neuf mois (de mars 2024 à janvier 2025), les habitants et les professionnels du territoire de la métropole nantaise ont été interrogés sur leur perception des impacts de l'IA. De ce dispositif d'écoute sont ressortis quelques sujets communs aux professionnels et au grand public, notamment l'agentivité vis-à-vis de l'IA, son impact environnemental, mais aussi le besoin d'identifier lorsque l'IA est utilisée (notamment dans des usages d'information et d'éducation), et enfin la demande d'un cadre de régulation.

— 3. L'IA ET LES RISQUES PSYCHOSOCIAUX, UNE QUESTION DE CURSEUR ?

Depuis plus de 10 ans, la collectivité travaille à l'identification et à la prévention des risques psychosociaux auxquels sont exposés ses agentes et agents. Dans cette perspective, un groupe de travail a été créé voici plusieurs années afin d'accompagner spécifiquement les sujets croisant transition numérique et qualité de vie au travail, tels que le phénomène de dématérialisation d'un certain nombre de services et d'actes administratifs, l'évolution des outils bureautiques, ou encore la mise en place du droit à la déconnexion. C'est donc naturellement, dans le prolongement de ces travaux, que le sujet de l'IA a été intégré aux prérogatives de ce groupe de travail.

À partir des familles de facteurs de risques psychosociaux définies par l'INRS[1] (Institut national de recherche et de sécurité pour la prévention des risques professionnels) et de cas concrets ou fictifs d'intelligences artificielles en contexte territorial, le groupe de travail a classé en trois grandes dimensions les impacts de l'IA.

La première dimension concerne les impacts de l'IA sur le travail lui-même, impacts qui se présentent en opportunités ou risques potentiels selon les conditions d'usage et les contextes professionnels dans lesquels ces outils à base d'IA s'insèrent. Par exemple, le gain de temps souvent mis en avant pour promouvoir l'utilisation de l'IA pourrait représenter une chance de réduire des tâches fastidieuses ou pénibles. Cependant, selon la manière dont ce gain de temps est employé ou réutilisé, cela peut entrainer un risque d'intensification du travail. Autre exemple, l'utilisation d'une intelligence artificielle invite fréquemment à accroitre le contrôle humain, ce qui pourrait favoriser la responsabilisation des utilisateurs. Toutefois on observe, y compris pour des outils ne requérant pas d'IA, que les tâches de contrôle ont tendance à se diluer dans le temps, une fois qu'une forme de confiance est installée vis-à-vis de l'outil. Cette dilution engendrerait alors un risque pour la responsabilisation des agents et la qualité du travail. Un dernier exemple dans cette dimension souligne les impacts sur les compétences des agentes et agents. L'utilisation de l'IA peut en effet permettre de développer de nouvelles compétences, soit vis-à-vis de l'IA elle-même (apprendre à rédiger un prompt, par exemple), soit grâce au temps libéré par l'outil pour développer d'autres activités. Néanmoins, à plus long terme et sans maitrise,

1 https://www.inrs.fr/risques/psychosociaux/facteurs-risques.html
Document source : le rapport du collège d'experts, https://travail-emploi.gouv.fr/mesurer-les-facteurs-psychosociaux-de-risque-au-travail-pour-les-maitriser

les tâches réalisées par l'IA et délaissées par les agents pourraient conduire à un affaiblissement des compétences (par exemple rédactionnelles), voire à une déqualification des agents.

La deuxième dimension a trait aux impacts de l'IA sur le rapport au collectif de travail. Sur cette dimension aussi se dessinent des opportunités et des risques en fonction du niveau de régulation de ces effets. Dans les relations interpersonnelles, vis-à-vis des collègues ou de l'encadrement, l'IA peut, dans certaines situations de travail, renforcer le collectif lorsque la mise en place du nouvel outil crée des interactions autour des formes d'usages et des retours de pratiques, renforçant ainsi les échanges et le développement d'un référentiel commun d'usage. De même, le gain de temps autorisé par l'IA peut favoriser les espaces d'échanges au sein de l'équipe si ce temps « gagné » est ainsi réutilisé et réorganisé pour soutenir ces temps collectifs. Cependant, à terme, si l'interaction entre l'agent et l'IA s'intensifie au détriment de l'échange entre collègues, il peut surgir un risque d'isolement et de délitement du collectif. Cette situation est d'ores et déjà observée chez les lycéens et étudiants qui ont tendance à interroger en première intention une IA générative plutôt que leur professeur ou leurs camarades. La situation se transpose aisément au monde du travail et appelle une vigilance. Autre exemple : dans les relations avec les usagers des services publics, le recours à l'IA pourrait contribuer à accroitre le niveau de service, par exemple en proposant une information accessible en temps réel, cette qualité de service conduisant à la qualité des relations entre usagers et agents. En revanche, si l'IA produit des informations erronées, les relations risquent de se tendre.

Enfin, les membres du groupe de travail ont repéré les opportunités et les risques liés à l'utilisation de l'intelligence artificielle dans le rapport au travail. Pour cette troisième dimension, les impacts pourraient se percevoir à plus longue échéance ou par effet rebond des impacts sur les deux dimensions précédentes. L'impact de l'IA sur le sens au travail – l'usage de l'IA pouvant le renforcer ou conduire à son délitement – semble lié à l'utilité perçue vis-à-vis de l'outil. L'usage d'outils à base d'IA peut soutenir l'utilité ressentie par les agents s'il permet d'augmenter la qualité du travail réalisé ou de proposer un service jusqu'alors non rendu. Mais l'usage de l'IA peut aussi dégrader ce sentiment d'utilité lorsque les agents interviennent uniquement pour contrôler l'outil et n'ont plus d'espace de création ou de production dans le travail. La question de l'adéquation entre les valeurs personnelles des agentes et agents et l'utilisation de ces outils est également à prendre en compte. Certains agents sont favorables à l'IA, convaincus de l'efficacité qu'elle autorise, notamment en termes d'amélioration de la qualité du service rendu ; d'autres préfèrent ne pas l'utiliser à cause de ses impacts environnementaux et sociaux. Du point de vue de l'évolution des tâches et des métiers, la mise en place d'un dispositif d'accompagnement sera décisive. L'absence d'accompagnement et un déploiement tous azimuts risquent de positionner les agents face à une évolution subie, source d'inquiétude et potentiellement de décrochage ou de désengagement.

Au-delà de ces trois dimensions, le groupe d'experts a identifié des enjeux managériaux relativement forts. Que l'usage de l'IA soit masqué ou officialisé, l'introduction de ces outils crée d'ores et déjà des collectifs de travail disparates. Entre les agentes et agents qui utilisent l'IA (officiellement ou non) et celles et ceux qui ne l'utilisent pas (par respect du cadre édicté par la collectivité, pour des questions de compétences numériques ou encore de valeurs personnelles), le rôle des managers dans la transformation digitale des organisations s'accroit nécessairement. Quelques questions managériales incontournables sont identifiées : l'organisation et la régulation de l'activité, la sécurisation des données et activités sensibles en veillant à la réversibilité en cas de panne de l'outil, la reconnaissance du travail, la cohésion et la régulation du collectif de travail.

À partir de cette analyse prospective, le groupe de travail a formulé la question du curseur (dans le déploiement de l'IA, l'accompagnement ou encore le niveau de régulation) comme centrale tout en reconnaissant que l'équilibre est délicat à trouver. C'est pourtant la seule voie pour que l'IA soit une chance pour les conditions de travail, sans payer le prix des dommages collatéraux.

À ce jour, peu d'outils à base d'IA sont utilisés dans la collectivité (en tout cas officiellement). Cependant, mieux vaut prévenir que réparer ; anticiper les impacts potentiels de l'IA sur les conditions de travail et instaurer des garde-fous, que prendre en charge des situations dégradées. L'évaluation des impacts de l'IA sur les conditions de travail réalisée dans le cadre de l'observatoire viendra confirmer ou infirmer les hypothèses du groupe de travail. Dans tous les cas, ces hypothèses constituent des points d'attention très utiles au développement et à la mise en place des outils à base d'IA. À partir de cette analyse des impacts, la collectivité nourrit également l'ambition d'identifier et de sélectionner les IA et les types d'usages vertueux pour la qualité de vie et les conditions de travail de ses agentes et agents.

DIALOGUER SUR L'IA ? RETOURS SUR UNE EXPÉRIMENTATION ORGANISATIONNELLE COLLABORATIVE : LE PROJET DIAL-IA

Odile Chagny
(IRES, Réseau Sharers & Workers)
Christophe Teissier
(ASTREES / Ultra Laborans)

Aujourd'hui au cœur de tous les débats, l'IA demeure un sujet émergent du dialogue social. La complexité et l'évolutivité des technologies couvertes par cette notion génèrent de multiples incertitudes quant à leurs impacts tous azimuts, de l'emploi à la préservation de l'environnement, en passant par les conditions et organisations du travail. Face à cette réalité nouvelle, les acteurs du dialogue social doivent être en mesure de s'organiser et d'innover. L'objet de cet article est ainsi de revenir sur une expérimentation organisationnelle et collaborative récente, le projet Dial-IA piloté par l'IRES et soutenu par l'Anact. Conduite pendant 18 mois, elle a visé, en écho aux productions du dialogue social européen, à réfléchir et concevoir les moyens pour soutenir le développement du dialogue social en France sur ce sujet. L'article décrit l'émergence et les modalités d'une approche méthodologique qui a souhaité faire toute sa place à un apprentissage collectif par les acteurs du dialogue social.

Mots-clés : dialogue social, apprentissage collectif, intelligence artificielle

— INTRODUCTION

Face au déploiement tous azimuts, présent et à venir, des systèmes d'IA au sein des organisations privées comme publiques, le projet Dial-IA est né du constat que le dialogue social et ses acteurs ne pouvaient rester inactifs. En dépit de la montée en puissance déjà ancienne du sujet « IA » au niveau européen[1] et des impulsions imprimées par les partenaires sociaux européens pour davantage de

1 Notamment illustré par une multiplicité d'initiatives de l'UE sur ces enjeux, d'une approche de soft law (2018- 2020) jusqu'à l'adoption de nouvelles réglementations, en particulier le règlement européen sur l'IA (2024) et la directive sur les conditions de travail des travailleurs de plateforme.

dialogue social en ces matières[2], peu de traductions concrètes étaient décelables dans le paysage français jusqu'à encore récemment. Ce constat avait déjà été posé par les porteurs du projet européen SeCoIA Deal (Servir la confiance dans l'IA par le dialogue), qui avait regroupé, du printemps 2021 au printemps 2023, plusieurs acteurs français et européens, entrainés par la CFE-CGC avec le soutien de l'IRES (Institut de recherches économiques et sociales au service des organisations syndicales en France) et d'ASTREES. Pour autant, au printemps 2022, le besoin de renforcer le dialogue social sur l'IA et ses impacts sur l'emploi et le travail restait fort en France aux yeux des futurs porteurs du projet Dial-IA. C'est donc pour répondre à ce besoin que Dial-IA a été conçu : fabriquer un outillage utilisable par les représentants du personnel mais aussi, en leur absence, par les dirigeants de TPE-PME pour développer, au sein de leurs organisations, un dialogue dédié à l'IA permettant de tirer un bénéfice effectif de ces outils dans les entreprises et administrations.

Le projet Dial-IA, soumis dans le cadre d'un appel à projets de l'Anact sur le dialogue social technologique[3] (2022), a été coordonné par l'IRES avec le soutien de l'association ASTREES / Ultra Laborans (18 mois en 2023-2024). Il visait à constituer un réseau élargi d'acteurs parties prenantes du dialogue social technologique pour opérer un travail de veille et de partage des pratiques, construire un référentiel méthodologique de déclinaison de l'accord-cadre européen sur la numérisation en France (concernant le volet sur l'IA) et identifier, à un niveau local, des cas d'usage en vue de tester la méthodologie de déclinaison définie.

L'ambition était dès lors grande et constituait d'emblée un défi : face à des outils nouveaux[4] et potentiellement très impactants sur de multiples dimensions (productivité, niveau d'emploi, conditions de travail et organisation du travail, sécurité informatique, impacts environnementaux, etc.), comment construire un outillage pertinent et effectivement mobilisable par une diversité d'acteurs à différents niveaux ?

C'est sur ce défi que cet article propose de revenir. Il ne s'agira donc pas de s'attarder sur les résultats du projet Dial-IA en tant que tels,[5] mais plutôt sur le processus du projet et sa dynamique collective, qui ont permis de déboucher sur des objets concrets, et qui constituent le cœur de l'aventure Dial-IA.

1. L'ÉMERGENCE D'UN PARTI PRIS MÉTHODOLOGIQUE « BOTTOM-UP » – L'APPROCHE SHARERS&WORKERS

Le projet Dial-IA est ancré dans une approche qui lui a préexisté : celle du réseau Sharers&Workers. Ce dernier a été initié en 2015 par l'IRES et l'association ASTREES. Face au bouleversement généré par le développement du numérique[6], l'objectif du réseau a été d'emblée de faire dialoguer des acteurs de la recherche, des acteurs de la transformation numérique en entreprise et des acteurs syndicaux afin d'appréhender collectivement les impacts de ces transformations sur le travail et le dialogue social, et les moyens d'y apporter des réponses adaptées. Le point de départ du réseau était la conviction que la transformation numérique amène nécessairement à renouveler nos façons de penser et d'agir sur le travail, les compétences, les relations de travail, les formes de représentation, les façons de partager la valeur, etc., et qu'il est nécessaire de croiser les points de vue et de mettre en relation l'ensemble des parties prenantes, qui n'ont pas toujours les mêmes approches ni intérêts, afin de mieux appréhender ces transformations. Concrètement, il s'agit par là de créer des espaces de dialogue informel et non

2 Accord-cadre autonome sur la transformation numérique des entreprises de juin 2020.
3 https://www.anact.fr/appel-manifestation-dinteret-developper-le-dialogue-social-technologique-au-travail
4 Le lancement du projet Dial-IA est contemporain de l'explosion très remarquée de l'IA générative.
5 Ils sont librement accessibles en ligne : https://dial-ia.fr/
6 En commençant dès 2015 par l'irruption de l'économie de plateforme.

institutionnel pour se placer en situation de s'attaquer à des dilemmes numériques qui, par nature, défient les cadres institutionnels existants. **Plus fondamentalement, il s'agit d'opter pour un parti pris méthodologique considérant que la pertinence des modalités d'élaboration de solutions ou d'outils revêt autant d'importance que la pertinence conceptuelle de ces derniers.**

Le développement des initiatives de Sharers&Workers est ainsi initialement empirique : il s'agit d'expérimenter les manières par lesquelles organiser et construire un dialogue multi-acteurs dont la nécessité est postulée par la complexité des évolutions technologiques et sociales. De ce point de vue, Sharers&Workers présente des similitudes fortes avec la notion d'« expérimentation organisationnelle[7] » proposée en 2020 par un collectif pluridisciplinaire de chercheurs[8]. Selon ces auteurs, « *les acteurs sociaux confrontés au défi cognitif de la transition numérique doivent rechercher de nouvelles façons d'organiser la production de normes et de règles grâce à la recomposition des ressources* ». Cela les invite à rechercher « *les moyens de développer leurs capacités collectives* » notamment, en premier lieu « *par l'expérimentation organisationnelle, où les acteurs sociaux recherchent de nouveaux types d'organisations, de réseaux et d'alliances, et réfléchissent, évaluent et tirent des enseignements de leurs expériences* ».

Considérant cette perspective, Sharers&Workers constitue une illustration de ce que l'expérimentation organisationnelle peut recouvrir en pratique : une fabrication, voire un bricolage d'arrangements entre une pluralité d'acteurs pour avancer en commun sur un sujet donné. Le projet Dial-IA illustre également cette démarche, dans la continuité de l'approche Sharers&Workers.

— 2. LA FABRIQUE DE DIAL-IA

Les considérations qui précèdent avaient conduit à une segmentation du projet en plusieurs phases : une phase de préparation, une phase de déploiement des activités et une phase de capitalisation. Sans livrer un tableau exhaustif des actions menées, nous évoquons à la suite les points-clés de ces phases pour éclairer la conduite des activités réalisées.

Il s'agissait en premier lieu de constituer une communauté d'acteurs agissante. Communauté « agissante », car il fallait mettre en ordre de marche une pluralité d'acteurs afin de constituer un collectif apte à produire des solutions au problème posé dès le départ : renforcer le dialogue social sur l'IA. En d'autres termes, l'enjeu était d'agréger les ressources nécessaires à l'expérimentation organisationnelle projetée. L'IRES avait noué dès le printemps 2022 un partenariat avec quatre organisations syndicales à un niveau interprofessionnel (CFE-CGC, UGICT CGT, FO Cadres, CFDT), membres de l'IRES. Ce noyau de départ a permis la constitution d'un comité de pilotage du projet. Ce dernier a été le levier principal pour mobiliser un premier cercle de participants au projet, à savoir des représentants du personnel issus de chaque organisation syndicale et appartenant à une variété d'entreprises et d'administrations. Le critère de recrutement était double : mobiliser des acteurs d'univers professionnels différents (entreprises privées, administrations et agences publiques, fédérations syndicales, organismes paritaires) et directement impliqués dans des initiatives ou problématiques relatives aux impacts de l'IA dans le milieu de travail. Il s'agissait ainsi de multiplier les ressources disponibles (regards, expériences et expertises) tout en préparant une large dissémination des résultats du projet. En parallèle, l'équipe d'animation du projet (IRES/ASTREES-Ultra Laborans) a travaillé à un enrichissement du cercle de participants au regard des

7 Nous remercions vivement Pierrette Howayek, doctorante durant les projets SecoIA Deal et Dial-IA, d'avoir partagé ces réflexions théoriques avec nous dans le cadre de ses travaux de recherche.

8 Murray, G., Levesque, C., Morgan, G.D. et Roby, N. (2020), « Disruption and re-regulation in work and employment: from organisational to institutional experimentation », in *Transfer: European Review of Labour and Research*, 26(2), 135-156.

mêmes critères de recrutement. Cela a permis d'élargir le cercle des organisations syndicales intégrées au comité de pilotage, pour l'étendre à d'autres organisations syndicales (Solidaires Finances publiques, CFTC, etc.), à des organisations patronales (UNAPL, CINOV Digital) et à des experts (de l'IA, de l'informatique, des relations sociales, du travail, etc.). À ce processus de recrutement exigeant s'est ajoutée la mise en place d'outils numériques collaboratifs devant garantir une communication et un travail en continu de la communauté. Enfin, un webinaire initial d'intégration des membres de la communauté a eu pour effet de poser les bases du travail en commun ainsi que de fournir quelques points de repère initiaux sur le dialogue social face à l'IA au niveau européen, en particulier la méthodologie de dialogue social promue par l'accord-cadre européen de juin 2020 et les contenus de l'accord européen sur la transformation numérique des administrations d'État et fédérales d'octobre 2022.

La phase de déploiement du projet devait initialement s'articuler autour de deux temps successifs :

— *Un temps d'acculturation et de montée en compétence collective*. Par ce biais, il s'agissait de construire une grammaire commune de l'IA, de ses impacts potentiels sur le travail et de ses formes de régulation : acquérir des fondamentaux partagés pour définir les objets et les conditions de mise en œuvre d'un dialogue social technologique adapté à l'IA, partager des retours d'expérience pour appréhender les objets à traiter, les confronter aux outils existants, nourrir la conception du périmètre, des objectifs et des formes du référentiel de déclinaison de l'accord-cadre européen visé. Ainsi, le caractère techniquement évolutif des SIA a pu être concrètement appréhendé afin de soutenir la nécessité de concevoir les modalités d'un dialogue social itératif, apte à répondre à une technologie en mouvement. Par ailleurs, des expériences spécifiques d'introduction de SIA dans des organisations ont été mobilisées pour constater bien souvent l'absence d'anticipation suffisante des conséquences des SIA sur les ressources humaines, par exemple au sein du ministère de l'Économie et des Finances, ou encore de celui de l'Intérieur.

— *Un temps progressif de co-construction du livrable du projet.*
Ce schéma initial a été bousculé dans la pratique : les deux temps se sont en partie chevauchés dans le calendrier du projet pour répondre au besoin exprimé par les participants d'itérer davantage sur certains sujets (tels que le volet technologique des SIA et l'analyse d'expériences supplémentaires). La phase de co-construction du livrable a par ailleurs nécessité des temps de travail collectif en présentiel (deux ateliers), non envisagés initialement, pour maximiser l'intelligence collective. Ces ajustements du plan de travail initial ont permis de renforcer l'engagement des participants dans la communauté en vue d'atteindre les objectifs du projet.

Cette situation apparait en cohérence avec le processus même d'expérimentation organisationnelle tel que mentionné précédemment. Si une partie du déploiement du projet a pu être délibérée et arrêtée au départ du projet, une autre a dû faire l'objet de renégociations, d'ajustements, au cours de la démarche et avec les acteurs engagés. Il y a là une forme de bricolage *in situ* indissociable de la nature de cette initiative qui supposait d'accorder et de composer avec un collectif multi-acteurs.

3. LES CONTROVERSES SOULEVÉES AU SEIN D'UN PROCESSUS D'APPRENTISSAGE COLLECTIF

Le déroulement du projet s'est accompagné, chemin faisant, d'un certain nombre de controverses au sein de la communauté. Il n'y a là rien de surprenant en soi, s'agissant d'une démarche de co-construction misant sur la recherche collective de ressources à réunir ou à concevoir face à la problématique posée. Elles ont porté tout à la fois sur la composition de la communauté, sur le fond

du sujet mais aussi sur les débouchés attendus du projet. Nous mentionnons ici les plus significatives pour rendre compte du processus d'apprentissage collectif opéré.

Controverses sur la composition de la communauté

Au départ, comme déjà indiqué, le projet Dial-IA est appréhendé comme une initiative intersyndicale. Pour autant, il s'agissait de constituer une communauté agissante dotée d'un nombre suffisant de personnes ressources pour poursuivre les objectifs assignés à la démarche. De cette situation résultent des discussions dès le démarrage du projet : quels acteurs peuvent rejoindre la communauté sans dénaturer le sens de la démarche ? Le sens assigné à la démarche est de dégager des voies pour développer un dialogue social – encore insuffisant – autour de l'IA. Dès lors, les organisations syndicales de salariés doivent-elles mener seules le projet ? N'est-il pas intéressant qu'il s'ouvre aussi aux organisations d'employeurs ou aux autorités publiques ? Quelle place réserver aux acteurs de la recherche ? L'objectif d'outiller concrètement les représentants des salariés sur le terrain rend-il tout apport théorique inutile ? Sur ces deux points, les arbitrages opérés par le collectif ont statué sur la recherche d'une ouverture la plus grande aux acteurs pertinents dans leur diversité (représentants syndicaux, représentants du personnel de terrain, représentants patronaux, responsables des ressources humaines, experts techniques, experts des relations sociales, etc.). Toutefois, si de nombreux efforts ont été déployés pour ouvrir la communauté, ils n'ont pas toujours été couronnés de succès. Il a ainsi été impossible de rallier certains acteurs à notre démarche : des représentants des autorités publiques faute de temps disponible, certaines organisations patronales au niveau interprofessionnel pour des raisons sans doute en partie liées à leur positionnement prudent quant au développement d'une régulation de l'IA, potentiellement perçue comme trop contraignante.

Controverses sur l'objet IA

La première des controverses soulevées révèle des difficultés à adopter un positionnement commun face à un objet nouveau et évolutif. L'IA n'est-elle pas simplement une évolution numérique de plus, à l'instar de celles connues par les organisations depuis la fin du 20e siècle ? Ce questionnement, partagé au départ par nombre de membres de notre communauté, est intéressant. Il suggère en effet qu'il n'y aurait rien de nouveau et donc, très opérationnellement, que les outils existants du dialogue social formel, en particulier des droits conférés aux représentants du personnel et aux organisations syndicales en matière d'information et consultation et de négociation collective au niveau des entreprises et des administrations, pourraient s'avérer suffisants pour y faire face. À cet égard, nos webinaires de montée en compétence ont notamment permis d'appréhender techniquement la spécificité des technologies IA, mais aussi les retours d'expériences sur des solutions IA précises. De nombreux débats ont porté sur la nécessité de déconstruire la notion d'intelligence artificielle, considérée de façon quasi unanime comme beaucoup trop anthromorphique, pour rappeler son ancrage socio-technique et économique. Les notions d'informatique et d'automatisme avancés ont régulièrement été invoquées comme des formulations alternatives, plus à même de couvrir les caractéristiques de ces systèmes. Pour autant, lorsqu'il a fallu adopter une définition commune, c'est vers celle du règlement européen de l'IA (article 3) que les participants se sont tournés, pour des raisons pratiques : c'est en effet cette définition qui s'impose pour identifier les systèmes d'IA et les classifier selon leur niveau de risque. Toutefois, afin de rendre compte de ces débats autour des définitions de l'IA, il a été décidé de dédier une partie du livrable (et donc une partie du site internet) à l'explicitation de ce qu'est l'IA (un produit sur un marché concurrentiel, un processus de traitement des données, des interactions homme-machine, un paradigme de réponse, une empreinte écologique), et de ce qu'elle n'est pas (objective, intelligente, universelle, dotée de sens commun, située). Progressivement, il est apparu également dans les discussions qu'il était nécessaire, au-delà de la définition « juridique » de l'IA, de bien distinguer les systèmes d'IA symboliques des systèmes connexionnistes, pour en appréhender

les spécificités, en particulier pour ce qui concerne la transparence, l'intelligibilité et la fiabilité[9]. On trouve ici une illustration de la construction d'une grammaire commune entre les participants. Cette dernière n'est pas neutre : elle a directement alimenté la réflexion sur les outils à construire et proposer.

Une autre controverse soulevée était directement afférente à la co-construction de notre livrable : peut-on imaginer des outils à visée opérationnelle en raisonnant sur le seul vocable « IA » ? Comment rendre tangible l'objet dont on parle pour développer un dialogue social ? À nouveau, les discussions ont permis de réaliser des arbitrages, en l'occurrence de préférer la notion de « système d'IA » ou de « projet de SIA », plus située, à celle, plus abstraite et indéfinie, d'IA. Ce choix illustre également le cheminement des participants : dès le départ, de par leur positionnement d'acteurs de terrain du dialogue social, la question de savoir comment l'IA se déploie dans l'entreprise ou l'administration, par quels choix et arbitrages, a pris beaucoup de place dans les discussions. Restait à savoir comment formaliser les diverses réponses à cette question. La rencontre, organisée par l'Anact dans le cadre de l'appel à projets, avec d'autres porteurs de projets qui travaillent la question de la discutabilité des projets informatiques[10], a été déterminante. Le choix opéré s'est traduit directement dans les résultats de notre projet : une partie du livrable est dédiée à la cartographie des situations de développement des SIA, en distinguant les phases de développement, leurs modalités (achat d'une solution sur étagère, conception et maitrise de l'ensemble des phases, adaptation et paramétrisation) et, surtout, en identifiant les acteurs concernés, tant internes qu'externes à l'organisation[11]. L'explicitation des situations de développement des SIA, qui s'est imposée dans les échanges, a nécessité un investissement considérable de la part des participants et des organisateurs, sur le plan tant conceptuel que graphique.

Controverses sur les livrables : quoi et comment ?

L'objectif initial du projet était de concevoir un référentiel méthodologique de déclinaison de l'accord autonome européen sur la numérisation. Ce cadrage initial était perçu comme suffisamment souple pour accueillir des idées et productions diverses et concrètes de la communauté. Et ce fut bien le cas.

Rapidement, et sans que cela ait été envisagé au départ, les organisations syndicales associées au pilotage du projet ont souhaité produire un manifeste public[12], actant leur positionnement commun face au développement de l'IA et le besoin de dialogue social associé. Le manifeste a été présenté à l'ensemble de la communauté, il a été validé par les bureaux des organisations syndicales pilotes du projet, puis a reçu l'adhésion d'une organisation syndicale non impliquée dans le portage du projet. Surtout, son contenu s'est nourri des échanges et informations mises en commun dans le cadre du projet. Le manifeste traduit ainsi une forme d'institutionnalisation[13] des travaux de la communauté. Ce document appelle à un « dialogue social au service des bons usages de l'IA et d'une nouvelle étape de progrès social dans les entreprises et les administrations ». Il pose un certain nombre de principes, et de modalités pour y parvenir, en rappelant en particulier la nécessité d'un juste équilibre entre, d'un côté, les actions visant à encourager ces technologies et, de l'autre, celles à même de créer les conditions de leur fiabilité et de leur finalité par la participation de toutes les parties prenantes de la mise en place de ces systèmes. Il appelle à mobiliser les potentialités de dialogue qui existent dans les entreprises comme dans les administrations, à prendre en compte la temporalité propre des SIA, leur « cycle de vie », à sortir d'une approche « statique » du dialogue social. Un point important à souligner

9 Ce qui, sur le terrain, renvoie à la difficulté souvent rencontrée par les représentants du personnel pour déterminer si tel ou tel logiciel en place dans l'entreprise est ou non assis sur l'IA. Un point qui justifie la demande de transparence quant aux technologies introduites dans une entreprise, cf. notre livrable en ligne.

10 Le projet APIDOR porté par l'Université Toulouse Capitole : https://methodes-dialogue-social-numerique.eu/

11 Voir notamment, sur le site du projet, la cartographie des situations de développement des SIA : https://dial-ia.fr/acculturation-a-l-ia/situations-de-developpement

12 https://dialia.alwaysdata.net/

13 Via l'entrée et le traitement dans les processus de décision des organisations syndicales porteuses du projet.

est la mention, dans le manifeste, de modalités élargies de dialogue, impliquant une multiplicité de parties prenantes (fournisseurs, prestataires, clients, usagers, entreprises, plateformes). Cette précision est importante : c'est la seule mention, dans le cadre du projet, d'un périmètre élargi de dialogue. Le fait que cette mention ne soit pas formulée dans d'autres livrables du projet s'explique par l'absence de consensus, cette fois de nature plutôt politique, entre les participants, sur la nécessité de réfléchir hors du périmètre des acteurs traditionnels du dialogue social.

Par ailleurs, les activités menées en commun ont rapidement conduit les participants à imaginer que le projet puisse permettre de forger un projet d'accord national interprofessionnel à soumettre aux organisations d'employeurs non incluses dans la communauté. On pouvait voir là une illustration concrète d'un outil de transposition, en France, de l'accord européen de juin 2020. Pour autant, était-ce bien le rôle d'une communauté expérimentale que de peser directement sur des processus institutionnels bien ancrés[14] ? Après des tâtonnements lors d'ateliers communs[15], la communauté a finalement renoncé à cette ambition, considérant que l'objectif principal du projet était davantage d'outiller les acteurs de terrain pour initier, dans leurs organisations, une démarche de dialogue social. Cela traduit la distance qui peut subsister entre une démarche d'expérimentation telle que Dial-IA et sa reprise par les institutions[16].

— CONCLUSION

Le projet Dial-IA relevant d'une expérimentation organisationnelle, il convient, pour conclure cet article, d'en évaluer les effets ou, plus modestement dans un format limité, d'énoncer certaines leçons qu'il est possible d'en tirer. En premier lieu, la mise en place d'une communauté agissante en dehors des cadres institutionnels établis, c'est-à-dire, ici, dans le cadre de la démarche d'expérimentation organisationnelle mise en œuvre, met en lumière le rôle-clé du tiers facilitateur, joué en l'espèce par l'équipe d'animation du projet[17]. Animé par une posture de neutralité, ce dernier semble indispensable pour relier des univers, perspectives et personnes distincts et accompagner un travail collectif orienté vers la co-construction de solutions. En second lieu, toute expérimentation questionne ses capacités de passage à l'échelle ou, plus précisément, de diffusion large de ses résultats dans un périmètre excédant les frontières de l'expérimentation. Dans le cas du projet Dial-IA, nous avons constaté l'importance prise par le processus d'acculturation collective dans le déroulement du projet. Ce dernier était incontournable compte tenu des innovations technologiques en cause, dont les impacts largement incertains obligent les acteurs à modifier leurs postures habituelles[18]. En parallèle, l'importance de cette phase d'acculturation questionne les possibilités effectives de passage à l'échelle, à d'autres niveaux et dans d'autres environnements que la communauté Dial-IA. Répondre à cette difficulté suppose donc de trouver et mobiliser des canaux institutionnels permettant de prolonger et donner vie aux propositions résultant du projet. Ces considérations générales ne sont pas propres au projet Dial-IA, elles se retrouvent pour toutes les formes d'expérimentations organisationnelles. On vise par là toutes les modalités par lesquelles des acteurs tentent de s'organiser pour agir en commun aux fins de peser sur des situations qu'ils jugent insatisfaisantes[19]. Ce type d'initiative questionne toujours, *in fine*, la capacité à innerver les institutions en place. Une tâche difficile qui ne remet cependant nullement en cause, dans un monde toujours plus fragmenté, la nécessité d'expérimenter collectivement hors des chemins balisés.

14 En l'occurrence ceux du dialogue social au niveau interprofessionnel.
15 Qui se sont notamment traduits par la production d'une trame d'accord-cadre interprofessionnel.
16 Ici celles du dialogue social français.
17 IRES et ASTREES/Ultra Laborans.
18 Y compris, très concrètement, dans leurs manières de penser le dialogue social nécessaire pour approcher les évolutions en cause.
19 On pourrait trouver beaucoup d'exemples de ce type, des conventions citoyennes nationales récentes à des initiatives telles que l'Assemblée populaire du Rhône qui a mobilisé pendant deux ans des citoyens pour inventer une nouvelle gouvernance, durable, du Rhône.

{IA, travail et emploi : Trajectoires annoncées et effets discutés}

121 L'IA GÉNÉRATIVE, L'EMPLOI ET LES CONDITIONS DE TRAVAIL : QUELQUES CONSIDÉRATIONS POUR LES DÉCIDEURS PUBLICS, LES PARTENAIRES SOCIAUX ET LES TRAVAILLEURS

Janine Berg et Pawel Gmyrek

132 APPRÉHENDER L'IMPACT DE L'IA SUR LE TRAVAIL À TRAVERS UNE APPROCHE PAR LES CHAÎNES DE VALEURS GLOBALES.

Clément Le Ludec, Maxime Cornet

142 L'IA GÉNÉRATIVE AU TRAVAIL : QUELLES TRANSFORMATIONS À L'HORIZON ?

Vincent Mandinaud, Yann Ferguson, Aída Ponce del Castillo

153 L'IA À LA RÉUNION : PERCEPTIONS ET STRATÉGIES DES PARTENAIRES SOCIAUX

Jean-Noël Amato, Corinne Dubois, Elisabetta Zibetti

L'IA GÉNÉRATIVE, L'EMPLOI ET LES CONDITIONS DE TRAVAIL : QUELQUES CONSIDÉRATIONS POUR LES DÉCIDEURS PUBLICS, LES PARTENAIRES SOCIAUX ET LES TRAVAILLEURS

Janine Berg et Pawel Gmyrek
(Bureau international du travail)

Cet article examine les implications de l'intelligence artificielle générative (IA générative ou GenAI) pour l'emploi et les conditions de travail, en s'appuyant sur des recherches récentes du Bureau international du travail. Il soutient que, bien que l'IA générative soit susceptible d'avoir des effets significatifs sur les marchés du travail, son impact principal sera la transformation des emplois plutôt qu'une destruction massive d'emplois. En utilisant une méthodologie affinée d'exposition professionnelle, les auteurs montrent qu'environ un quart de l'emploi mondial est exposé à l'IA générative à des degrés variables, avec une exposition plus élevée chez les femmes et dans les pays à revenu élevé. L'analyse met en évidence la façon dont la variabilité des tâches influence le fait que les emplois soient à risque d'automatisation complète, ou plus susceptibles d'être transformés, la technologie ne remplaçant que certaines des tâches réalisées. L'article apporte également un éclairage sur l'utilisation de l'IA dans les fonctions managériales et les préoccupations liées à l'intensification du travail, à la réduction de l'autonomie et à l'augmentation de la surveillance algorithmique. Il met en évidence les contraintes structurelles dans les pays en développement qui peuvent limiter l'adoption et freiner les gains de productivité. Les auteurs insistent sur l'importance du dialogue social et de cadres politiques inclusifs pour gérer les transitions et garantir que les gains de productivité se traduisent par de meilleures conditions de travail et des résultats équitables sur le marché du travail.

Mots-clés : IA générative, emploi, conditions de travail, dialogue social

INTRODUCTION

Le lancement de ChatGPT en novembre 2022, conjugué à d'autres avancées rapides en intelligence artificielle, a intensifié les préoccupations quant aux destructions d'emplois et aux transformations du monde du travail. Les mandants tripartites de l'Organisation internationale du travail (OIT) – les gouvernements, les organisations d'employeurs et de travailleurs de 187 pays – ont sollicité une assistance technique afin de comprendre ces évolutions et de mieux gérer la transition.

Étant donné que la recherche existante et le débat politique se sont principalement concentrés sur les économies à revenu élevé, notamment aux États-Unis et en Europe, l'OIT a cherché à apporter une perspective mondiale à l'analyse de l'IA et des marchés du travail, et à intégrer d'autres enjeux dans le débat, comme la fracture numérique, les effets sur les conditions de travail et le rôle essentiel du dialogue social dans la gestion de la transformation. À l'automne 2024, l'OIT a lancé un Observatoire sur l'IA et le travail dans l'économie numérique pour servir de référentiel à ses travaux de recherche en cours dans ce domaine.

Dans cet article, nous mettons en lumière certains des principaux enseignements des recherches de l'OIT concernant les conséquences possibles de l'intelligence artificielle générative (IA générative ou GenAI) sur la quantité d'emplois, mais aussi sur les conditions de travail. En règle générale, on distingue deux grands types d'applications des technologies d'IA sur le lieu de travail. Le premier vise à automatiser les tâches réalisées par les travailleurs, en particulier les tâches routinières et répétitives qui peuvent être efficacement prises en charge par des machines. Le second consiste à utiliser des analyses et des algorithmes fondés sur l'IA pour remplacer ou « augmenter » les fonctions de gestion : recrutement, surveillance, supervision et formation des travailleurs, ainsi que la planification des horaires et des pauses – ce que l'on désigne communément par « gestion algorithmique ». Ces deux formes d'utilisation ont des implications tant sur la quantité d'emplois (le nombre d'emplois) que sur la qualité de l'emploi, y compris le respect des principes et droits fondamentaux au travail.

En ce qui concerne l'utilisation des technologies d'IA pour automatiser les tâches, une telle automatisation n'implique pas nécessairement des suppressions d'emplois, car la technologie peut également compléter ou « augmenter » le travail humain lorsque seules certaines tâches sont automatisées. Les emplois peuvent être décrits comme un « ensemble de tâches » (Autor, 2015) ; ainsi, le fait que l'adoption de la technologie conduise à une automatisation (perte d'emplois) ou à une « augmentation » (complémentarité avec le travail humain) dépendra de l'importance d'une tâche donnée dans une profession, ainsi que de la variabilité des tâches effectuées au sein de cette profession.

L'un des principaux constats des recherches de l'OIT est que, si l'IA générative est susceptible d'avoir des effets significatifs sur le monde du travail, en particulier pour ce qui concerne le « travail de connaissance », ces effets prendront principalement la forme d'une transformation des emplois plutôt que d'une destruction massive d'emplois. Il y aura des pertes d'emplois, mais elles seront plus probablement concentrées dans certaines professions et se manifesteront sur le marché du travail par un ralentissement des embauches (Crane et Soto, 2026). L'ampleur de ces transformations a des implications profondes du point de vue des réponses politiques à apporter et de la manière dont les gouvernements et les organisations de travailleurs et d'employeurs abordent le changement technologique. Si la plupart des emplois sont transformés et non supprimés, le principal défi consiste à gérer les transitions afin que les évolutions de l'organisation du travail améliorent, et non détériorent, les conditions de travail.

1. L'EXPOSITION PROFESSIONNELLE À L'INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE

L'OIT a publié en 2023 ses estimations concernant l'exposition potentielle des emplois à l'intelligence artificielle générative (Gmyrek et al., 2023), puis a actualisé ces estimations en 2025 en utilisant une méthodologie plus affinée (Gmyrek et al., 2025). En 2023, la méthodologie reposait sur les descriptions de tâches issues de la Classification internationale type des professions à quatre chiffres (CITP-08) ainsi que sur le modèle GPT-4 d'OpenAI pour estimer des scores d'exposition à l'IA générative au niveau des tâches et des professions. Les estimations de 2025 ont intégré des informations plus détaillées sur les tâches effectuées et introduit un niveau supplémentaire de validation par des experts. En outre, elles se sont appuyées sur la classification nationale des professions à six chiffres de la Pologne, ce qui a permis d'augmenter le nombre de tâches analysées à plus de 30 000 – soit 10 fois plus qu'au niveau à quatre chiffres de la CITP-08. Les nouveaux scores ajustés ont alors été intégrés dans plusieurs modèles d'IA qui ont généré de nouveaux scores en tenant compte de ce jugement humain. L'étape finale a consisté à recalculer les scores pour les tâches de la CITP-08[1].

Les scores ont ensuite été intégrés dans un cadre de référence s'appuyant sur deux critères pour classer chaque profession à quatre chiffres : son score professionnel moyen et l'écart-type (ET) de ses scores au niveau des tâches (Tableau 1). Les scores des 436 professions ont été classés selon un spectre d'exposition comportant quatre niveaux d'exposition croissants.

— Le gradient 1 correspond aux professions présentant une faible exposition globale à l'IA générative et une forte variabilité des tâches (*faible exposition, forte variabilité des tâches*). Bien que certaines tâches au sein de ces fonctions puissent présenter un fort potentiel d'automatisation, ces professions comportent de nombreuses tâches qui continuent de nécessiter une intervention humaine, ce qui correspond étroitement à l'idée de renforcement des capacités humaines.
— Le gradient 2 regroupe des professions caractérisées par une exposition modérée à l'IA générative et un mélange de tâches fortement exposées et faiblement exposées, entrainant des effets inégaux où certaines tâches peuvent être impactées tandis que d'autres restent inchangées (*exposition modérée, forte variabilité des tâches*).
— Le gradient 3 concerne des professions où une part importante des tâches est de manière constante exposée à l'IA générative, signalant des risques croissants d'automatisation et nécessitant des stratégies d'adaptation pour les travailleurs (*exposition significative, forte variabilité des tâches*).
— Le gradient 4 met en évidence les professions présentant la plus forte proportion de tâches exposées à une automatisation potentielle par l'IA générative, avec une forte homogénéité de cette exposition entre les tâches (*exposition très élevée, faible variabilité des tâches*). Ce gradient correspond étroitement à la notion de risque d'automatisation.
— En outre, deux autres catégories sont distinguées : **exposition minimale** et **non exposé**, où il peut exister une interaction limitée avec l'IA générative sans modification significative de la nature des fonctions, ou bien aucune possibilité observable d'automatiser les tâches (Figure 1).

L'utilisation de gradients d'exposition croissants permet de rendre compte de la manière dont l'IA générative peut affecter les professions à différents niveaux d'exposition, en fonction de l'exposition moyenne à l'automatisation et de la variabilité des tâches. La flexibilité de ce cadre renforce également son applicabilité à des contextes nationaux spécifiques, permettant aux décideurs publics de mieux comprendre la répartition des impacts potentiels de l'IA générative et de prioriser les interventions en faveur des groupes les plus concernés au niveau national. Nous soulignons que

1 Pour plus de détails sur la méthodologie, voir le document de travail n° 140 de l'OIT (Gmyrek et al., 2025).

ces classifications sont uniquement indicatives, car tout système d'évaluation des tâches comporte une part de subjectivité. En outre, les scores ne tiennent pas compte des progrès technologiques futurs ni de la question de savoir si ces technologies seront effectivement mises en œuvre dans la pratique.

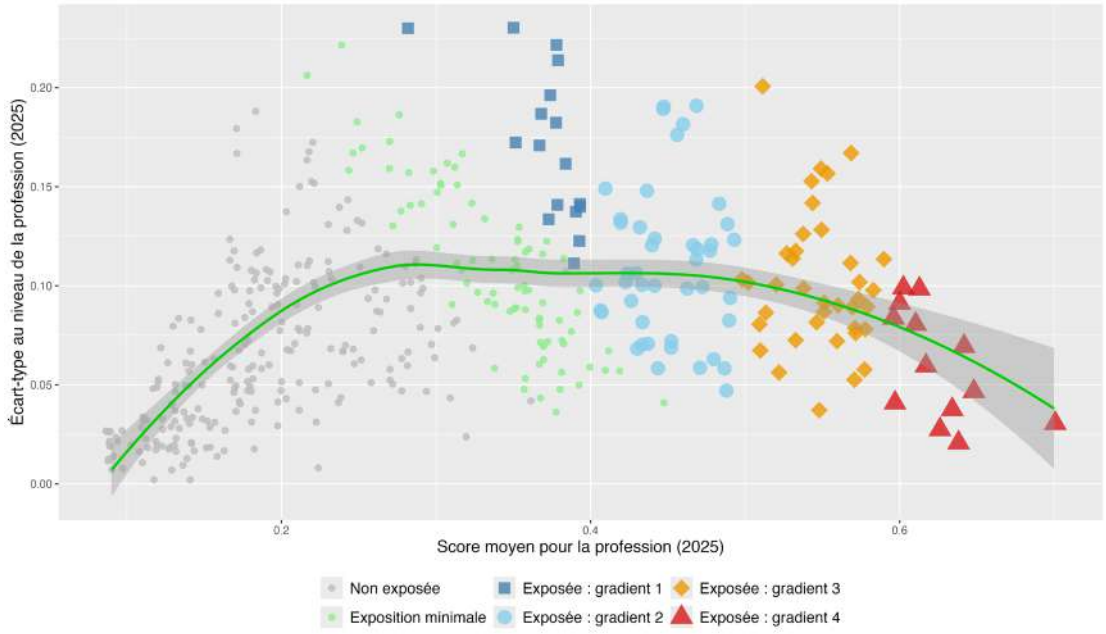


Figure 1 : Gradients d'exposition à l'IA générative : professions au niveau à quatre chiffres de la CITP-08. Source : Gmyrek et al. 2025.

EXPOSITION À L'IA GÉNÉRATIVE (GENAI)	DÉFINITION (MOYENNE ET ÉCART-TYPE DES NOTES AUX TÂCHES)	INTERPRÉTATION
Exposé : Gradient 4 (exposition la plus élevée, faible variabilité des tâches)	$\mu \geq 0,6$ et $\mu - \sigma >= 0,5$	Exposition élevée et homogène à l'IA générative pour les tâches au sein de la profession. La plupart des tâches actuelles de ces emplois présentent un fort potentiel d'automatisation, avec une faible variabilité de l'exposition au niveau des tâches.
Exposé : Gradient 3 (exposition significative, forte variabilité des tâches)	$0,5 \leq \mu < 0,6$ et $\mu + \sigma \geq 0,5$	Exposition professionnelle supérieure à la moyenne : bien que certaines tâches restent moins exposées, le potentiel global d'automatisation des tâches actuelles par l'IA générative est en augmentation dans ces professions.
Exposé : Gradient 2 (exposition modérée, forte variabilité des tâches)	$0,4 \leq \mu < 0,5$ et $\mu + \sigma \geq 0,5$	Exposition modérée à l'IA au niveau professionnel, avec une forte variabilité des tâches. Ces professions comprennent un mélange de tâches exposées et non exposées à l'IA générative, ce qui entraine des effets inégaux.
Exposé : Gradient 1 (faible exposition, forte variabilité des tâches)	$\mu < 0,4$ et $\mu + \sigma \geq 0,5$	Faible exposition globale à l'IA générative au niveau de la profession et forte variabilité des tâches. Certaines tâches présentent un potentiel d'automatisation élevé, même si la profession dans son ensemble reste largement fondée sur des tâches peu automatisables.
Faible exposition (exposition minimale, variabilité modérée des tâches)	$\mu < 0,5$ et $\mu + \sigma > 0,4$	Professions présentant une faible exposition à l'IA générative, où certaines tâches montrent un potentiel d'automatisation modéré, mais où l'exposition globale de la profession reste limitée.
Non exposé	Les professions qui ne remplissent aucune des conditions susmentionnées.	Professions dans lesquelles la plupart des tâches restent relativement peu affectées par l'IA générative, avec une faible variabilité des tâches et un score moyen d'exposition stable et faible.

Tableau 1 : Gradients d'exposition à l'IA générative par profession.
Source : Gmyrek et al. 2025.

Comme dans les estimations de 2023, les professions administratives continuent de présenter les niveaux d'exposition les plus élevés. Toutefois, des changements notables sont intervenus dans la distribution des moyennes par profession entre 2023 et 2025 (Figure 2). Plusieurs professions qui avaient auparavant obtenu les scores les plus élevés ont vu leur score moyen diminuer, ce qui met en évidence la distinction entre le potentiel théorique d'automatisation d'une tâche et les enseignements pratiques tirés de deux années d'expérimentation avec les outils d'IA générative.

Si des tâches telles que la prise de notes en réunion ou la planification de rendez-vous peuvent bénéficier de manière significative du soutien de l'IA générative, les scores de 2023 – qui atteignaient pour certaines tâches jusqu'à 0,9 – reflétaient une évaluation trop optimiste du potentiel d'automatisation complète. Dans les estimations de 2025, le score maximal au niveau des tâches est de 0,76 et la moyenne professionnelle la plus élevée est de 0,7 (gradient 4), ce qui signifie que, même dans ces professions à plus haut risque, certaines tâches nécessitent encore une intervention humaine[2].

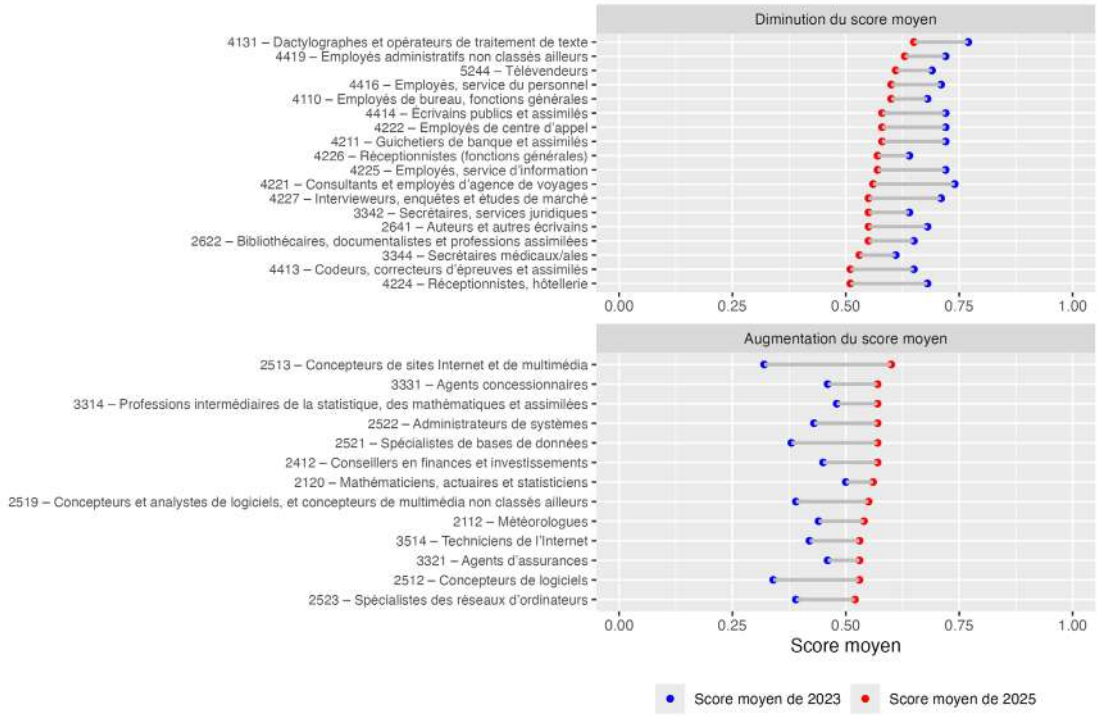


Figure 2. Écart entre les moyennes des professions classées au niveau à quatre chiffres de la CITP-08 ayant connu les variations les plus importantes entre 2023 et 2025.
Source : Gmyrek et al., 2025.

Néanmoins, plusieurs professions fortement numérisées – telles que les développeurs Web et multimédia, les spécialistes des statistiques et des bases de données, ainsi que les métiers liés à la finance et aux logiciels – ont vu leur score moyen augmenter par rapport à 2023. Cette hausse s'explique en partie par l'expansion rapide des capacités de l'IA depuis notre étude de 2023. À cette époque, les grands modèles de langage (LLM) étaient principalement considérés comme des générateurs de texte avancés. Depuis lors, ils ont acquis des capacités multimodales leur permettant de traiter et de générer du texte, des images, de l'audio et de la vidéo. De plus, ces modèles ont développé certaines capacités agentiques qui leur offrent la possibilité d'exécuter de manière autonome des tâches en plusieurs étapes, d'interagir avec des environnements logiciels et de prendre des décisions sur la base d'informations contextuelles. Ces avancées élargissent considérablement l'éventail des tâches pouvant être au moins partiellement automatisées grâce à l'IA générative, les secteurs du développement logiciel et du multimédia figurant parmi les principaux concernés.

Il convient toutefois de garder à l'esprit que, comme lors des transitions technologiques passées, de nouveaux rôles au sein des professions ainsi que de nouvelles professions devraient émerger parallèlement

2 Pour consulter la liste complète des scores pour chaque groupe professionnel à quatre chiffres de la CITP-08, voir l'annexe du document de travail n° 140 (Gmyrek et al., 2025).

aux avancées technologiques. Dès lors, une question-clé pour les politiques publiques est de savoir comment garantir que les travailleurs actuellement employés dans ces professions puissent être formés, si nécessaire, afin de s'y maintenir, et quels seront les besoins en compétences des professions futures. En outre, des interrogations subsistent quant à la manière dont cette transformation pourrait influencer l'évolution de la qualité de l'emploi. Les travailleurs disposeront-ils de davantage de temps pour effectuer un travail créatif ou, au contraire, l'automatisation entrainera-t-elle une plus grande standardisation des tâches et une réduction de l'autonomie ?

— 2. LES PROFESSIONS EXPOSÉES EN PART DE L'EMPLOI

L'étape suivante de notre analyse utilise les 112 professions identifiées comme exposées (gradients 1 à 4) et estime la part de l'emploi que ces professions représentent aux niveaux mondial, régional et selon les groupes de revenu. Nous appliquons la même méthode que dans Gmyrek et al. (2023), qui repose sur le modèle d'estimation globale de l'OIT. Ce modèle utilise la base de microdonnées harmonisées de l'OIT ainsi qu'un processus d'imputation hiérarchique intégrant les données à travers la structure de la CITP-08, en mobilisant le niveau de détail le plus élevé disponible[3].

Au niveau mondial, environ un quart de l'emploi total se situe dans l'un des quatre gradients d'exposition, avec des différences significatives entre les femmes et les hommes, en particulier dans les deux gradients d'exposition les plus élevés (Figure 3). Parmi les hommes, environ un cinquième des emplois relèvent de l'un des gradients d'exposition, dont 3,1 % dans le gradient 3 et 2,4 % dans le gradient 4, la catégorie d'exposition la plus élevée. En revanche, non seulement la part totale des femmes en emploi exposées à l'IA générative est nettement plus élevée, mais l'écart se concentre dans les deux gradients supérieurs, avec 5,7 % des femmes dans le gradient 3 et 4,7 % dans le gradient 4. Ces disparités sont encore plus marquées dans les pays à revenu élevé.

Les différences d'exposition entre groupes de pays selon le niveau de revenu sont également frappantes, les pays à revenu élevé présentant la plus forte part d'emploi relevant de l'un des quatre gradients d'exposition (34 %) et donc la plus faible part d'emplois faiblement ou non exposés à l'automatisation des tâches par l'IA générative (Tableau 1). La part totale de l'emploi exposé diminue nettement à mesure que le niveau de revenu baisse, pour atteindre seulement 11 % dans les pays à faible revenu. Des niveaux de revenu plus faibles réduisent par ailleurs les écarts d'exposition entre les sexes, principalement en raison de la moindre concentration des emplois dans les gradients d'exposition les plus élevés (gradients 3 et 4). À l'inverse, ces écarts sont plus marqués dans les régions plus riches, en particulier en Europe et en Asie centrale, où 39 % de l'emploi féminin relève de l'un des quatre gradients d'exposition, contre 26 % pour les hommes. Ces tendances reflètent à la fois la ségrégation professionnelle entre les sexes et le fait que les emplois exposés à l'IA générative – tels que les emplois administratifs, financiers et de service à la clientèle – sont davantage concentrés dans les pays à revenu élevé.

Il est cependant important de se rappeler que ces estimations ne tiennent pas compte des obstacles à l'application des technologies d'IA générative ni d'autres facteurs susceptibles d'en influencer l'utilisation. En particulier, les pays en développement peuvent faire face à des contraintes liées aux infrastructures physiques nécessaires, telles qu'un approvisionnement en électricité peu fiable ou un accès limité au haut débit, ou encore à une insuffisance de compétences numériques pour utiliser ces technologies. Les organisations qui n'ont pas numérisé leurs documents ou dossiers de travail tireront

3 Les calculs des agrégats d'emploi ont été réalisés par David Bescond, Département de recherche et de statistiques, BIT.

peu de valeur de ces technologies. Par ailleurs, le cout constitue un facteur à prendre en compte : ces technologies ne sont pas gratuites et, dans les pays où le cout du travail est faible, leur utilisation peut ne pas générer de gains financiers immédiats.

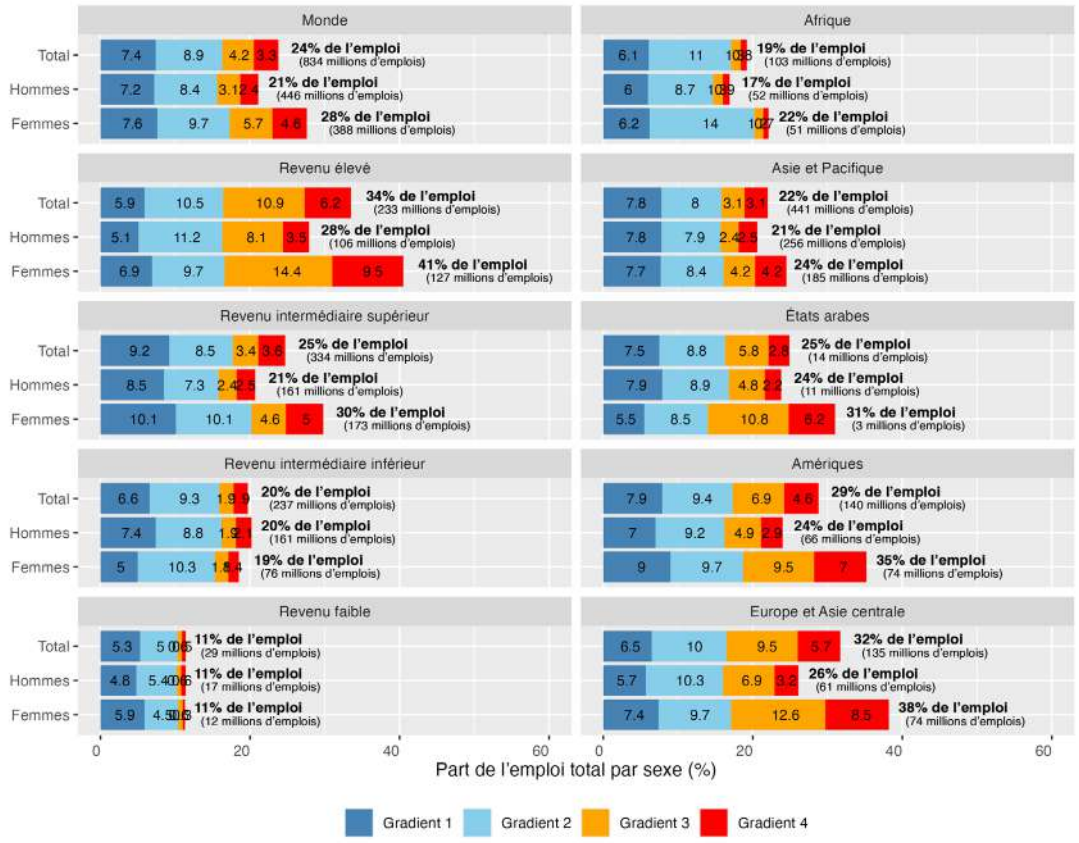


Figure 3 : Estimations mondiales des professions potentiellement exposées à l'IA générative (% de l'emploi, par sexe).
Source : Gmyrek et al., 2025.

Ces estimations ne tiennent pas compte des avancées technologiques futures, qui pourraient encore faciliter l'application de ces technologies aux tâches humaines. Elles ne prennent pas non plus en considération les nouveaux emplois susceptibles d'être créés. S'il peut être difficile d'imaginer la nature de ces derniers, il faut se souvenir qu'au cours de l'histoire, les progrès technologiques ont conduit à l'émergence de nouvelles professions. Ainsi, voici 20 ans, il n'existait pas de gestionnaires de réseaux sociaux, et voici 30 ans, les concepteurs de sites Web étaient peu nombreux. Comme le montrent Autor et al. (2021), environ 60 % des emplois aux États-Unis en 2018 correspondaient à des professions qui n'existaient pas dans les années 40.

Les obstacles à l'utilisation de l'intelligence artificielle dans les pays en développement constituent une source de préoccupation. Comme l'indiquent les recherches de l'OIT, certaines professions pourraient bénéficier de l'IA, notamment lorsque certaines tâches sont remplacées, ce qui peut améliorer l'efficacité des travailleurs. Toutefois, ces gains de productivité dépendent de la disponibilité des infrastructures physiques nécessaires pour soutenir ces technologies, ainsi que des compétences numériques. Deux études menées par l'OIT et la Banque mondiale sur l'IA générative et les pays en développement (Gmyrek et al., 2026 ; Gmyrek et al., 2024) révèlent que l'utilisation des ordinateurs au travail est faible dans de nombreux pays en développement. De plus, la nature et la répartition des tâches diffèrent selon les pays : les travailleurs y effectuent beaucoup moins de tâches analytiques non routinières ;

or, ce sont précisément ces tâches qui pourraient être automatisées grâce à l'IA générative. Ainsi, ces études montrent que près de la moitié des emplois susceptibles de bénéficier de l'IA générative ne peuvent exploiter pleinement ce potentiel. À l'inverse, les métiers menacés d'automatisation utilisent majoritairement un ordinateur au travail, ce qui suggère l'absence de protection contre les effets négatifs potentiels des suppressions d'emplois.

3. LE POTENTIEL ACCRU DE TRANSFORMATION MET EN ÉVIDENCE L'IMPORTANCE DES ENJEUX LIÉS À LA QUALITÉ DE L'EMPLOI

L'analyse de l'OIT montre clairement que l'effet le plus probable de l'intégration de l'IA dans le monde du travail est la transformation des professions, étant donné qu'une part bien plus importante de l'emploi se concentre dans les gradients 1 et 2, ou dans la catégorie d'exposition minimale. Ces transformations ont des implications importantes pour la qualité de l'emploi. Si l'IA peut « augmenter » certaines tâches et améliorer la productivité, elle peut également intensifier le travail, réduire l'autonomie et accroitre la surveillance. Les systèmes de gestion algorithmique, souvent intégrés dans des outils basés sur l'IA, peuvent mener à un suivi plus strict des performances, à une surveillance en temps réel et à une évaluation des travailleurs fondée sur les données, parfois avec une transparence limitée.

Lorsqu'ils sont appliqués à la gestion des performances, les systèmes intégrant l'IA sont conçus pour mesurer en continu l'activité des travailleurs et utiliser les principes du taylorisme et de l'organisation scientifique du travail afin d'analyser les comportements et les résultats. Dans des emplois hautement standardisés et répétitifs – tels que la logistique d'entrepôt et l'industrie manufacturière –, l'IA autorise un suivi fin et en temps réel de la vitesse, des mouvements et des performances des travailleurs, se traduisant généralement par une augmentation des objectifs de performance et une intensification de la pression au travail. Même lorsque les indicateurs de performance semblent objectifs, ils sont influencés par des choix de conception qui déterminent ce qui est considéré comme une activité « productive », en négligeant fréquemment des facteurs contextuels tels que les aléas dans la réalisation du travail et les considérations de santé et de sécurité. En conséquence, ces systèmes peuvent contribuer à une intensification du travail, à un risque accru de blessures et à un moindre contrôle des travailleurs sur le rythme et l'organisation de leurs tâches (Berg et Johnston, 2025).

En outre, les progrès en matière de collecte et de traitement des données ont étendu la gestion des performances fondée sur l'IA à des professions moins standardisées, y compris le travail de connaissance et les emplois impliquant des interactions interpersonnelles. Les outils utilisant la reconnaissance vocale, le traitement du langage naturel et les technologies de « reconnaissance des émotions » sont désormais capables de surveiller non seulement ce que font les travailleurs, mais aussi la manière dont ils le font – en fournissant un retour d'information en temps réel tout en limitant leur marge de décision. L'essor du télétravail depuis la pandémie de Covid-19 a encore accéléré l'adoption de systèmes de surveillance qui suivent les frappes au clavier, l'activité à l'écran et la présence en ligne, en s'appuyant souvent sur des indicateurs indirects pour évaluer la qualité des performances. Or, ces indicateurs ne sont pas toujours fiables ni transparents et peuvent conduire à des évaluations erronées, à des incitations inadaptées et à un affaiblissement de la capacité managériale à fournir un encadrement pertinent. Dans certains cas, des systèmes de notation algorithmique opaques laissent les travailleurs dans l'incertitude quant aux critères d'évaluation de leur performance et aux moyens de l'améliorer. Pris ensemble, ces développements mettent en évidence le risque que les systèmes d'IA, s'ils sont mal conçus ou mal encadrés, détériorent la qualité de l'emploi, ce qui renforce la nécessité de garanties solides et d'un dialogue social (Berg et Johnston, 2025).

Gérer la transition : l'importance du dialogue social

Les gouvernements et les partenaires sociaux jouent un rôle central dans la manière dont l'intelligence artificielle générative façonne le monde du travail. Les réponses politiques doivent viser à élargir les possibilités de croissance de la productivité tout en promouvant le travail décent et en atténuant les effets négatifs associés au changement technologique. Un pilier essentiel de ces efforts est un dialogue social efficace, indispensable pour garantir que les transitions technologiques soient inclusives.

Le dialogue social remplit quant à lui une mission déterminante pour garantir que l'intégration des nouvelles technologies produise des résultats positifs tant pour les entreprises que pour les travailleurs. Associer les travailleurs au processus d'adoption technologique est non seulement équitable, mais aussi efficace, car ce sont eux qui connaissent le mieux leur travail et peuvent ainsi indiquer quelles technologies sont les plus utiles, quels processus se prêtent à l'automatisation et quels types de formation sont nécessaires. À l'inverse, le déploiement de technologies sans implication des travailleurs mène fréquemment à des résultats sous-optimaux, soit parce que les systèmes sont mal adaptés aux réalités opérationnelles, soit parce qu'ils affaiblissent le sens, l'engagement et l'implication au travail. L'approche la plus efficace pour intégrer les technologies repose donc sur un dialogue structuré et continu entre les employeurs, les travailleurs et leurs représentants.

Ce dialogue peut se dérouler dans des cadres institutionnels formels – tels que les comités d'entreprise ou via des dispositions intégrées aux conventions collectives – ou à travers des mécanismes plus informels dans des lieux de travail caractérisés par une forte participation des salariés. Les données comparatives issues de l'Europe indiquent que les pays dotés de systèmes de consultation au travail bien établis et coopératifs, en particulier les pays nordiques et l'Allemagne, présentent généralement des niveaux plus élevés d'acceptation des technologies par les travailleurs. Plus largement, le dialogue social est indispensable pour faire face aux multiples implications de l'intelligence artificielle générative, qu'il s'agisse de gérer le chômage causé par les transformations technologiques ou de veiller à ce que les gains de productivité s'accompagnent d'améliorations de la qualité de l'emploi.

Dans le contexte de l'automatisation, les politiques publiques devraient privilégier le redéploiement et la formation plutôt que les suppressions d'emplois, en particulier dans les secteurs les plus exposés à l'intelligence artificielle générative. Le dialogue social est essentiel à cet égard, car il favorise des approches collaboratives entre les employeurs, les représentants des travailleurs et les pouvoirs publics afin de concevoir des mesures visant à éviter ou à limiter les licenciements. Lorsque des pertes d'emploi se produisent, il est indispensable de veiller à ce que les travailleurs concernés bénéficient d'une protection sociale adéquate et aient accès à des possibilités de reconversion efficaces. Les politiques doivent également prendre explicitement en compte la dimension genrée des impacts sur le marché du travail, en reconnaissant que l'IA générative est susceptible d'affecter différemment les femmes et les hommes, et en répondant aux besoins spécifiques à chaque sexe tout au long de la transition. Parallèlement, les investissements devraient être orientés vers des secteurs sous-financés présentant un fort potentiel de création d'emplois de qualité, notamment dans les secteurs des soins et de l'économie verte.

Enfin, la coopération internationale sera essentielle pour garantir que les bénéfices de l'intelligence artificielle générative soient largement partagés. Le soutien aux pays à faible revenu – par le biais du transfert de technologies, de la coopération au développement et de mesures telles que la restructuration et l'allègement de la dette – peut permettre les investissements nécessaires dans les infrastructures afin de tirer parti du potentiel de gains de productivité offert par l'IA générative. La création d'emplois de qualité liée aux transformations technologiques doit être au cœur de toutes les discussions politiques.

RÉFÉRENCES

— Autor, D. (2015), « Why Are There Still So Many Jobs? The History and Future of Workplace Automation », in *Journal of Economic Perspectives,* 29 (3): 3–30.
— Autor, D., Salomons, A. et Seegmiller, B. (2021), « New Frontiers: The Origin and Content of New Work, 1940 – 2018 », MIT Working Paper, juillet.
— Berg, J. et Johnston, H. (2025), « AI in human resource management: The limits of empiricism », ILO Working Paper 154 (Genève, ILO). https://doi.org/10.54394/NMSH7611
— Crane, L. et Soto, P. (2026), « AI and Coder Employment: Compiling the Evidence », Federal Reserve Board, Working Paper.
— Gmyrek, P., Berg, J. et Bescond, D. (2023), « Generative AI and jobs: A global analysis of potential effects on job quantity and quality », ILO Working Paper 96 (Genève, ILO). https://doi.org/10.54394/FHEM8239
— Gmyrek, P., Winkler, H. et Garganta, S. (2024), « Buffer or Bottleneck? Employment Exposure to Generative AI and the Digital Divide in Latin America », ILO Working Paper 121 (Genève, ILO et la Banque mondiale). https://doi.org/10.54394/TFZY768
— Gmyrek, P., Berg, J., Kamiński, K., Konopczyński, F., Ładna, A., Nafradi, B., Rosłaniec, K. et Troszyński, M. (2025), « Generative AI and Jobs: A Refined Global Index of Occupational Exposure », ILO Working Paper 140 (Genève, ILO). https://doi.org/10.54394/HETP0387
— Gmyrek, P., Viollaz, M. et Winkler, H. (2026), « Disruption without Dividend?: How the Digital Divide and Task Differences Split GenAI's Global Impact », ILO Working Paper 166 (Genève, ILO). https://doi.org/10.54394/0003314

APPRÉHENDER L'IMPACT DE L'IA SUR LE TRAVAIL À TRAVERS UNE APPROCHE PAR LES CHAÎNES DE VALEURS GLOBALES.

LES CONSÉQUENCES DE LA DIVISION INTERNATIONALE DU TRAVAIL DE L'IA SUR LA PRÉCARISATION DU TRAVAIL

Clément Le Ludec
(Télécom Paris, Institut polytechnique de Paris)
Maxime Cornet
(THEMA, CY Cergy Paris Université)

L'intelligence artificielle est souvent présentée comme une technologie susceptible de transformer en profondeur le travail, principalement à travers l'automatisation des tâches. Pourtant, ces analyses peinent à saisir les transformations concrètes du travail liées au déploiement des systèmes d'IA, en particulier lorsqu'elles négligent les dimensions organisationnelles, sociales et géographiques de leur production. À partir d'une enquête qualitative et quantitative menée entre la France et Madagascar, cet article propose d'analyser l'impact de l'IA sur le travail à travers une approche par les chaines de valeur globales. Il montre que le développement des systèmes d'IA repose sur un travail humain intensif, largement externalisé vers les pays des Suds, en particulier dans les activités d'annotation et de maintenance des modèles. Loin de conduire à une disparition du travail humain, l'automatisation algorithmique s'accompagne ainsi d'un allongement des chaines de production et d'une reconfiguration des hiérarchies du travail, s'inscrivant dans des dynamiques de précarisation des travailleurs qui rendent les systèmes d'IA opérationnels.

Mots-clés : Intelligence artificielle, travail, chaines de valeur globales, externalisation, précarisation

Certains économistes présentent l'intelligence artificielle (IA) comme une technologie de changement structurel pour le travail (Frey et Osborne, 2017 ; Brynjolfsson et al., 2017 ; Acemoglu et Restrepo, 2019). Les effets des systèmes automatisés sur le travail restent cependant difficiles à évaluer, en raison de la diversité des contextes d'application et du manque de données sur leur déploiement dans les environnements professionnels. Par ailleurs, si ces analyses mettent en avant le lien entre la *faisabilité de l'automation* (possibilités techniques) et la décomposition des emplois en tâches automatisables, elles négligent les dimensions sociales, organisationnelles et institutionnelles (Benedetto-Meyer et Boboc, 2021 ; Krzywdzinski, 2022), qui jouent pourtant un rôle crucial dans les transformations du travail liées à l'automatisation.

Notre contribution vise précisément à déjouer le piège du déterminisme technologique en mobilisant la théorie socio-économique de l'automatisation au travail (Socioeconomic Company-Level Theory of Automation at Work). Avec ce cadrage théorique, Martin Krzywdzinski (2022) identifie « *l'interaction entre l'architecture du produit (complexité du produit) et la complexité du processus* », ainsi que les « *stratégies de profit des entreprises et les relations de pouvoir entre la direction et les travailleurs* » comme étant les facteurs déterminants des conséquences de l'automatisation au travail. Dans cette perspective, nous pensons qu'il est nécessaire d'articuler l'analyse du travail de fabrication des systèmes d'IA avec l'analyse des conséquences de ces systèmes sur le travail de manière plus générale.

Le développement de l'intelligence artificielle repose en effet sur un travail humain intensif, en particulier dans l'annotation de données et l'entrainement des modèles (Gray et Suri, 2019 ; Tubaro et al., 2020 ; Le Ludec et al., 2023 ; Girard-Chanudet, 2025). Ce processus renforce une structure duale dans laquelle le travail informatique visible et valorisé se concentre dans les économies développées, tandis que le travail essentiel mais invisible qui soutient les systèmes d'IA est relégué à des travailleurs du Sud global (Anwar et Graham, 2020 ; Le Ludec et Cornet, 2023 ; Tubaro et al., 2025).

Sur la base d'une enquête de terrain qualitative et quantitative auprès de salariés de start-up d'IA françaises et d'annotateurs de données à Madagascar, notre contribution vise à prolonger ces perspectives de recherche en examinant trois arguments qui affectent les possibilités d'automatisation dans la fabrication des modèles d'IA. Premièrement, la complexité des processus d'automatisation et des produits et services basés sur l'automatisation limite les possibilités d'automation complète et conduit donc à un maintien de travailleurs humains dans le processus de production. Ensuite, dans leurs stratégies de maximisation des profits, les entreprises d'IA s'appuient sur des chaines de valeur postcoloniales, dans lesquelles les inégalités historiques ne sont pas seulement perpétuées, mais activement reconfigurées pour répondre aux besoins croissants du secteur (Le Ludec, 2024 ; Cornet, 2026). *In fine*, l'impact central de l'automatisation algorithmique porte davantage sur la précarisation des travailleurs tant le marché du travail de l'IA reste structuré par des institutions héritées de l'époque coloniale, qui continuent de définir les hiérarchies de compétences, les flux de main-d'œuvre et les dépendances économiques (Mezzadra et Neilson, 2013).

1. UNE ENQUÊTE ENTRE PARIS ET ANTANANARIVO : RENDRE VISIBLES LES CHAINES DE PRODUCTION DE L'IA

Cet article s'appuie sur une enquête consacrée aux chaines de production des systèmes d'intelligence artificielle (SIA) développés par des entreprises françaises. L'approche retenue analyse la production de l'IA comme un processus organisationnel distribué, impliquant une pluralité d'acteurs situés à différents niveaux de la sous-traitance et articulant plusieurs sites d'enquête.

Le choix des terrains s'inscrit dans le prolongement de la littérature sur l'externalisation des services numériques, qui met en évidence la structuration transnationale de ces activités selon des logiques Nord/Sud, mais aussi linguistiques et culturelles (Mitter et Pearson 1992 ; Coe, 1997 ; Huws, 2014 ; Graham et al., 2017). Plusieurs travaux montrent que la maitrise de langues comme l'anglais et le français constitue un facteur déterminant dans l'implantation des services numériques externalisés, favorisant des formes d'externalisation fondées sur des proximités linguistiques, culturelles et historiques plutôt que sur les seuls différentiels de couts (Gonzalez et al., 2006 ; Gefen et Carmel, 2008).

Partant de ces hypothèses, nous avons analysé les données de l'Organisation mondiale du commerce concernant les exportations de services informatiques de 16 pays d'Afrique francophone et de l'océan Indien, utilisées comme indicateur de l'implantation d'entreprises de sous-traitance numérique. Pour 2023, ces données mettent en évidence le rôle central de Madagascar dans l'externalisation francophone des services numériques (145 millions de dollars d'exportation) derrière le Maroc (1159 millions de dollars) et la Tunisie (378 millions de dollars)

Entre février 2020 et novembre 2021, une première phase de l'enquête est venue confirmer cette perspective empirique. Elle a consisté en 25 entretiens menés auprès de salariés et dirigeants de 14 entreprises développant des projets d'IA en France, afin de documenter les usages du travail des données et les modalités d'externalisation mises en œuvre. Cette phase a montré que, contrairement aux attentes issues de la littérature sur le micro-travail, les entreprises étudiées recouraient majoritairement à des prestataires formels situés à Madagascar, plutôt qu'à des plateformes fonctionnant comme des marchés ouverts, telles qu'Amazon Mechanical Turk. L'accès aux travailleurs des données a dès lors nécessité une négociation explicite avec les entreprises de sous-traitance et leurs clients.

La seconde phase de l'enquête a été menée à Madagascar sous la forme d'enquêtes réalisées sur site au sein d'entreprises de sous-traitance du travail des données. Au total, 147 annotateurs des données ont été interrogés, ainsi que 17 cadres, managers et dirigeants appartenant à sept entreprises. Les entretiens, d'une durée moyenne d'environ une heure, portaient sur les tâches quotidiennes, les dispositifs organisationnels et techniques encadrant le travail, la gestion des erreurs et des incertitudes, ainsi que sur les trajectoires professionnelles et les conditions d'emploi. Cette phase qualitative a été complétée par la diffusion d'un questionnaire auprès de 294 travailleurs, visant à documenter leurs caractéristiques socio-démographiques et leurs conditions de travail.

Si l'enquête a été menée en partie à distance, elle repose avant tout sur un terrain de six semaines réalisé à Antananarivo, qui a permis de saisir l'*hypervisibilité* du secteur de l'externalisation du travail des données à Madagascar (Chandhiramowuli et al., à paraitre). Loin d'être un secteur marginal ou invisible, l'annotation de données apparait comme une activité largement diffusée au sein de la jeunesse urbaine éduquée, mais aussi parmi les expatriés français impliqués dans les entreprises de sous-traitance. Cette visibilité s'observe dans des situations ordinaires : lors de notre séjour à Antananarivo, nous rencontrons ainsi une cheffe cuisinière malgache, figure médiatique reconnue pour ses émissions consacrées à la cuisine malagasy, qui raconte avoir réalisé des tâches de rédaction de contenus en ligne afin de compenser la fermeture de son restaurant durant la crise sanitaire liée au Covid-19. Cet exemple invite à considérer le travail des données comme une activité socialement banalisée. De manière similaire, l'accès au terrain s'est parfois opéré par des canaux ordinaires : l'entrée dans l'une des entreprises enquêtées s'est ainsi faite par l'intermédiaire d'un directeur de centre d'appels français récemment recruté, rencontré de manière informelle dans l'hôtel où nous résidions, confirmant le caractère localement visible, socialement diffus et quotidiennement accessible de ce secteur dans la capitale malgache.

2. L'ALLONGEMENT DES CHAINES DE PRODUCTION DU NUMÉRIQUE

La production de données pour l'IA s'inscrit dans un secteur présent de longue date sur l'ile, et qui accompagne le développement de l'économie numérique, celui de la production de « services » aux entreprises, appelé sur place secteur de « l'externalisation ». Depuis la fin des années 90, des sociétés, généralement fondées par des Français, proposent aux entreprises un ensemble de services « délocalisables » ayant trait à la numérisation de l'administration des entreprises. De nombreuses entreprises françaises délèguent progressivement aux entreprises de sous-traitance (BPOs) malgaches un large éventail de tâches liées entre autres à la numérisation de documents administratifs (factures, fichiers clients, chèques, etc.), au support client (notamment par téléphone), à la rédaction et à la modération d'espaces en ligne (sections commentaires des sites d'e-commerce par exemple). L'activité des entreprises occidentales s'appuie progressivement sur ces chaines de travail de bureau externalisées, qui permettent la production et la maintenance à bas cout de services numériques.

Ce secteur constitue de plus un axe de développement potentiel majeur pour l'ile. En effet, si les BPOs ne représentent qu'une fraction du chiffre d'affaires annuel réalisé par les acteurs du numérique et des télécommunications sur l'ile (en 2020, ces entreprises réalisent 17 % des 365 millions d'euros dégagés par le secteur au total, soit moins de 1 % du PIB), le secteur de l'externalisation est identifié comme l'un des plus dynamiques de l'ile (Direction générale du Trésor, 2021), et comme offrant des perspectives de croissance importante. Le nombre d'entreprises actives dans le secteur, par exemple, a plus que doublé entre 2016 et 2020, passant de 287 à 656 (Kacenelenbogen et al., 2020) selon des données fournies par l'Economic Development Board of Madagascar (EDBM). Le développement Web et de sites d'e-commerce assure la majeure partie de cette hausse d'activité, mais la BPO (qui regroupe l'annotation, la saisie, la modération et les centres d'appels) décolle fortement depuis 2011. De plus, selon François Kacenelenbogen et ses collègues, le secteur dans sa totalité représente 22 600 emplois formels directs sur l'ile, soit 3,4 % des emplois formels au niveau national. Les employés des BPOs composent la grande majorité de ces emplois (Kacenelenbogen et al., 2020). Dans ce contexte, les emplois proposés par les entreprises d'externalisation aux jeunes Malgaches diplômés constituent l'une des rares entrées disponibles dans l'économie formelle.

Ceux-ci sont ainsi intégrés aux lignes de production des entreprises technologiques occidentales. Une start-up parisienne qui voudrait entrainer un modèle de vision assistée par ordinateur pourrait ainsi disposer de moins de 20 employés à Paris, *data scientists*, développeurs et personnel support, mais de plusieurs dizaines d'annotateurs au sein d'une BPO à Madagascar. Ces derniers sont ainsi virtuellement rattachés à l'entreprise cliente pour la durée du projet. En effet, ils n'opèrent généralement que sur un projet à la fois, la durée de celui-ci pouvant varier de quelques semaines à plusieurs années. Ils sont encadrés sur place par un chef de projet, qui prend la plupart du temps en charge le relais des instructions client et la formation des travailleurs, et parfois par un employé dédié au sein de l'entreprise cliente, qui suit leur activité au quotidien. Cette structure de production complexifie grandement le suivi et la transparence tout au long des chaines de production du numérique.

En effet, un SIA peut parfois mobiliser trois ou quatre niveaux de production distincts, sans qu'aucun des acteurs n'ait de visibilité globale sur l'ensemble de la chaine. Une entreprise française peut par exemple sous-traiter la production d'un SIA à une entreprise innovante, qui elle-même délègue l'annotation nécessaire à l'entrainement des modèles à une BPO située à Madagascar. En fonction de la force de travail disponible, l'entreprise de BPO peut ensuite décider de ré-externaliser la production des données auprès d'une entreprise tierce, ou d'un réseau de travailleurs informels, complexifiant considérablement le suivi du process de production des données d'entrainement pour l'IA. Ce suivi est d'autant plus important que, nous allons le voir, les annotateurs font partie intégrante de la production des SIA tout au long de leur cycle de vie, et non pas seulement durant les phases préliminaires d'entrainement des modèles.

3. QU'EST-CE QUE L'ANNOTATION ? UN TRAVAIL COGNITIF AU CŒUR DU DÉVELOPPEMENT DE L'IA

L'analyse de l'annotation ne saurait se limiter à une description de l'organisation du micro-travail ou aux dispositifs de sous-traitance qui l'encadrent. Les catégories de « micro-travailleurs » ou de « travailleurs des données » renvoient respectivement à une définition fondée sur l'organisation du travail par les plateformes et à une approche issue du point de vue des entreprises d'annotation. Or, l'enquête montre que le travail d'annotation mobilise des qualifications bien plus larges que ne le suggèrent ces catégories. Les annotateurs disposent le plus souvent d'une formation universitaire (75 %) et acquièrent, au sein des entreprises, des compétences spécifiques liées aux projets, aux outils et aux exigences de qualité. L'annotation recouvre en outre une pluralité de postes et de responsabilités, certains jouant un rôle central dans la gestion des doutes, des erreurs et du contrôle qualité.

Cette dimension apparait de manière particulièrement nette dans un projet de reconnaissance d'images de plateaux-repas destiné à une caisse de cantine automatisée. Les annotateurs décrivent un travail de classification très détaillé, avec « *plus de 3 000 ingrédients à mémoriser* » (Pota, annotatrice de données) et plusieurs niveaux d'annotation :

> « *(La classification est) très très détaillée, parce que tout ce qui est dans l'assiette, on doit le classifier.* »
> *(Naivo, annotateur)*

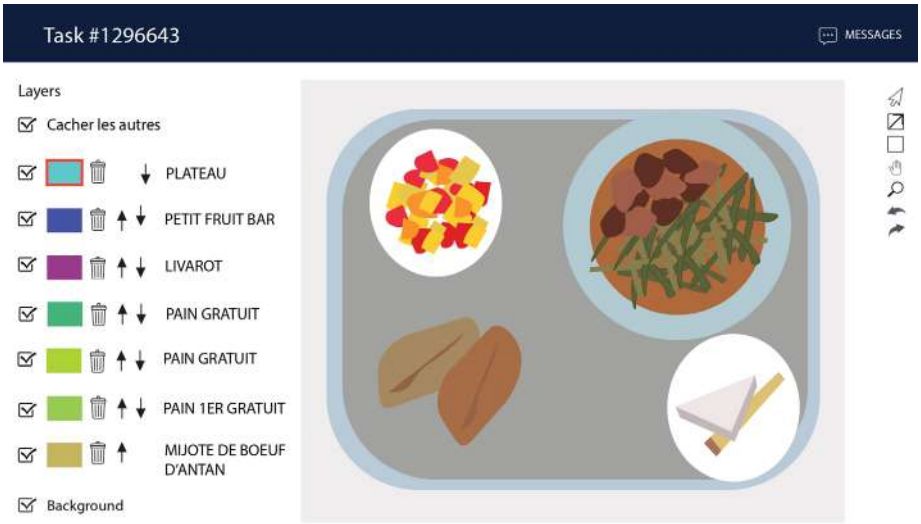


Image n° 1. Représentation graphique d'une interface d'annotation. L'image montre un plateau de cantine comportant plusieurs éléments auxquels sont associés différents labels sur la gauche : « assiette », « petit fruit », « pain en libre-service », « ragoût de bœuf traditionnel ». © Aurore Le Ludec (Le Ludec, 2024), cité dans Wirth et al., 2025.

Cette activité ne peut pas davantage être comprise comme un simple « travail du doigt », car elle implique un travail cognitif soutenu, nécessaire à la catégorisation d'éléments présents dans des données hétérogènes et à la gestion de situations imprévues. Les annotateurs sont directement confrontés à ce que l'on peut qualifier de « dérive conceptuelle » des modèles, c'est-à-dire à l'écart entre les catégories définies en amont et les données effectivement rencontrées. Ils jouent alors un rôle actif dans la reformulation des problèmes posés aux modèles, notamment à travers des dispositifs collectifs de remontée d'incertitudes :

> « *Oui (ça nous arrive de faire remonter.) (…) lorsqu'on avait des questions ou qu'on trouvait des nouvelles choses on le mettait dans ce document.* »
> *(Soly, annotateur)*

Cette complexité suppose la mobilisation de nombreux supports de formation et de référentiels, élaborés conjointement par les équipes en France et à Madagascar. Ces supports restent toutefois incomplets et doivent être continuellement enrichis à partir des cas rencontrés. Loin d'un travail routinier, l'annotation apparait ainsi comme une activité collective de classification et de « problématisation », distribuée le long de la chaine de valeur (Le Ludec et al., 2023). Cette analyse conduit à considérer les frontières entre modélisation et annotation comme largement poreuses : à plusieurs niveaux, annotateurs, chefs de projet, ingénieurs et clients participent conjointement à un même travail sur les données, au-delà du seul encerclement d'objets sur des images. À ce titre, l'annotation ne consiste pas uniquement à appliquer un système de classification existant, mais aussi à en éprouver les limites et à le compléter, en rendant visibles des éléments laissés dans l'ombre (Bowker et Star, 1999).

Ainsi, le rôle des annotateurs ne se cantonne pas uniquement aux phases d'entrainement des modèles. En effet, ceux-ci interviennent dans l'ensemble des phases de production, de suivi et de maintenance. Ces derniers prennent aussi en charge la vérification des sorties et prédictions du modèle – pour s'assurer que celui-ci fonctionne en accord avec les besoins du service – et réalisent parfois les tâches que ce dernier ne peut accomplir (Tubaro et al., 2020). C'est par exemple le cas lorsque des annotateurs rencontrés à Antananarivo se retrouvent à identifier les voleurs potentiels dans des supermarchés occidentaux « à la place » du système d'IA. Ce cas s'inscrit dans un phénomène plus large : dès 2019, près de 40 % des start-up européennes d'IA déclaraient des systèmes pour lesquels aucune preuve fonctionnelle n'avait pu être établie (Vincent, 2019).

Ils accompagnent ainsi le modèle tout au long de sa durée de vie. Ils participent également à la redéfinition constante et cyclique des catégories qui composent la représentation du monde réel ou « *ground truth* » (« vérité terrain ») que cherchent à reproduire les SIA (Le Ludec et al., 2023 ; Girard-Chanudet, 2025). En effet, les modèles sont peu adaptés au traitement de situations nécessitant un degré de contextualisation élevé pour opérer une catégorisation. Un modèle opérant une remontée automatique d'alertes de vol à l'étalage à partir d'images de caméra de surveillance aura par exemple un taux d'erreur beaucoup plus important qu'un modèle de lecture automatique de plaque d'immatriculation. Dans le premier cas, la définition de la catégorie dépend d'une foule de facteurs particulièrement complexes : est-ce qu'une volonté de dissimulation de produit constitue un vol ? Est-ce qu'un placement de produit dans un sac à dos constitue une dissimulation ? Etc. Dans le second, l'action du modèle repose sur la reconnaissance d'un ensemble de signes sur un support extrêmement standardisé : les plaques minéralogiques des véhicules. En fonction des cas d'usage, un nombre élevé de travailleurs des données présents en continu peut donc être nécessaire pour surveiller l'action du modèle et, par extension, la conformité de la production avec le service proposé. Dans le cas de situations peu standardisables, on peut affirmer avec un assez haut degré de confiance que ces travailleurs ne sont pas voués à disparaitre.

— 4. DES CHAINES GLOBALES DE PRÉCARITÉ ?

Cette dépendance des entreprises technologiques à un travail externalisé à faible cout pose des questions, en particulier si l'on s'intéresse au contexte local dans lequel évoluent les travailleurs. En effet, pour ces travailleurs jeunes, urbains et éduqués, l'entrée dans le secteur de l'externalisation ne se fait pas par choix. Ils décrivent majoritairement des conditions de travail assez précaires. Et il n'est pas rare que, lorsqu'un projet s'arrête, une grande partie des travailleurs opérant sur ce dernier soient licenciés.

Le métier est par ailleurs décrit comme pénible et répétitif, avec peu de perspectives d'évolution de carrière. Le salaire perçu, entre 80 et 120 euros avec une médiane à 98 euros au sein de notre échantillon, s'il est supérieur au salaire mensuel minimum malgache (50 euros), est en moyenne inférieur au salaire perçu par les travailleurs formels ayant effectué un passage par l'université en population générale, qui est de 145 euros en 2022 (INSTAT, 2024). Il en résulte un sentiment de carrière empêchée, où les travailleurs nous décrivent un choix par défaut, face à l'impossibilité de pouvoir exercer dans le secteur pour lequel ils ont été formés.

> « *Mon rêve, c'est de devenir avocat, parce que j'ai étudié le droit. D'avoir un vrai job, une vraie profession. (À Madagascar), nous utilisons un argot, nous appelons (l'annotation de données) : "sérum", parce que c'est (simplement) un travail qui vous empêche de mourir.* »
> *(Tiary, annotateur)*

L'analyse des chaines de production de l'IA conduit à s'interroger sur la répartition de la valeur produite par le travail d'annotation. Dans le cadre de négociations salariales entre une start-up d'IA et l'un de ses sous-traitants, un employé de la start-up nous a donné accès à des éléments détaillant la structure des couts salariaux de l'entreprise de sous-traitance. Ces données offrent un aperçu rare des mécanismes d'appropriation de la valeur au sein du secteur.

Au moment de l'enquête, cette entreprise de taille moyenne emploie 110 salariés sous contrat de prestation de services et facture ses services 3,30 euros par heure et par agent, soit environ 530 euros par mois et par agent. Sur cette base, le chiffre d'affaires annuel peut être estimé entre 650 000 et 700 000 euros, pour des bénéfices nets compris entre 390 000 et 440 000 euros. Les couts fixes apparaissent limités : les locaux appartiennent au dirigeant, et les principaux postes de dépenses concernent la connexion Internet (environ 165 euros par mois pour 6 téraoctets de données et un débit de 100 Mbits par seconde), l'entretien des locaux et le matériel informatique. Malgré ces couts, l'activité se révèle particulièrement rémunératrice pour les propriétaires de l'entreprise.

Ce cas correspond toutefois à une configuration spécifique, caractérisée par un statut informel, c'est-à-dire qui opère avec des contrats de travailleurs indépendants non déclarés et une taille intermédiaire. Dans le cas d'une autre entreprise, cette fois formelle et de plus grande taille, un annotateur indique que la prestation est facturée 75 euros par jour et par agent, soit environ 1 500 euros par mois pour 20 jours ouvrés. Ce tarif varie selon les compétences requises par les projets d'IA. Bien que cette entreprise supporte des couts plus élevés – notamment la location de locaux importants à Antananarivo pour environ 350 postes –, les marges demeurent significatives, comme le souligne l'un de ses clients :

> « *Par contre, j'en suis convaincu, des boîtes comme ça (…) ils te proposent un budget journalier qui est en général 20 fois supérieur à ce que la personne va réellement toucher (…) et comme tout truc de vente de service, ils font 80 % de profit, donc ce truc-là est complètement disproportionné.* »
> *(Cédric, Lead Data Team, AIMedic)*

Les données recueillies permettent également de situer les entreprises d'IA au sein de cette chaine de valeur. En 2021, une employée indique que la start-up Automatik a facturé environ 200 000 euros à l'un de ses clients pour le développement d'une caisse de cantine automatisée. Dans le même temps, les documents transmis par le sous-traitant montrent qu'Automatik débourse environ 100 000 euros par an pour le travail d'annotation. Les 19 annotateurs affectés au projet ne travaillent cependant pas à temps plein sur celui-ci ; les entretiens permettent d'estimer qu'ils y consacrent environ un quart de leur temps, soit un cout d'environ 25 000 euros pour ce projet spécifique. À ces montants s'ajoutent les couts

d'infrastructure, principalement liés à l'hébergement des données et à la puissance de calcul, qu'une employée d'Automatik estime à environ 100.000 euros par an pour l'ensemble des projets de l'entreprise.

Pris ensemble, ces éléments mettent en évidence une forte asymétrie dans la distribution de la valeur le long de la chaine de production de l'IA. Bien que le travail des annotateurs malgaches constitue une composante essentielle du développement des modèles, il demeure faiblement valorisé financièrement. Pour les travailleurs, y compris ceux disposant d'un foyer et d'enfants, les revenus issus de l'annotation sont vécus comme des revenus de subsistance, insuffisants pour couvrir l'ensemble des besoins du ménage ou permettre une réelle accumulation de capital. Atteindre une relative stabilité financière dans le secteur de l'externalisation privée suppose dès lors l'acceptation d'un fort risque de licenciement, facilité par le cadre du droit du travail malgache, ainsi que le développement d'activités secondaires destinées à compléter le revenu principal. Entre « travail d'à-côté » (Weber, 2009) et activités entrepreneuriales « reproductives » (Wilson et Yochim, 2017), ces pratiques constituent des stratégies à la fois professionnelles et domestiques de gestion de la précarité.

Pendant ce temps, les entrepreneurs de la sous-traitance, les entreprises d'IA et les grandes entreprises du numérique fournissant les infrastructures captent l'essentiel de la valeur produite. Cette configuration renvoie plus largement à un modèle économique désormais central dans le secteur de l'IA. Une illustration paradigmatique est celle de Scale AI, start-up américaine spécialisée dans l'annotation et la préparation de données pour l'entrainement des modèles d'IA, dans laquelle Meta Platforms a investi environ 14 milliards de dollars pour acquérir une participation de 49 %, valorisant l'entreprise à près de 29 milliards de dollars. Ces sommes, colossales par rapport aux revenus des annotateurs, révèlent une captation de la valeur à plusieurs échelons de la chaine, où des acteurs globalisés accumulent des ressources importantes pour des activités reposant massivement sur du travail humain externalisé.

L'analyse des chaines de production de l'IA entre la France et Madagascar permet de déplacer le regard porté sur les effets de l'automatisation. Loin de conduire à une disparition du travail humain, le développement des systèmes d'IA repose sur une redistribution du travail à l'échelle globale, marquée par l'allongement des chaines de sous-traitance et par la centralité du travail d'annotation, de vérification et de maintenance des modèles.

Cette configuration s'accompagne d'une reconfiguration des hiérarchies du travail à l'échelle internationale. Tandis que les activités les plus visibles et valorisées se concentrent dans les pays du Nord global, les tâches les plus intensives en travail, mais aussi les plus précaires, sont externalisées vers des économies du Sud global. Ce processus ne se limite pas à une simple division technique du travail : il s'inscrit dans des dynamiques historiques et institutionnelles plus larges, qui contribuent à reproduire, voire à reconfigurer, des rapports de dépendance hérités.

Les résultats mettent également en évidence une forte asymétrie dans la distribution de la valeur produite le long de ces chaines. Malgré leur rôle central dans le développement des systèmes d'IA, les travailleurs des données bénéficient d'une faible reconnaissance économique et professionnelle. Cette situation alimente des trajectoires marquées par l'instabilité, la faiblesse des revenus et des perspectives limitées d'évolution, contribuant à ancrer ces activités dans des formes durables de précarisation.

Au-delà du cas étudié, ces résultats invitent à repenser les effets de l'intelligence artificielle sur le travail en articulant analyse des systèmes techniques et analyse des organisations productives qui les rendent possibles. Ils suggèrent que l'automatisation algorithmique ne peut être comprise indépendamment des conditions sociales de sa production, et qu'elle tend moins à supprimer le travail qu'à en transformer la répartition, les formes et la visibilité.

RÉFÉRENCES BIBLIOGRAPHIQUES

— Acemoglu, D. et Restrepo, P. (2019), « Automation and New Tasks: How Technology Displaces and Reinstates Labor », in *Journal of Economic Perspectives*, 33(2), 3-30. https://doi.org/10.1257/jep.33.2.3

— Anwar, M.A. et Graham, M. (2020), « Digital labour at economic margins: African workers and the global information economy », in *Review of African Political Economy*, 47(163), 95-105. https://doi.org/10.1080/03056244.2020.1728243

— Benedetto-Meyer, M. et Boboc, A. (2021), *Sociologie du numérique au travail*, Malakoff, Armand Colin, 240 p.

— Bowker G.C., Star S.L. (1999), *Sorting things out. Classification and its consequences*, Cambridge, MA, MIT Press.

— Brynjolfsson, E., Rock, D. et Syverson, C. (2017), « Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics » (No. W24001; p. w24001), National Bureau of Economic Research. https://doi.org/10.3386/w24001

— Chandhiramowuli, S., Cornet, M., Le Ludec, C., Torres Cierpe, J., Vogiatzis, I., Wirth, A., Yang, T. et Zefeng, B. (2026), « Looking at "ghost work": Revisiting and interrupting the invisibility in data work », in *Handbook of Critical Data Studies*.

— Coe, N. (1997), « US Transnationals and the Irish Software Industry: Assessing the Nature, Quality and Stability of a New Wave of Foreign Direct Investment », in *European Urban and Regional Studies*, 4(3), 211-230. https://doi.org/10.1177/096977649700400302

— Cornet, M. (2026), « Travail de la donnée et données du travail: Gestion de l'incertitude et délégation des risques dans les chaînes de production de l'Intelligence Artificielle » (thèse, Institut polytechnique de Paris). https://theses.fr/s253434?trk=public_post_comment-text

— Frey, C.B. et Osborne, M.A. (2017), « The future of employment: How susceptible are jobs to computerisation? », in *Technological Forecasting and Social Change*, 114, 254-280. https://doi.org/10.1016/j.techfore.2016.08.019

— Gefen, D. et Carmel, E. (2008), « Is the World Really Flat? A Look at Offshoring at an Online Programming Marketplace », in *MIS Quarterly*, 32(2), 367-384. JSTOR.

— Girard-Chanudet, C. (2025), « Ground-truth is law: The invisible conceptual work behind AI », in *Big Data & Society*, 12(2), 20539517251352823.

— Gonzalez, R., Gasco, J. et Llopis, J. (2006), « Information systems offshore outsourcing: A descriptive analysis », in *Industrial Management & Data Systems*, 106(9), 1233-1248. https://doi.org/10.1108/02635570610712555

— Graham, M., Hjorth, I. et Lehdonvirta, V. (2017), « Digital labour and development: Impacts of global digital labour platforms and the gig economy on worker livelihoods », in *Transfer: European Review of Labour and Research*, 23(2), 135-162. https://doi.org/10.1177/1024258916687250

— Gray, M.L. et Suri, S. (2019), *Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass*, Boston, Houghton Mifflin Harcourt, 288 p.

— Huws, U. (2014), « Labor in the global digital economy: The cybertariat comes of age », in *Monthly Review Press*.

— INSTAT (2024), « EPM | Enquête Permanente auprès des Ménages 2021-2022 » (p. 426), Institut national de la statistique. https://www.instat.mg/p/epm-enquete-permanente-aupres-des-menages-2021-2022

— Kacenelenbogen, F., Martin Meier, C., Rahaingoson, I. et Reboza, T. (2020), « Madagascar, Écosystème numérique : Cartographie et propositions de mesures d'atténuation des effets du COVID-19 et d'appui au secteur » (p. 175), TPSDE Facility.

— Krzywdzinski, M. (2022), « Toward a Socioeconomic Company-Level Theory of Automation at Work », in *Weizenbaum Journal of the Digital Society*, 2(1), Article 1. https://doi.org/10.34669/wi.wjds/2.1.5
— Le Ludec, C. (2024), « Des humains derrière l'intelligence artificielle. La sous-traitance du travail de la donnée entre la France et Madagascar » (thèse de doctorat, Institut polytechnique de Paris). https://theses.fr/2024IPPAT001
— Le Ludec, C. et Cornet, M. (2023), « Derrière l'IA, le travail précaire au Sud », in *The Conversation*, 26.
— Le Ludec, C., Cornet, M. et Casilli, A.A. (2023), « The problem with annotation. Human labour and outsourcing between France and Madagascar », in *Big Data & Society*, 10(2), 20539517231188723. https://doi.org/10.1177/20539517231188723
— Mezzadra, S. et Neilson, B. (2013), *Border as method, or, the multiplication of labor*, Durham, Duke University Press, 384 p.
— Mitter, S. et Pearson, R. (1992), « Global information processing: The emergence of software services and data entry jobs in selected developing countries », ILO Working Papers (No. 992897713402676 ; ILO Working Papers), International Labour Organization. https://ideas.repec.org/p/ilo/ilowps/992897713402676.html
— Tubaro, P., Casilli, A.A., Cornet, M., Le Ludec, C. et Torres Cierpe, J. (2025), « Where does AI come from? A global case study across Europe, Africa, and Latin America », in *New Political Economy*, 30(3), 359-372. https://doi.org/10.1080/13563467.2025.2462137
— Tubaro, P., Casilli, A.A. et Coville, M. (2020), « The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence », in *Big Data & Society* (Sage UK : London, England). https://doi.org/10.1177/2053951720919776
— Vincent, J. (2019), « Forty percent of "AI startups" in Europe don't actually use AI, claims report », in *The Verge*, 5 mars. https://www.theverge.com/2019/3/5/18251326/ai-startups-europe-fake-40-percent-mmc-report
— Weber, F. (2009), *Le travail à-côté: Une ethnographie des perceptions* (nouvelle éd. revue et augm.), Aubervilliers, Éditions de l'École des hautes études en sciences sociales, 238 p.
— Wilson, J.A. et Yochim, E.C. (2017), *Mothering through Precarity: Women's Work and Digital Media*, Durham, Duke University Press, 232 p.
— Wirth, A., Le Ludec, C., Girard-Chanudet, C. et Cornet, M. (2025), « Who Decides When Artificial Intelligence Works? Algorithmic Production as a Continuous Evaluation Process », in *Digital Society*, 4(3), 65. https://doi.org/10.1007/s44206-025-00224-4

L'IA GÉNÉRATIVE AU TRAVAIL : QUELLES TRANSFORMATIONS À L'HORIZON ?

Vincent Mandinaud (Anact), **Yann Ferguson** (Inria), **Aída Ponce del Castillo** (ETUI)

À partir d'une revue critique des travaux récents et de l'analyse des trajectoires actuelles de déploiement de la GenAI, cet article distingue les initiatives portées par les employeurs des usages, souvent informels, développés par les salariés. Il montre que les promesses de gains de productivité et d'amélioration des conditions de travail restent largement conditionnées aux processus d'appropriation, aux choix organisationnels et aux formes de dialogue qui accompagnent l'introduction de ces technologies. Les usages individuels révèlent un potentiel d'autonomisation et d'innovation, mais soulèvent également des enjeux de compétences, de reconnaissance, de coopération et de qualité du travail. L'article défend ainsi l'idée qu'un dialogue social technologique, articulé au travail réel, constitue une condition essentielle pour orienter les usages de l'IA vers des transformations favorables aux travailleuses et travailleurs comme aux organisations.

Mots-clés : intelligence artificielle générative, transformation du travail, dialogue social technologique, conditions de travail

— INTRODUCTION : DE LA DISPARITION DES EMPLOIS À LA TRANSFORMATION DU TRAVAIL

Aujourd'hui, des systèmes d'intelligence artificielle (SIA) peuvent orchestrer et contrôler les tâches, mesurer les performances ou encore appuyer certaines décisions relatives à la gestion du personnel. L'IA générative (GenAI) s'inscrit dans cette dynamique de transformation du travail, où elle est amenée à jouer un rôle compte tenu de sa facilité d'utilisation, de la variété de fonctions qu'elle peut remplir, et de services qu'elle peut rendre. Il reste à savoir si cette technologie et ses usages peuvent améliorer la situation des travailleurs, et, le cas échéant, à quelles conditions juridiques, organisationnelles et sociales.

La littérature récente converge sur un point essentiel : la GenAI est susceptible de transformer en profondeur les activités et les relations de travail, même si ses effets macro-économiques demeurent incertains. Si certains emplois pourraient être automatisés et donc disparaitre, la majeure partie des

emplois devraient voir leurs tâches, leurs conditions de réalisation et leurs modalités d'accomplissement pratique considérablement transformées. De plus, la majorité des études qui se penchent sur les effets de la GenAI sur la croissance économique soulignent qu'elle a le potentiel d'améliorer la productivité dans certains secteurs et/ou métiers. Les appréciations varient toutefois quant à la mesure de cette augmentation et de ses effets sur le PIB.

Les simulations du FMI réalisées par Cerutti et al. (2025) illustrent deux trajectoires contrastées : un scénario d'adoption et de diffusion rapide et un autre plus lent. Dans le premier scénario, la productivité mondiale progresserait de 2,4 % en 10 ans, entrainant une hausse du PIB mondial de 4 % par rapport à la trajectoire sans IA. Dans le second, la progression de la productivité et celle du PIB seraient limitées à l'horizon de 10 ans à seulement 1,8 % et environ 1,3 % respectivement, en raison d'une adoption partielle. Néanmoins, Simons et al. (2024) insiste sur le fait que les gains de productivité ne sont pas automatiques mais dépendent de la réorganisation des processus de production, de l'adaptation des compétences et du renouvellement du capital.

Acemoglu et al. (2024) adopte quant à lui une position plus prudente. En modélisant l'IA à travers un cadre basé sur les tâches aux États-Unis, les auteurs estiment que la contribution de l'IA aux gains de productivité pourrait ne pas dépasser 0,7 point de pourcentage sur 10 ans. L'impact sur le PIB, bien que légèrement supérieur en raison des effets induits (investissement, consommation), resterait modeste, à 1,1 %. Ces estimations suggèrent que l'IA pourrait ne produire qu'un effet positif incrémental sur la croissance annuelle, à moins d'un changement structurel qui porterait notamment la création de nouveaux types d'activités. Acemoglu et al. (2024) explique par ailleurs que si l'IA d'aujourd'hui peut améliorer la productivité moyenne, elle risque de remplacer de nombreux travailleurs et de dégrader la qualité de l'emploi de ceux qui restent. Sous cet angle, elle évolue dans une mauvaise direction parce que « nous l'utilisons trop pour automatiser et pas assez pour apporter de l'expertise et des informations aux travailleurs ». L'IA ne deviendra sans doute pas, miraculeusement, une technologie qui agit « en complémentarité » avec les travailleurs, sauf si l'industrie consacre beaucoup de temps et d'énergie à la faire évoluer dans ce sens (Dizikes, 2024). Acemoglu, Autor et Johnson (2026) prolonge cette analyse en plaidant pour une « pro-worker AI », c'est-à-dire une IA conçue pour renforcer les capacités des travailleurs, compléter leur expertise et favoriser la création de nouvelles tâches, plutôt que de poursuivre principalement des stratégies d'automatisation.

Les effets de la GenAI varient fortement selon les usages retenus, les contextes organisationnels et les catégories de travailleurs concernés. Les experts soulignent des variabilités non seulement sur ce que la GenAI vient faire (automatiser ou assister, compléter ou augmenter), mais aussi sur la répartition des avantages et des inconvénients entre les travailleurs. Ainsi, Goldman Sachs Research (Briggs et al., 2023) estime que deux tiers des métiers aux États-Unis pourraient être exposés à une certaine forme d'automatisation par l'IA, et que la proportion de travail potentiellement remplaçable par l'IA varie de 25 % à 50 %. Cependant, l'OIT souligne que la GenAI devrait principalement transformer les activités et moins s'y substituer, concerner principalement les pays à haut niveau de revenu élevé en raison d'une plus grande part d'emploi de bureau, et en conséquence impacter prioritairement les femmes qui occupent une part importante de ces emplois (Berg et al., 2025)[1]. Des travaux du MIT (Brynjolfsson et al., 2023), en se concentrant sur le secteur du service client, constatent que le recours à la GenAI a des impacts assez inégaux sur les travailleurs selon qu'ils sont novices ou expérimentés. Ainsi, l'usage de l'IA augmente sensiblement la productivité des travailleurs moins qualifiés et expérimentés, mais a des effets minimes sur les travailleurs plus qualifiés ou expérimentés, pouvant même contribuer à une dégradation de la qualité du travail et des compétences, ou de la réflexion critique. Des résultats

1 Voir l'article dans ce dossier.

similaires sont constatés dans une étude d'Harvard (Dell'Acqua et al., 2023) à propos de consultants, dès lors que l'IA sait faire. Mais lorsqu'elle ne sait pas faire, ceux qui l'utilisent réussissent moins bien à produire des conseils avisés.

Les impacts de la GenAI au travail dépassent ainsi largement la seule question de la productivité. En revanche, son intégration formelle dans les organisations emprunte des trajectoires hétérogènes, marquées par des stratégies d'entreprise, des expérimentations ponctuelles et des usages parfois non institués. L'article propose d'éclairer les façons dont se déploie la GenAI dans les entreprises, en distinguant les initiatives des employeurs et celles des employés. Ce faisant, il interroge à la fois les modalités et les finalités de son introduction, mais aussi les pratiques plus ou moins instituées de dialogue qui participent à la régulation de ses usages, et leur contribution à l'amélioration des conditions de travail et à l'amélioration des changements technologiques et organisationnels.

1. RETOURS SUR DES INITIATIVES D'EMPLOYEURS : DES PROMESSES D'IA ENCORE PEU TENUES DANS LA PRATIQUE

Si, pour la majorité des utilisateurs, la sortie mondiale de l'agent conversationnel ChatGPT, fondé sur un Large Language Model (LLM), constitue l'an 1 de l'intelligence artificielle, les expérimentations de l'IA d'apprentissage à base de données datent du mitan du XXe siècle et se sont multipliées depuis les années 2010. Cette méthode marquerait en effet un « *point d'inflexion* » vers ce qu'Erik Brynjolfsson et Andrew McAfee (2015) appellent le « *second âge de la machine* » permettant une automatisation croissante de tâches cognitives. Cette forme d'IA serait en mesure de briser le fameux « paradoxe de Polanyi », considéré comme le plafond de verre de l'automatisation de l'intelligence des travailleurs, dont la part tacite augmente avec l'expertise. Les travailleurs déploient en effet des arts de faire, une « *invention du quotidien* », du « *braconnage* » (Certeau, 1990), qui tendent à s'accumuler avec l'expérience, tout en s'internalisant et s'automatisant car ils sont incorporés (Nonaka et von Krogh, 2009). Mais l'IA d'apprentissage à base de données, sous réserve d'avoir été convenablement alimentée et entrainée (Casilli, 2019), serait en capacité de développer sa propre empirie. Autrement dit, il ne serait plus nécessaire de formaliser les connaissances et les pratiques professionnelles : grâce aux données, la GenAI élaborerait ses tactiques particulières fondées sur des « *patterns* » et le traitement de données provenant de très grandes bases de données.

La représentation dominante de la GenAI au travail à l'œuvre dans les cas d'usage à l'initiative d'employeurs demeure relativement classique. On la retrouve par exemple dans le célèbre article de Frey et Osborne (2013), l'étude du Conseil d'orientation pour l'emploi (COE, 2017) ou le rapport de la Mission Villani (2018) : les systèmes d'IA automatiseraient les routines, c'est-à-dire les tâches ne sollicitant pas de capacité d'adaptation, de résolution de problème, d'intelligence sociale, de créativité, etc. En faisant gagner en efficience, la GenAI permettrait des gains de productivité aux organisations, et dans le même temps redonnerait du sens au travail en « *désautomatisant* » le travailleur (Villani et al., 2018), « *en remettant de l'humain dans le travail* » (Malakoff Médéric, 2018), de façon à « *innover, avoir de l'imagination, prendre des risques contrôlés, entreprendre. Toutes ces capacités qui, dans une certaine mesure, sont le symbole de notre humanité* » (Dejoux, 2018, p. 25).

Les organisations pionnières dans l'intégration des systèmes d'IA se sont logées dans cette division du travail : automatisation de tâches répétitives telles que répondre aux questions courantes, trier des CV et repérer des défauts, l'optimisation de processus comme la réduction de la consommation d'eau et d'énergie, ou l'accélération de processus (traitement de dossiers administratifs). Qualifiées de tâches « à faible valeur ajoutée », leur automatisation libérerait du temps pour des activités à « haute

valeur ajoutée », « cœur de métier », proclamées plus épanouissantes par les manageurs. Telle est la narration classique, dans la continuité des discours sur la transformation numérique des entreprises disant remettre « l'humain au centre ».

Pour actualiser et soutenir cette perspective, de nombreuses démarches ont été engagées en faveur d'un cadre éthique ainsi que juridique. Par exemple, en 2019, un groupe d'experts européens de haut niveau (Commission européenne, 2019) a développé des « *lignes directrices en matière d'éthique pour une IA digne de confiance* » sur la base des principes convenus au sein du groupe d'experts. Sept « *exigences essentielles* » ont été posées :

— Le contrôle humain : assurer une prise de décision éclairée.
— La robustesse et la sécurité : maitriser les risques et les incertitudes.
— Le bien-être sociétal et environnemental : assurer la soutenabilité sociale et écologique.
— L'équité et la diversité : garantir la non-discrimination.
— La transparence : expliquer les modèles.
— La responsabilité : maitriser les responsabilités en cas d'incident ou de contentieux.
— Vie privée : respecter l'anonymat des personnes.

Le rapport associe une série de questions que les acteurs d'un projet d'IA, en particulier les concepteurs, doivent se poser. Cette « *éthique à principes* » (Ganascia, 2022) se réfère largement aux droits de l'homme pour élaborer une IA « *responsable* » (OCDE, 2019), « *centrée sur l'humain* ». On constate que ces principes sont souvent repris et affichés dans beaucoup d'initiatives d'employeurs qui affirment ainsi leur caractère responsable tout en formulant une double promesse : gagner en productivité et améliorer la qualité de vie et des conditions de travail.

Ces travaux ont jeté les bases du règlement européen sur l'IA (AI Act) fondé sur une approche par risques, que les employeurs, en tant que fournisseurs ou déployeurs, doivent suivre (Commission européenne, 2024). Les systèmes avec un risque inacceptable sont interdits, sauf exception. Ceux identifiés comme présentant un risque élevé (les SIA destinés à être utilisés pour le recrutement ou la sélection de personnes physiques, et ceux destinés à être utilisés pour prendre des décisions concernant les conditions des relations de travail, la promotion ou la résiliation des relations contractuelles liées au travail) doivent se conformer à des obligations légales : mettre en place un système de gestion des risques, assurer une gouvernance des données de haute qualité, établir une documentation technique pour démontrer la conformité et fournir les informations nécessaires à l'évaluation de cette conformité, etc.

Bien qu'enthousiasmantes pour les entreprises – comme le révélaient le niveau d'investissement (IDC, 2017) ou encore l'initiative de 29 États engagés dans la constitution d'un Partenariat mondial sur l'IA (PMIA) en 2020 –, les promesses tardent à s'actualiser. La majorité des cas d'usage ne passent pas la période d'essai. Les analyses économiques montrent que 85 % environ des *proofs of concept* (PoC) ne sont pas déployées (Weiner, 2021 ; Escobar, McGovern et Morales-Menendez, 2021). Le PMIA, dans son rapport de 2021, s'interrogeait en ce sens : « *Pourquoi les PoC réussis ne sont presque jamais déployés ?* » Les chercheurs et les ingénieurs pointent des problèmes techniques (fiabilité, explicabilité, certifiabilité, qualité des données, etc.), les économistes un déficit de retour sur investissement (ROI), notamment la difficulté de « coster » un projet d'IA dont les itérations et les temps d'apprentissage sont complexes à estimer. Les sciences du travail (sociologie, psychologie, ergonomie) mettent plutôt en avant des enjeux d'appropriation qui ne sont pas suffisamment pris en compte en phase de conception, et insuffisamment ou incorrectement pris en compte lors des phases d'accompagnement (Barcellini et al., 2024).

Le centre de ressources et d'expérimentations sur l'intelligence artificielle dans le milieu professionnel du ministère du Travail et des Solidarités (LaborIA) et Inria (Institut national de recherche en sciences et technologies du numérique), dans une étude longitudinale de plusieurs cas d'usage, pointent « *un conflit de rationalités* » entre une logique gestionnaire de rationalisation du travail à l'œuvre et le travail réel (Borel, Ferguson et Condé, 2024). Autrement dit, les SIA se rapportent encore trop largement au travail prescrit, ce qui génère d'importants décalages entre les recommandations, notifications et autres alertes de l'IA et les situations réellement vécues au travail. Ce même rapport évoque par ailleurs des situations où un « *compromis de rationalités* » semble trouvé, mais au détriment du sens du travail pour le professionnel. L'analyse de ces cas d'usage, initiatives d'employeurs, montre qu'ils sont loin de conduire à des « *configurations capacitantes* ». La mobilisation – et d'autant plus la simple invocation – de principes éthiques n'est pas suffisante en tant que telle.

Des méthodes, des outils et des pratiques effectives de dialogues social et professionnel loyaux et constructifs à l'occasion et tout au long du déploiement de systèmes utilisant l'IA sont nécessaires pour construire des équilibres, compte tenu de la nature évolutive et apprenante de l'IA, mais aussi des organisations dans lesquelles elles s'inscrivent, et des individus et des collectifs qui sont amenés à en faire usage. En France, ces méthodes, outils et pratiques de dialogue social technologique restent à ce stade balbutiants (Mandinaud et Keusch, 2025). Pourtant, les accords-cadres européens sur la transformation numérique de 2020 et 2022 fournissent des points d'appui utiles. Il reste toutefois à les décliner et à les opérationnaliser. Certains projets s'y emploient, tels SeCoIA Deal (Chagny et al., 2023), porté en France par la CFE-CGC et l'U2P, ou encore Dial-IA[2], porté et coordonné par l'IRES (Institut de recherches économiques et sociales).

Ainsi, la trajectoire d'appropriation ou de socialisation ne relève pas exclusivement des propriétés techniques de l'IA en question, ni des considérations éthiques s'y rapportant, mais tout autant d'un processus de dialogue social technologique qui est en pratique nécessaire pour orienter, qualifier et réguler collectivement l'usage de ces technologies et leurs effets sur les réalités du travail et des organisations. Faute de quoi, ce sont les promesses d'augmentation de la productivité et d'amélioration des conditions de travail qui ne sauraient être tenues.

— 2. DES INITIATIVES D'EMPLOYÉS, INSTRUCTIVES MAIS LIMITÉES

Le lancement mondial de ChatGPT d'OpenAI, version gratuite d'un LLM, a fait basculer l'IA au travail dans une nouvelle dimension que l'on peut qualifier à l'aune de trois dynamiques : un passage à l'échelle (si moins de 1 % des travailleurs français pouvaient témoigner d'une utilisation de l'IA au travail, ils seraient plus de 20 % aujourd'hui[3] (Beauchene et al., 2024)) ; un changement de trajectoire (d'initiatives d'employeurs à des initiatives d'employés, soutenues par l'initiative du fournisseur de la technologie) ; d'un sujet réservé à un club d'experts de grandes organisations à une conversation sociétale. Cette conversation sociétale alterne « panique morale » (Cohen, 1972), quand il s'agit d'évoquer une potentielle destruction d'emplois (Briggs et al., 2023) ou la fraude académique (Sciences Po, 2023), et pratiques fourmillantes mais discrètes, y compris au travail, à travers des usages personnels qui passent sous les radars. ChatGPT est l'application qui a atteint les 100 millions d'utilisateurs le plus rapidement de l'histoire du numérique : deux mois.

2 https://dial-ia.fr/ - Voir également l'article dans ce numéro.
3 https://www.talan.com/france/fr/Ifop-Talan-barometre-2025-francais-ia-generatives

De grands cabinets de consultants et autres instituts de sondage produisent de façon quasi hebdomadaire des données décrivant l'ampleur et l'orientation des pratiques. Si chaque étude doit être maniée avec précaution, leur macro-analyse révèle des invariants :

— Des usages majoritairement non institués : la majorité des usages de l'IA générative au travail se font à l'insu de la hiérarchie.
— Des pratiques diversifiées : l'efficacité déclarée se répartit équitablement entre productivité et créativité.
— Des pratiques qui n'évincent pas la peur de l'IA, ni des employeurs. Les usagers disent s'inquiéter du futur de leur emploi.
— Des pratiques installées : un tiers environ des usagers disent qu'ils ne souhaitent plus effectuer les tâches concernées autrement qu'avec l'IA.

Ces résultats qui décryptent les usages personnels de l'IA, par contraste avec les usages institutionnels issus des cas d'usage portés par les employeurs, sont doublement révélateurs. D'une part, ils soulignent une évolution radicale de la valeur des SIA au travail, mise en lumière par l'OCDE (2023) : contre le discours dominant des années 2010 qui faisait de l'IA un outil d'automatisation des tâches répétitives, manuelles et cognitives de faible niveau, la GenAI se révèle mieux adaptée aux tâches cognitives de haut niveau non répétitives. Ainsi, toujours selon l'OCDE, les manageurs, ingénieurs et directeurs généraux sont les professions ayant le degré de correspondance le plus élevé entre les capacités de l'IA et leurs tâches (OCDE, 2025) Autrement dit, la diffusion de la GenAI concerne en premier lieu les emplois qualifiés du tertiaire. D'autre part, ils mettent en évidence une trajectoire d'appropriation de ces systèmes par les travailleurs cherchant tout autant à améliorer la qualité de leur travail que leur qualité de vie et celle de leurs conditions de travail.

Cet intérêt spontané pour le potentiel de la GenAI au travail montre que les trajectoires et les motifs de son introduction dans une organisation et dans une activité doivent être pris en considération en tant qu'éléments de la chaine socio-technique qui font tenir ensemble les objets, les sujets et leurs milieux. Tandis que la productivité des organisations est une valeur prioritaire pour leurs dirigeants, il est démontré que la qualité du travail, le travail bien fait, est facteur de bien-être au travail (Clot, 2010) pour les travailleurs. Aussi, en tant qu'ils braconnent (Certeau, 1990), bricolent (Lévi-Strauss, 1962) ou butinent (Plantard, 2011), les travailleurs font avec ce qu'ils ont, et font preuve de créativité en se jouant des codes et des règles, en détournant des objets, en contournant des obstacles pour se dégager des marges de manœuvre. Leurs usages clandestins de l'IA générative leur permettent de vivre une certaine expérience de renforcement de leur autonomie, de libération des pratiques, d'encapacitation dans un espace d'action ordinairement contraint. Toutefois, ces usages, bien que générant un sentiment d'efficacité individuelle, exposent à des risques collectifs : sécurité des données, non-conformité, biais, atomisation des collectifs, affaiblissement des apprentissages et accroissement des inégalités (Ferguson, 2025).

L'analyse de ces usages invite à prendre du recul sur la dichotomie entre l'IA d'employeurs tournée vers la productivité d'une part, et l'IA de travailleurs orientée vers la qualité du travail, la qualité de vie et des conditions de travail d'autre part. Les enjeux d'appropriation doivent beaucoup à la façon dont les travailleurs gardent la main sur les pratiques pour être les principaux bénéficiaires de la valeur d'usage de la technologie. Ce qui ne signifie pas qu'ils abandonnent l'enjeu de productivité au profit de celui de qualité. Au contraire, ils peuvent, à travers l'usage de la GenAI, gagner secrètement en productivité, et leurs productions gagner possiblement en qualité, au moins parfois en apparence. Ce qui peut leur permettre de libérer du temps qu'ils peuvent ensuite réinvestir, ou non, dans diverses autres activités, mais aussi de faire valoir de nouvelles qualités pour tenter de jouir d'une nouvelle reconnaissance.

Alors, il ne s'agit pas tant – ou pas seulement – de chercher à discuter des impacts de l'IA sur le travail et l'emploi selon une logique techno-déterministe du numérique, que d'équilibrer l'interprétation en partant du travail réel et en adoptant une lecture socio-technique prenant davantage en compte les rapports de force et les intelligences pratiques dans lesquels s'inscrit la GenAI, et auxquels elle contribue. Ce faisant, il s'agit également de sortir du piège de la pensée technique qui se refermerait sur les technologies numériques proprement dites, quand bien même on leur reconnaitrait une dimension sociale dans leur conception et à travers leurs architectures et leurs scripts notamment. Mais aussi d'acter un pluralisme des techniques et des usages, et d'intégrer le fait que les techniques recouvrent bien d'autres dimensions avec lesquelles les technologies numériques doivent compter pour se développer et se déployer, et avec lesquelles elles doivent se combiner pour être appropriées, socialisées, civilisées (Mauss, 2012 ; Morozov, 2024 ; Tang et al., 2024).

Bien que particulièrement instructif sur les conditions sociales d'appropriation de l'IA, le bricolage personnel avec l'IA ne peut être considéré comme une situation de travail pleinement satisfaisante. Tout d'abord, ces pratiques recourent largement à l'outil dominant du marché, ChatGPT, et sont le plus souvent le fait de professionnels non acculturés à l'IA. Selon un sondage Ifop (2025), seuls 15 % des salariés français ont été formés à l'usage de l'IA dans une optique professionnelle[4]. La domination de cette technologie sur les usages interroge moins sur les tentatives de domestication de cet agent conversationnel réalisées par les travailleurs ou les employeurs, que sur l'ensauvagement des pratiques des fournisseurs de technologies qui, à grand renfort de communication, ouvrent au grand public des produits ni vraiment fiables ni vraiment conformes au droit. À quoi cela peut-il bien servir ? On peut émettre l'hypothèse qu'au-delà de faire la course en tête dans l'espoir de conserver une position dominante et de générer ainsi un haut niveau d'investissement en leur faveur, il s'agit aussi d'installer dans les entreprises, par la voie des employeurs ou des travailleurs, des manières de « faire le job » en tentant d'imposer leurs propres vues sur la façon dont il conviendrait de réaliser le travail, de le réorganiser. Au fond, ces manœuvres d'accoutumance ne viseraient-elles pas une forme de mise en dépendance, et ne constitueraient-elles pas une façon de domestiquer le travail et les entreprises, et de les mettre au service du développement des fournisseurs de technologies eux-mêmes ?

En deçà de ces interrogations très générales, l'actualisation d'une division technique et morale du travail à travers le développement et le déploiement de la GenAI peut aussi être saisie à des échelles beaucoup plus micro.

Par exemple, l'usage de la GenAI par des employés de centres d'appels ou de cabinets de conseil, qui offre aux travailleurs les moins expérimentés d'atteindre rapidement le niveau de productivité de ceux qui sont considérés comme des experts, pose des questions redoutables : faut-il y voir vraiment une montée en compétence ou en expertise, ou plutôt un simulacre qui ne tient qu'à la capacité des machines à délivrer plus rapidement les bonnes explications ou les bons conseils ? Que feront ces travailleurs si d'aventure ils se retrouvent privés de l'appui du SIA ? Auront-ils vraiment appris de cet appui, ou se seront-ils laissés porter ? Et les travailleurs considérés comme experts vont-ils eux-mêmes considérer cette forme de mise en concurrence comme loyale, alors même qu'ils ont mis des années d'expérience et d'efforts pour atteindre cette qualité de travail, pour construire cette expertise ? Y verront-ils une façon de dévaloriser leur expertise, d'affaiblir la position des experts dans l'organisation, et du même coup les possibilités d'échanger des « tours de main » contre des « coups de main » ? Et plus encore la possibilité même d'une construction collective d'une certaine professionnalité, d'une autre économie au sein même des organisations productives (Caillé et Grésy, 2014). L'analyse des usages montre en outre que de nouveaux réflexes techniques, « demander à l'IA », colonisent d'anciens

4 https://www.ifop.com/article/3eme-vague-du-barometre-sur-la-perception-et-lusage-des-ia-generatives-par-les-francais/

réflexes sociaux, « demander aux collègues » (Retkowsky, Hafermalz et Huysman, 2024). Or, ces gestes, coordonnés ni par l'organisation du travail ni par les pairs, entravent la circulation des connaissances et la définition partagée du « bon travail ». Tandis que l'inquiétude du remplacement des travailleurs domine l'arène publique, elle invisibilise celle du remplacement du collectif de travail. Ce ne sont donc pas uniquement les enjeux de développement des compétences, de reconnaissance des efforts, de cohésion des collectifs de travail qui sont mis en question avec le développement des usages de l'IA générative au travail. Le développement de ces usages soulève aussi la question des conditions d'un discernement sur ce qui fait la qualité et la valeur du travail, sans la confondre avec la quantité des productions et des livraisons. Lorsque ce discernement fait défaut, le risque est double : une dévaluation du travail (Nguyen et Mateescu, 2024) et un renforcement de la défiance envers les individus, les organisations et les institutions, par la dissimulation ou la naturalisation des pratiques qui façonnent les assemblages sociotechniques (Latour, 2006).

— CONCLUSION : QUELS DÉFIS POUR UN DIALOGUE SOCIAL TECHNOLOGIQUE AUSSI NÉCESSAIRE QUE BALBUTIANT ?

Le déploiement de la GenAI dans les milieux de travail constitue un enjeu d'innovation autant que de régulation. En effet, il ne s'agit pas seulement d'innover pour innover, ou de monter dans le train technologique par peur de rester sur le quai sans connaitre ni décider de la destination (Aggeri, 2023). Il ne s'agit pas non plus de réguler pour réguler, ou de réguler pour freiner, empêcher ou encore soutenir toutes formes d'innovation (Mandinaud et Ponce del Castillo, 2024). Mais plutôt de réguler pour maintenir et développer un cadre de travail favorable à la santé à travers un développement économique soutenable (Barcellini et al. 2024), et à l'autonomie des individus, des collectifs de travail et des entreprises elles-mêmes (Acemoglu et al., 2024).

Cela impose aux acteurs du dialogue social de sortir d'une certaine torpeur les ayant conduits ces dernières années à introduire le sujet de l'IA dans environ deux accords pour mille (Greenan et al., 2024). Mais sortir du « *somnambulisme technologique* » (Winner, 1986) pour retrouver des prises sur la situation et le cours de l'histoire du travail n'est pas simple. Et la démocratisation des usages de la GenAI ne saurait être réduite à la massification des usages individuels. Elle doit pouvoir passer par une revitalisation des processus collectifs qui soutiennent les possibilités d'une démocratie technique (Callon et al., 2001 ; Lequin et Lamard, 2025) et d'un design organisationnel (Boullier, 2017).

Pour cela, il est nécessaire de déverrouiller l'imaginaire socio-technique. D'autres qualités technologiques et dialogiques sont possibles. Ces possibilités dépendent de la capacité à créer les conditions d'apprentissage organisationnel susceptibles de transformer les usages individuels et collectifs des technologies, et les organisations elles-mêmes. Ces apprentissages doivent concerner l'ensemble des acteurs, et pas uniquement les opérateurs de bout de chaine : employeurs, manageurs, représentants du personnel, mais aussi concepteurs et acheteurs de technologies, designers d'organisation. Ils ne doivent pas être restreints aux seules conditions d'utilisation des technologies, mais s'étendre aux usages organisationnels et institutionnels eux-mêmes. Cela nécessite de questionner aussi, en lien avec l'introduction de SIA, les manières de manager le travail, de gouverner les organisations, de piloter les changements, d'envisager le rapport à la technologie et à ses fournisseurs, et d'impliquer les salariés et leurs représentants dans ces processus. Cela suppose de rouvrir l'horizon d'action, d'assumer la recomposition du champ d'expérience et la reconfiguration du système de relations de travail. C'est à cette condition que le dialogue social technologique peut dépasser un rôle consultatif ou défensif, pour devenir un véritable dispositif de gouvernance.

Les pratiques d'enquête, d'expérimentation, d'évaluation collective semblent être des clés intéressantes pour soutenir ce genre de dynamique, à même de mieux combiner les critiques des technologies, des organisations et du travail, plutôt que d'assimiler ces dimensions dans une mystique productiviste du progrès technologique (Musso, 2017) au service de la croissance et du bien-être.

— BIBLIOGRAPHIE

— Acemoglu, D., Autor, D. et Johnson, S. (2026), « Building pro-worker artificial intelligence », MIT Department of Economics.
https://economics.mit.edu/sites/default/files/2026-03/Building%20Pro-Worker%20Artificial%20Intelligence.pdf"

— Acemoglu, D., Bursztyn, L., Demirer, M., Fahey, L., Noy, S., Peng, S., Regier, J. et Zhang, W. (2024), « The simple macroeconomics of AI », National Bureau of Economics Research, MIT.

— Aggeri, F. (2023), *L'innovation mais pour quoi faire ? Essai sur un mythe économique, social et managérial*, Paris, Éditions du Seuil, 256 p.

— Barcellini, F., Gamkrelidze, T., Greenan, N., Jolivet, A. et Zouinar, M. (2024), « Le travail et l'emploi à l'épreuve de l'IA », CNAM.

— Beauchene, V., Laverdiere, R., Duranton, S., Walters, J., Lukic, V. et de Bellefonds, N. (2024), « AI at work : friend and foe », Boston Consulting Group.
https://www.bcg.com/publications/2024/ai-at-work-friend-foe

— Borel, S., Ferguson, Y. et Condé, J. (2024), « Étude des impacts de l'IA sur le travail. Rapport d'enquête LaborIA Explorer », LaborIA.

— Boullier, D. (2017), « Le problème n'est pas le numérique en lui-même, mais le capitalisme financier numérique », in *Revue des conditions de travail*, n° 6, « Mieux travailler à l'ère du numérique », p.56-62, Anact.

— Briggs, J., Kodnani, D., Hatzius, J. et Pierdomenico, G. (2023), « The Potentially Large Effects of Artificial Intelligence on Economic Growth », Goldman Sachs.

— Brynjolfsson, E. et McAfee, A. (2015), *Le deuxième âge de la machine. Travail et prospérité à l'heure de la révolution technologique*, Paris, Éditions Odile Jacob, 336 p.

— Brynjolfsson, E., Li, D. et Raymond, L.R. (2023), « Generative AI at Work », working paper, National Bureau of Economics Research.

— Caillé, A. et Grésy, J.-E. (2014), *La révolution du don. Le management repensé à la lumière de l'anthropologie*, Paris, Éditions du Seuil, 256 p.

— Callon, M., Lascoumes, P. et Barthe, Y. (2001), *Agir dans un monde incertain, Essai sur la démocratie technique*, Paris, Éditions du Seuil, 358 p.

— Casilli, A. (2019), *En attendant les robots*, Paris, Éditions du Seuil, 400 p.

— Certeau (de), M. (1990), *L'invention du quotidien*, établi et présenté par Luce Giard, Paris, Éditions Gallimard.

— Cerutti, E., Garcia Pascual, A., Kido, Y., Li, L., Melina, G., Mendes Tavares, M. et Wingender, P. (2025), « The Global Impact of AI : Mind the Gap », FMI Working Paper n° 2025/076, Fonds monétaire international.

— Chagny, O., Teissier, C., Bertholon, R., Blanc, N. et Matthys, L. (2023), « Servir la confiance dans l'intelligence artificielle par le dialogue », CFE-CGC.

— Chui, M. et al. (2023), « Le potentiel économique de l'IA générative : la prochaine frontière de la productivité », McKinsey Digital.

— Conseil d'orientation pour l'emploi (2017), rapport « Automatisation, numérisation et emploi » (trois tomes).

— Cohen, S. (1972), *Folk Devils and Moral Panics*, Londres, Routledge, 328 p.

— Commission européenne (2024), « Règlement sur l'intelligence artificielle ». https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=OJ:L_202401689
— Commission européenne, Direction générale des réseaux de communication, du contenu et des technologies (2019), « Lignes directrices en matière d'éthique pour une IA digne de confiance », Publications Office.
— Clot, Y. (2010), *Le travail à cœur. Pour en finir avec les risques psychosociaux*, Paris, Éditions La Découverte, 188 p.
— Dejoux, C. (2018), « Manager augmenté avec l'Intelligence Artificielle ? », in *La Revue RH&M*, n° 68.
— Dell'Acqua, C., McFowland, E., Mollick, E., Lifshitz-Assaf, H., Kellogg, K.C., Rajendran, S., Krayer, L., Candelon, F. et Lakhani, K.R. (2023), « Navigating the jagged technological frontier : field experimental evidence of the effects of AI on knowledge worker productivity and quality », working paper, Harvard Business School.
— Dizikes, P. (2024), « What do we know about the economics of AI? », MIT Economics.
— Escobar, C., McGovern, M.E. et Morales-Menendez, R. (2021), « Quality 4.0 : a review of big data challenges in manufacturing », in *Journal of Intelligent Manufacturing*, 32(8), 2319-2334.
— Ferguson, Y. (2025), « L'intelligence artificielle au travail. Accompagner & sécuriser les initiatives collaborateurs », rapport, Inria-Datacraft. https://datacraft.paris/wp-content/uploads/2025/10/Rapport-INRIA-_compressed.pdf
— Frey, C.B. et Osborne, M.A. (2013), « The future of employment : How susceptible are jobs to computerization ? », Oxford Martin School Working Paper. https://www.oxfordmartin.ox.ac.uk/publications/the-future-of-employment
— Ganascia, J.-G. (2022), *Servitudes virtuelles*, Paris, Éditions du Seuil, 320 p.
— Getapp (2023), « IA générative : comment les entreprises devraient-elles aborder cette révolution ? ». https://www.getapp.fr/blog/3801/ia-generative-guide-pour-pme
— Greenan, N., Napolitano, S. et Pillosio, J. (2024), « L'IA dans les entreprises : que révèlent les accords négociés ? », in *Connaissance de l'emploi*, n° 200, CNAM/CEET.
— Gmyrek, P., Berg, J., Kamiński, K., Konopczyński, F., Ładna, A., Nafradi, B., Rosłaniec, K. et Troszyński, (2025), « Generative AI and jobs : A global analysis of potential effects on job quantity and quality », ILO working paper. https://www.ilo.org/fr/publications/intelligence-artificielle-generative-et-emploi-revision-2025
— IDC (2017), « Worldwide semiannual big data and analytics spending guide », International Data Corporation.
— Latour, B. (2006), « Le prince : machines et machinations », in Akrich, M., Callon, M. et Latour, B. (dir.), *Sociologie de la traduction* (1), Paris, Presses des Mines, 401 p.
— Lequin, Y.-C. et Lamard, P. (dir.) (2025), *Éléments de démocratie technique*, Pôle éditorial de l'université de technologie de Belfort Montbéliard.
— Lévi-Strauss, C. (1962), *La pensée sauvage*, Paris, Éditions Plon, 395 p.
— Malakoff Médéric (2018), « Intelligence Artificielle et capital humain. Quels défis pour les entreprises ? », Boston Consulting Group.
— Mandinaud, V. et Keusch, M. (2025), « Le développement du dialogue social technologique en France », Anact.
— Mandinaud, V. et Ponce del Castillo, A. (2024), « AI, risks and working conditions », in Ponce del Castillo, A. (ed.), *Artificial intelligence, labour and society*, Bruxelles, ETUI, 259 p.
— Mauss, M. (2012), *Techniques, technologie et civilisation*, édition et présentation par Nathan Schlanger, Paris, Presses universitaires de France, 504 p.
— Misch, F., Park, B., Pizzinelli, C. et Sher, G. (2025), « Artificial Intelligence and Productivity in Europe », working paper 25/67, Fonds monétaire international.
— Morozov, E. (2024), « Une autre intelligence artificielle est possible », in *Le Monde diplomatique*, n° 845.

— Musso, P. (2017), *La Religion industrielle. Monastère, manufacture, usine. Une généalogie de l'entreprise*, Paris, Éditions Fayard, 800 p.
— Nguyen, A. et Mateescu, A. (2024), « Generative AI and Labour : Power, Hype and Value at Work », in *Data & Society*, 4 décembre.
— Nonaka, I. et von Krogh, G. (2009), « Tacit knowledge and knowledge conversion : Controversy and advancement in organizational knowledge creation theory », in *Organization Science*, 20 (3), p. 635-652.
— Partenariat mondial sur l'IA (2021), « Future of work, AI observatory at the workplace », GPAI Paris Summit.
— OCDE (2023), « Perspectives de l'emploi de l'OCDE 2023. Intelligence artificielle et marché du travail ».
— OCDE (2019), « Recommandation du Conseil sur l'intelligence artificielle », OECD/LEGAL/0449.
— OCDE (2025), « Algorithmic Management in the Workplace. New Evidence from an OECD Employer Survey », OECD Artificial Intelligence Paper n° 31.
— OpenAI (2025), « How your data is used to improve model performance », Help Center, FAQ Policy.
— Plantard, P. (2013), « E-inclusion : braconnage, bricolage, butinage », in *Place Publique*, n° 25, p. 16-22. hal-01739761
— Retkowsky, J., Hafermalz, E. et Huysman, M. (2024). « A ChatGPT-empowered workforce : Understanding its affordances and side effects », in *Business Horizons*, n° 67, Issue 5, p. 511-523.
— Sciences Po (2023), « Sciences Po interdit l'utilisation de l'outil ChatGPT », communiqué de presse.
— Simons, W., Turrini, A. et Vivian, L. (2024), « Artificial Intelligence : Economic Impact, Opportunities, Challenges, Implications for Policy », discussion paper 210, Commission européenne. https://economy-finance.ec.europa.eu/publications/artificial-intelligence-economic-impact-opportunities-challenges-implications-policy_en#disclaimer
— Tang, E. et al. (2024), *Plurality. The Future of Collaborative Technology and Democracy*, 584 p.
— Villani, C. et al. (2018), « Donner un sens à l'intelligence artificielle. Pour une stratégie nationale et européenne », rapport officiel, mission confiée par le Premier ministre.
— Weiner, J. (2021), « Five Methods to Avoid Common Pitfalls » (p. 59-61), in Weiner, J., *Why AI/Data Science Projects Fail : How to Avoid Project Pitfalls*, New York, Springer International Publishing, 81 p.
— Winner, L. (1986), *La baleine et le réacteur. À la recherche de limites au temps de la haute technologie*, Herblay, Éditions Libre, 332 p.

L'IA À LA RÉUNION : PERCEPTIONS ET STRATÉGIES DES PARTENAIRES SOCIAUX

Jean-Noël AMATO (Aract Réunion), **Corinne DUBOIS** (Aract Réunion), **Elisabetta ZIBETTI** (Université Paris 8)

Cet article examine les perceptions et positionnements d'acteurs locaux et de partenaires sociaux réunionnais face à l'intelligence artificielle (IA). À travers des entretiens avec des organisations syndicales (OS), patronales (OP) et des institutions, l'étude révèle une vision ambivalente de l'IA, perçue comme à la fois un outil d'efficacité et une technologie risquée. Les OS expriment des craintes concernant l'emploi et l'éthique, tandis que les OP voient l'IA comme un levier de performance. Les acteurs institutionnels interrogés soulignent son utilité mais insistent sur la nécessité d'un cadre réglementaire. L'article aborde également la souveraineté numérique et l'importance de développer des infrastructures technologiques adéquates pour soutenir l'adoption de l'IA dans un contexte insulaire. Enfin, il met en lumière l'importance du dialogue social et de la formation pour une adoption responsable et durable de l'IA, adaptée aux spécificités locales de La Réunion.

Mots-clés : intelligence artificielle, perception sociale, dialogue social, éthique, insularité

L'intelligence artificielle (IA) s'impose progressivement comme un levier de transformation majeur dans de nombreux secteurs économiques et sociaux à travers le monde (OCDE, 2025 ; World Economic Forum, 2025). À La Réunion, cette technologie émergente suscite à la fois des espoirs et des interrogations parmi les acteurs locaux. Dans un contexte marqué par des spécificités socio-économiques et géographiques uniques (ultra-périphérie et insularité, économie dominée par le secteur tertiaire, faiblesse du tissu industriel et de la R&D locale, forte croissance démographique, « rattrapage » économique, chômage structurellement élevé), l'adoption de l'IA soulève des questions cruciales quant à ses implications notamment pour l'emploi, les conditions de travail, les relations professionnelles et le cadre réglementaire (Amato et Dubois, 2024).

Plusieurs caractéristiques différencient La Réunion des autres territoires français, ce qui peut potentiellement avoir des répercussions sur les défis et opportunités perçus par les acteurs de terrain. Tout d'abord, La Réunion présente un contexte socio-économique marqué par un taux de chômage élevé : 16,7 % au troisième trimestre de 2025 (Insee, 2025a). En particulier, les moins de 25 ans représentent environ 24 120 demandeurs d'emploi sur un total de 158 620 inscrits dans les catégories A, B et C au troisième trimestre de 2025 (Ah-Woane, 2023 ; Fabre, 2015 ; France Travail et Dares, 2025). Un taux de

chômage élevé peut rendre la population plus sensible aux questions de « remplacement d'emplois » par l'IA. Son économie est fortement dépendante de secteurs tels que le tourisme et l'agriculture. Sa structure économique est dominée par le secteur tertiaire (85 %) et les entreprises de petite taille (92,2 % comptant de 0 à 19 salariés, dont 12,6% sans salarié[1]), ce qui détermine les besoins et les capacités d'adoption de l'IA. Une majorité des entreprises de La Réunion ont entre un et neuf salariés, ce qui représente une part significative des établissements locaux : 71 % (Insee, 2025b). Ces facteurs viennent influencer la perception et les capacités d'appropriation de l'IA.

Enfin, la situation géographique de La Réunion, en tant que territoire ultra-périphérique, a un impact sur l'accès aux technologies et aux compétences nécessaires pour l'implémentation de l'IA. Cela affecte notamment la connectivité et l'infrastructure numérique, soulignant l'importance de la souveraineté numérique et de la mise en place d'infrastructures technologiques adaptées.

Dans ce contexte nous avons mené, en avril 2025, une étude exploratoire sur les perceptions et les stratégies d'onze partenaires sociaux et acteurs institutionnels réunionnais (Tableau 1) face au déploiement de l'IA[2]. À travers une série d'entretiens qualitatifs menés auprès de cinq membres d'organisations syndicales, trois membres d'organisations patronales et trois acteurs d'institutions locales, nous avons cherché à comprendre comment ils appréhendent les opportunités et les défis liés à l'IA.

L'ensemble de ces acteurs sont, à divers niveaux, impliqués dans les dynamiques sociales et institutionnelles liées à l'intelligence artificielle à La Réunion. À l'exception de la FSU, les partenaires sociaux interrogés sont membres du comité paritaire régional (CPR) de l'Aract Réunion, qui est une instance de dialogue social territorial dédiée aux conditions de travail et aux transformations professionnelles. La CFE-CGC, membre du CPR, n'a pas participé aux entretiens. La participation de la FSU s'est effectuée à la suite d'une prise de contact spontanée consécutive à une publication diffusée sur les réseaux sociaux portant sur les enjeux de l'IA à La Réunion. Concernant le Medef Réunion, la personne interrogée a été identifiée en raison de son engagement local dans le champ du numérique et de l'intelligence artificielle ; il est par ailleurs ambassadeur de l'IA dans le cadre du plan national « Osez l'IA » depuis juillet 2025. Pour la CFTC, le secrétaire général adjoint a également participé aux échanges.

Tableau 1 : Description des acteurs consultés.

CATÉGORIE	ORGANISATIONS	PRÉCISIONS
Organisations syndicales (OS)	FSU (France Travail), CFDT (secteur sanitaire et social), FO, CFTC (grande distribution), CGTR	La FSU (France Travail) a sollicité l'entretien spontanément. La CFTC a été représentée par son secrétaire général adjoint.
Organisations patronales (OP)	CPME, Medef Réunion, UNAPL	Le représentant du Medef est ambassadeur de l'IA (plan « Osez l'IA ») depuis juillet 2025.
Acteurs institutionnels	DEETS Réunion, CGSS Réunion (Carsat), CESER de La Réunion	Les personnes interrogées sont impliquées, au titre de leur institution, dans des projets liés à la thématique de l'IA.

1 https://www.insee.fr/fr/statistiques/2011101?geo=DEP-974#chiffre-cle-12

2 Cette étude exploratoire constitue la première investigation menée sur le territoire réunionnais portant sur les relations entre intelligence artificielle, dialogue social et transformations organisationnelles. À ce jour, aucune donnée empirique préexistante n'était disponible sur ce sujet à La Réunion.

Chaque entretien a été enregistré avec le consentement des participants puis retranscrit afin de permettre une analyse qualitative. La grille d'entretien était articulée autour de quatre grands axes : (1) la perception, la représentation et la définition de l'intelligence artificielle ; (2) l'intégration de l'IA au sein des organisations syndicales et professionnelles et au sein des institutions ; (3) les types d'IA actuellement utilisés ou envisagés ; (4) la transformation numérique et les accords d'entreprise.

Cet article, s'appuyant sur l'analyse de ces entretiens, vise à éclairer les dynamiques de déploiement de l'IA à La Réunion, en mettant en lumière les défis et opportunités perçus par les acteurs de terrain.

1. DES DÉFINITIONS ET PERCEPTIONS DIVERSES DE L'INTELLIGENCE ARTIFICIELLE PARMI LES PARTENAIRES SOCIAUX ET ACTEURS INSTITUTIONNELS LOCAUX

L'analyse des entretiens menés auprès de partenaires sociaux de La Réunion révèle une diversité de perceptions et de définitions concernant l'intelligence artificielle. Ces perceptions varient non seulement en fonction des catégories d'acteurs, mais aussi en fonction de leur niveau de familiarité et d'utilisation de cette technologie.

1.1. Les définitions et perceptions sur l'IA

Les partenaires sociaux définissent l'IA de manière variée, souvent influencée par leur domaine d'activité et leur expérience de cette technologie. Pour certains, l'IA est perçue comme un algorithme capable d'apprendre et de s'adapter, recueillant des données pour améliorer ses performances. Cette vision met en avant la capacité de l'IA à inférer et à déduire des patterns[3], la distinguant des technologies déterministes[4] « traditionnelles ». Cependant, une confusion persiste fréquemment entre l'IA et l'automatisation, avec une méconnaissance des différences entre les technologies déterministes et les IA génératives comme les Large Language Models (LLM). Pour beaucoup, l'IA est synonyme de LLM, tels que ChatGPT, sans une compréhension approfondie des mécanismes sous-jacents. Cette perception réductrice peut conduire à une vision de l'IA comme une « boite noire magique », sans une pleine conscience de ses capacités et de ses limites.

Globalement, les perceptions de l'IA parmi les partenaires sociaux et les acteurs institutionnels sont ambivalentes. Une majorité des interviewés (6 sur 11) expriment un sentiment positif, voyant dans l'IA un outil facilitateur capable d'améliorer l'efficacité et l'innovation. Ces acteurs soulignent le potentiel de l'IA pour automatiser des tâches répétitives, libérer du temps pour des activités à plus forte valeur ajoutée et améliorer la précision des analyses et des prédictions. Cependant, 3 sur 11 manifestent des préoccupations, principalement liées aux risques de « remplacement d'emplois » et aux questions éthiques. Les questions éthiques citées incluent les biais algorithmiques, la protection des données personnelles et la transparence des décisions prises par les systèmes d'IA. Enfin, 2 interviewés sur 11 sont restés neutres, manifestant soit une méconnaissance soit une indécision quant à l'impact de l'IA. Ces acteurs reconnaissent le potentiel de l'IA, tout en insistant sur la nécessité d'une meilleure compréhension et d'un encadrement réglementaire pour garantir un usage responsable et bénéfique (cet aspect est développé dans Amato et Dubois, 2024).

3 Le terme « pattern » désigne une régularité, une structure récurrente ou un schéma identifiable dans un ensemble de données, permettant de repérer des similitudes, des tendances.

4 Une technologie déterministe est un système dans lequel le comportement est entièrement déterminé par des relations fixes et préétablies entre les causes et les effets. En d'autres termes, un système déterministe produit toujours le même résultat pour une entrée donnée, sans aucune variation ni entropie.
Par exemple, les systèmes de navigation par satellite utilisent des calculs déterministes pour déterminer la position à partir des signaux satellite, même si les conditions environnementales peuvent varier. On peut également citer les calculatrices, qui effectuent des opérations mathématiques de manière déterministe : l'addition de 2 et 2 donnera toujours 4.

Les représentants des organisations syndicales interrogés expriment plus souvent des sentiments « négatifs » (3 sur 5), avec des préoccupations marquées concernant le « remplacement d'emplois » et les implications éthiques de l'IA. Elles reconnaissent que l'IA peut améliorer l'efficacité et la productivité, mais insistent sur les « garde-fous », c'est-à-dire, selon eux, l'importance de la formation et de la régulation. Ces acteurs soulignent la nécessité d'un cadre réglementaire « strict » pour protéger les emplois et garantir des conditions de travail équitables. Cette attention prend à La Réunion une acuité compte tenu du taux de chômage élevé sur l'ile.

À l'inverse, les représentants des organisations patronales interrogés adoptent une vision plus positive de l'IA (3 sur 3), mettant en valeur son potentiel pour améliorer la performance et la rentabilité des entreprises. Cette unanimité positive peut refléter une confiance dans la capacité de l'IA à transformer et optimiser les processus de travail, tout en créant de nouvelles opportunités économiques. Elles ajoutent que l'IA peut aider les entreprises à se différencier sur le marché, en améliorant la qualité des produits et services. Au regard du nombre de TPE et des difficultés économiques des entreprises à La Réunion, ces acteurs perçoivent avant tout le potentiel économique que représenterait l'IA, même si ce potentiel reste une promesse.

Les acteurs institutionnels interrogés, pour leur part, expriment une perception le plus souvent positive de l'IA (2 sur 3), la considérant comme un outil pour l'exécution des tâches et le soutien à l'expertise. Néanmoins, tous évoquent des inquiétudes liées au remplacement d'emplois et la nécessité d'une certaine prudence.

1.2. Une appropriation encore émergente de l'IA par les acteurs du territoire

Les membres des OS, OP et les acteurs institutionnels réunionnais interrogés sont pour la plupart en phase initiale d'exploration et d'expérimentation de l'IA. Certaines organisations ont commencé à intégrer des solutions d'IA dans leurs processus, mais cette intégration est généralement limitée et non systématique. Par exemple, les entretiens révèlent que certaines organisations recourent à des outils numériques, parfois présentés comme intégrant de l'intelligence artificielle, pour automatiser des tâches administratives ou améliorer la communication interne via des applications mobiles (Figure 1).

Figure 1. Illustration d'applications numériques utilisées à La Réunion par les partenaires sociaux (Medef, CPME, FSU) pour l'automatisation de tâches et la communication interne. FSU Emploi est une application nationale.

Cependant, l'adoption de l'IA reste inégale et souvent non officielle, avec une utilisation « grise » qui prédomine, c'est-à-dire une utilisation non encadrée par des politiques ou des directives claires.

Parmi les personnes interrogées, les OP apparaissent comme les acteurs ayant le plus intégré l'IA dans leurs organisations. Les membres d'OP interrogés témoignent par exemple utiliser l'IA pour des tâches spécifiques comme la reconnaissance optique de caractères (OCR) et la numérisation de documents, ce qui facilite la gestion des archives et des données. Pour leur organisation mais aussi plus globalement pour les entreprises, ces acteurs voient dans cette technologie un levier stratégique pour augmenter la compétitivité et ouvrir de nouveaux marchés (Grollier, 2024 ; Rédaction Leader Réunion, 2025). Lorsque les OP expriment des craintes, celles-ci sont davantage liées à des préoccupations concernant la confidentialité des données et les implications éthiques de l'usage de l'intelligence artificielle. Ces acteurs se disent particulièrement attentifs aux risques de fuites de données et aux biais algorithmiques susceptibles d'affecter « les décisions prises par les systèmes d'IA ».

Les organisations syndicales qui utilisent l'IA évoquent essentiellement des expériences d'usage de l'IA générative pour la rédaction de rapports, des IA spécifiques pour l'analyse de données (notamment juridique), ainsi que des dispositifs d'automatisation pour les tâches répétitives (notamment des comptes rendus). Pour ce type de tâches, ces acteurs considèrent que ces outils peuvent contribuer à automatiser certains processus et à « libérer du temps pour des tâches à plus forte valeur ajoutée » (3 sur 5). Ces technologies sont perçues, dans ces contextes d'utilisation, comme un moyen d'améliorer l'efficacité sans nécessairement recourir à des solutions d'IA plus avancées. Parallèlement à cet usage de « premier niveau », les organisations syndicales de salariés (OS) interrogées manifestent majoritairement des sentiments « négatifs » (3 sur 5) à l'égard de l'IA. Cette crainte est le plus souvent attribuée à des préoccupations relatives à la confidentialité des données ou à une potentielle « perte d'emplois ». Une autre difficulté exprimée concerne le manque de connaissance et de compréhension de l'IA. Ces acteurs soulignent ainsi la nécessité de mettre en œuvre des actions de formation et de sensibilisation, susceptibles de lever la défiance ou les craintes liées à l'intégration de l'IA, mais aussi de favoriser le développement d'usages vertueux au sein de leur organisation et, plus largement, dans les entreprises.

Du côté des acteurs institutionnels, les entretiens font apparaitre une posture plus équilibrée, articulant reconnaissance des opportunités et exigences de régulation. L'IA est principalement envisagée comme un outil d'appui à l'expertise et d'amélioration de l'efficacité des services publics, sous réserve d'un encadrement réglementaire garantissant la protection des données et la cybersécurité. Cette posture traduit une fonction de vigilance et de régulation propre à ces acteurs, distincte à la fois des logiques de performance portées par les OP et des préoccupations centrées sur l'emploi exprimées par les OS.

Selon l'ensemble des acteurs interrogés, les ressources « limitées » – tant en termes de budget que de compétences techniques – disponibles sur le territoire constituent un obstacle significatif à l'adoption de l'IA. Les entretiens exploratoires font émerger des besoins soulignant l'importance de l'accompagnement et du dialogue social pour garantir un usage responsable et bénéfique de l'IA, tout en tenant compte des spécificités locales.

— 2. UNE REPRÉSENTATION PARTAGÉE DE L'IMPORTANCE DU DIALOGUE SOCIAL

Les acteurs interrogés font également valoir l'importance du dialogue social dans le déploiement de l'IA à La Réunion. Les partenaires sociaux s'accordent sur la nécessité d'un cadre réglementaire « solide » (OP) pour « encadrer l'utilisation de l'IA » (institutions locales). Ce cadre doit non seulement « protéger les emplois » (OS) et garantir des conditions de travail équitables, mais aussi promouvoir l'innovation et l'efficacité. En outre, la nécessité de la formation est un enjeu partagé, exprimé de manière récurrente. Si cette approche commune du dialogue social s'impose, les acteurs portent néanmoins des points spécifiques.

Les membres des OS insistent sur la question de l'emploi, et notamment de la « protection » des emplois liée à la crainte du « remplacement » par l'IA, sur les effets sur les conditions de travail et les métiers, ainsi que sur la nécessité de maintenir une intervention humaine dans les situations de travail.

Les OP mettent l'accent sur les accords d'entreprise, la formation et la performance économique. Elles reconnaissent les défis liés à l'encadrement de l'utilisation de l'IA tout en insistant sur l'importance de la formation pour maintenir la compétitivité.

Ces positions sont importantes, car l'implication des partenaires sociaux joue un rôle essentiel dans l'accompagnement des transitions numériques, en veillant à ce que ces transformations technologiques se fassent de manière équitable et responsable (Brugière et al., 2020). Un autre enjeu encore est à prendre en compte : les perceptions et les attentes des différents groupes dans le cadre des négociations et des accords d'entreprise liés à la transformation numérique.

3. LES OFFRES DE SERVICES PROPOSÉES PAR LES PARTENAIRES SOCIAUX ET LES ACTEURS INSTITUTIONNELS À LA RÉUNION : ATTITUDES ET APPROCHES

Les trois catégories d'acteurs interviewés insistent sur les besoins en termes de formation, mais selon des angles différents.

Les OS adoptent une approche majoritairement prudente (Ferry, 2025). Les syndicats reconnaissent le potentiel de l'IA pour faciliter certaines tâches, mais pointent la « nécessité de cadrer et de réglementer son utilisation » pour éviter des « dérives ». Le besoin de formation des membres des OS et des élus à l'utilisation de l'IA est un thème récurrent. En réponse, les OS s'engagent progressivement dans la formation de leurs membres, cherchant à sensibiliser les salariés aux caractéristiques de l'IA et à minimiser les « impacts négatifs sur l'emploi ». Cependant, elles disent faire face à un manque d'expérience pratique. Les opportunités envisagées incluent entre autres l'amélioration de la communication sur les enjeux liés à l'IA.

Les OP, voyant dans l'IA un levier pour améliorer la productivité et la compétitivité (Roder, 2019), expriment leur intérêt pour les applications à base d'IA et leur volonté d'explorer leurs opportunités. Elles soulignent néanmoins les besoins en formation à l'usage de ces applications s'agissant de tous les acteurs du monde du travail. Elles se disent par ailleurs conscientes des enjeux éthiques et des risques, tels que la « dépendance excessive à la technologie » et la « perte de compétences humaines ». Elles cherchent à collaborer avec des organismes spécialisés pour une intégration responsable.

Quant aux acteurs institutionnels interrogés, ils soulignent la nécessité de former les élus et les professionnels pour faire de l'IA un « facilitateur » plutôt qu'un « remplaçant du travail humain ». Ils insistent sur l'importance de l'éthique et de la déontologie des IA, portant le projet d'une IA pour l'aide à la décision laissant toute sa place au jugement humain dans les processus décisionnels (Kai-Fu et Chen, 2022). Bien qu'ils en reconnaissent le potentiel, ils mettent l'accent sur la nécessité d'un encadrement approprié, d'une formation adaptée et d'une régulation claire, afin d'en tirer pleinement parti tout en en limitant les risques (Rédaction Mémento, 2024).

La coopération entre les organisations syndicales de salariés, les organisations patronales, le CESER et la DEETS est un levier essentiel pour préparer le monde professionnel aux changements induits par l'intelligence artificielle (Andler, 2023), en favorisant à la fois l'innovation et la protection des travailleurs (Dufour, 2024).

CONCLUSION

Cette étude exploratoire des perceptions et des positionnements des partenaires sociaux réunionnais face à l'intégration de l'IA témoigne d'une dynamique complexe, marquée par des opportunités et des défis.

Les entretiens révèlent une perception ambivalente de l'IA, vue à la fois comme un « outil facilitateur » (OP) et une « technologie complexe et risquée » (OS), notamment en ce qui concerne l'IA générative. Les acteurs soulignent tous la nécessité d'un esprit critique face aux contenus produits par l'IA (Lee et al., 2025), tout en reconnaissant son potentiel pour améliorer l'efficacité et l'innovation. Les membres des OS et des OP interrogés manifestent une certaine prudence vis-à-vis de l'IA, mais pour des raisons différentes. En revanche, tous s'accordent, comme les acteurs institutionnels, sur l'importance de la formation et du dialogue social.

Cependant, si des normes et décisions juridiques contribuent à une clarification du cadre institutionnel, cela contraste avec la faible mobilisation observée du dialogue social autour de l'IA dans les organisations (Aract Réunion, 2025b). Une méconnaissance des cadres réglementaires et jurisprudentiels encadrant le déploiement de l'intelligence artificielle continue de caractériser les stratégies et pratiques des acteurs (Pascal, 2025 ; Sereno, 2025 ; Zerillo, 2025). Or, l'AI Act européen constitue désormais un référentiel disponible et potentiellement structurant (Union européenne, 2025), et plusieurs décisions judiciaires récentes sont venues préciser les obligations des employeurs en matière de dialogue social. Par une ordonnance du 14 février 2025, le tribunal judiciaire de Nanterre a ainsi jugé que le déploiement d'outils d'IA sans information-consultation préalable du comité social et économique (CSE) constituait une entrave à ses prérogatives (Baron-Charbonnier, 2025 ; Germain, 2025). Cette position a été confirmée par une ordonnance de référé du 15 juillet 2025 rendue par le tribunal judiciaire de Créteil, puis par une décision du 2 septembre 2025 du tribunal judiciaire de Paris, réaffirmant que l'introduction de nouvelles technologies relève du champ de l'information-consultation du CSE (Chanu, 2025). Le dialogue social, fondé sur une approche paritaire et participative, permet de concilier les intérêts en présence et d'aborder collectivement les défis et les opportunités liés à l'intégration de l'IA dans le monde du travail. En cela, les partenaires sociaux réunionnais occupent une place-clé dans l'accompagnement des transitions numériques.

On voit aussi que, pour une adoption responsable de l'IA à La Réunion, il convient de proposer des stratégies spécifiques qui tiennent compte des particularités locales. Cela pourrait inclure des recommandations destinées aux syndicats, aux employeurs et aux institutions, ainsi que des programmes de formation et d'éducation visant à mieux sensibiliser, former et accompagner la main-d'œuvre locale à l'adoption de l'IA (Berezaie, 2025). Une collaboration entre les secteurs public et privé apparait essentielle pour favoriser un écosystème technologique résilient et autonome, capable de répondre aux défis actuels et futurs liés à l'implantation et au déploiement de solutions d'IA.

En cohérence avec plusieurs études réalisées (Mandinaud et Keusch-Bessard, 2022 ; Rousseau et al., 2024), cette étude exploratoire met en évidence deux leviers essentiels pour encourager une utilisation à la fois responsable et innovante de l'intelligence artificielle au service du développement économique et social de La Réunion (Aract Réunion, 2025a).

1. Renforcer le dialogue social et promouvoir des politiques ainsi que des pratiques favorisant une appropriation éthique et créative de l'IA.
2. Accompagner la transition par la mise en place de formations ciblées, en adéquation avec les besoins exprimés par chacune des trois catégories d'acteurs interrogées.

BIBLIOGRAPHIE

— Ah-Woane, M. (2023, 1er juin), « À La Réunion, malgré des diplômes de plus en plus élevés, les jeunes accèdent toujours tardivement à leur autonomie », Insee. https://www.insee.fr/fr/statistiques/7623989

— Amato, J.-N. et Dubois, C. (2024), « L'IA : Kwa sa i lé ? 10 réponses sur l'IA, en 2024 » (p. 50), Aract Réunion. https://www.calameo.com/read/0073506287834d83e3891

— Andler, D. (2023), *Intelligence artificielle, intelligence humaine : la double énigme*, Paris, Gallimard, 432 p. https://www.gallimard.fr/catalogue/intelligence-artificielle-intelligence-humaine-la-double-enigme/9782072792885

— Aract Réunion (2025a, 26 mai), « Retour sur le CRIS de l'Aract Réunion : Focus sur l'IA », Anact. https://www.anact.fr/retour-sur-le-cris-de-laract-reunion-focus-sur-lia

— Aract Réunion (2025b, 26 septembre), « Comment les entreprises réunionnaises utilisent l'intelligence artificielle ? Résultats d'une consultation régionale », Anact. https://www.anact.fr/limpact-de-lintelligence-artificielle-sur-le-travail

— Baron-Charbonnier, C. (2025, 30 avril), « Décision du Tribunal judiciaire de Nanterre du 14 février 2025 : quel rôle pour le CSE dans les phases d'implantation de l'IA en entreprise ? », Lamy Liaisons. https://www.lamy-liaisons.fr/eclaireurs-du-droit/decision-du-tribunal-judiciaire-de-nanterre-du-14-fevrier-2025-quel-role-pour-le-cse-dans-les-phases-dimplantation-de-lia-en-entreprise/

— Berezaie, Y. (2025, 5 juillet), « "Osez l'IA" : Le plan de l'état français pour diffuser l'IA dans toutes les entreprises », Lebkblog. https://www.berezaie.fr/post/osez-l-ia-le-plan-de-l-état-français-pour-diffuser-l-ia-dans-toutes-les-entreprises

— Brugière, A., Bugand, L., Guibert, A., Parmentier, C. et Mandinaud, V. (2020), « 10 questions sur la maîtrise des transformations numériques », Anact. https://www.anact.fr/10-questions-sur-la-maitrise-des-transformations-numeriques

— Chanu, A. (2025, 23 septembre), « Droits des travailleurs et intelligence artificielle : le CSE doit être consulté », *L'Humanité*. https://www.humanite.fr/social-et-economie/chronique-juridique/droits-des-travailleurs-et-intelligence-artificielle-le-cse-doit-etre-consulte

— Dufour, M. (2024), *L'IA au service du travail et des entreprises*, Paris, L'Harmattan, 172 p. https://www.editions-harmattan.fr/catalogue/livre/lia-au-service-du-travail-et-des-entreprises/16116

— Fabre, E. (2015, 21 mai), « Les jeunes Réunionnais peinent à acquérir leur autonomie », Insee. https://www.insee.fr/fr/statistiques/1285798

— Ferry, L. (2025), *IA : grand remplacement ou complémentarité ?*, Paris, Éditions de l'Observatoire, 336 p. https://editions-observatoire.com/livre/IA-%3A-grand-remplacement-ou-complementarite-%3F/557

— France Travail et Dares (2025), « Les chiffres du marché du travail du 3ème trimestre 2025 – Réunion » (p. 8), France Travail, Dares, DEETS Réunion. https://reunion.deets.gouv.fr/Les-chiffres-du-marche-du-travail-du-3e-trimestre-2025

— Germain, S. (2025), « CSE, exigez la procédure d'information-consultation ! », in *Santé & travail*, juillet 2025. https://www.sante-et-travail.fr/cse-exigez-procedure-dinformation-consultation

— Grollier, B. (2024, 1er novembre), « Le tour de France de l'IA a commencé à La Réunion », in *Eco Austral - Actualités économiques et entreprises de l'Océan Indien*, décembre 2024-janvier 2025(393). https://ecoaustral.com/le-tour-de-france-de-lia-a-commence-a-la-reunion/

— Insee (2025a, 19 juin), « Taux de chômage localisé par département – La Réunion ». https://www.insee.fr/fr/statistiques/serie/010751340#Tableau

— Insee (2025b, juillet 10), dossier complet – Département de La Réunion (974). https://www.insee.fr/fr/statistiques/2011101?geo=DEP-974#consulter-sommaire

— Kai-Fu, L. et Chen, Q. (2022), *IA 2042*, Paris, Les Arènes Éditions, 544 p. https://arenes.fr/livre/ia-2042/
— Lee, H.-P. (Hank), Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R. et Wilson, N. (2025, 1er avril), « The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers », CHI 2025. https://www.microsoft.com/en-us/research/publication/the-impact-of-generative-ai-on-critical-thinking-self-reported-reductions-in-cognitive-effort-and-confidence-effects-from-a-survey-of-knowledge-workers/
— Mandinaud, V. et Keusch-Bessard, M. (2022), « Accord européen sur la transformation numérique des entreprises : Pour un dialogue social technologique » (Anticiper n° 2, Cahier de l'Anact, p. 24), Anact. https://www.anact.fr/accord-europeen-sur-la-transformation-numerique-des-entreprises-pour-un-dialogue-social
— OCDE (2025), « Intelligence artificielle ». https://www.oecd.org/fr/themes/intelligence-artificielle.html
— Pascal, A. (2025), « L'intelligence artificielle dans les ressources humaines », in *Chroniques du travail*, 15, Article 15. https://doi.org/10.4000/159lu
— Rédaction Leader Réunion (2025, 1er mars), « La Réunion à l'heure de l'IA », in *Leader Réunion*, mars 2025(239). https://www.leaderreunion.fr/la-reunion-a-lheure-de-lia/
— Rédaction *Mémento* (2024, 9 décembre), « Dr Luc Julia, invité d'honneur : éclairage sur l'intelligence artificielle à La Réunion », in Mémento, janvier 2025(545). https://www.memento.fr/article_09-12-2024-dr-luc-julia-invite-d-honneur-eclairage-sur-l-intelligence-artificielle-a-la-reunion
— Roder, S. (2019), *Guide pratique de l'intelligence artificielle dans l'entreprise. Anticiper les transformations, mettre en place des solutions*, Paris, Éditions Eyrolles, 198 p. https://www.editions-eyrolles.com/livre/guide-pratique-de-l-intelligence-artificielle-dans-l-entreprise
— Rousseau, T., Mandinaud, V., Chagny, O., Salis-Madinier, F., Garrigou, S. et Pecoste, C. (2024, 19 septembre), « Replay Intelligence artificielle : Comment en faire un sujet de dialogue social ? », Anact. https://www.anact.fr/video-intelligence-artificielle-dialogue-social
— Sereno, S. (2025), « Réglementer l'IA dans l'entreprise », in *Chroniques du travail*, 15, Article 15. https://doi.org/10.4000/159lw
— Union européenne (2025), « Loi sur l'intelligence artificielle de l'UE : Développements et analyses actualisés de la loi européenne sur l'IA » [EU Artificial Intelligence Act]. https://artificialintelligenceact.eu/fr/
— World Economic Forum (2025, janvier), *The Future of Jobs Report 2025*. https://www.weforum.org/publications/the-future-of-jobs-report-2025/
— Zerillo, E. (2025), « La protection des droits fondamentaux et la transformation de l'emploi salarié à l'épreuve de l'intelligence artificielle », in *Chroniques du travail*, 15, Article 15. https://doi.org/10.4000/159lt

{Entretiens}

Points de vue experts et institutionnels : fabriquer, réguler et évaluer les impacts écologiques

164 DE LA TRANSPARENCE À LA RÉGULATION DES ALGORITHMES

Entretien avec Benoit Rottembourg

171 UNE CONVENTION DE L'OIT POUR RÉGULER LE TRAVAIL DES PLATEFORMES

Entretien avec Cyril Cosme et Philippe Marcadent

179 QUELLES MÉTHODES POUR MESURER LES IMPACTS DE L'IA ?

Entretien avec Éric Fourboul et Gauthier Roussilhe

Points de vue de partenaires sociaux : quelles conditions pour un progrès économique et social ?

187 LA NÉCESSITÉ D'UNE RÉGULATION DE L'IA AU TRAVAIL

Entretien avec Franca Salis-Madinier

193 ENCADRER L'IA AU TRAVAIL PAR UNE DIRECTIVE EUROPÉENNE

Entretien avec Nayla Glaise

200 FAIRE DE L'IA, UN LEVIER D'INNOVATION POUR LES ENTREPRISES

Entretien avec Yahya Fallah

DE LA TRANSPARENCE À LA RÉGULATION DES ALGORITHMES

Entretien avec Benoit Rottembourg

Benoit Rottembourg a créé le projet REGALIA (REGulation des ALgorithmes d'IA), au sein de Inria (Institut national de recherche en sciences et technologies du numérique). Titulaire d'un doctorat en mathématiques appliquées, il a travaillé pendant plus de vingt ans dans le développement de logiciels et la data science au sein de grands groupes internationaux. Il consacre aujourd'hui ses travaux à l'audit et à la régulation des algorithmes. Ses recherches portent sur les biais, l'explicabilité et les conditions de contrôle des systèmes d'intelligence artificielle, en lien avec les enjeux de travail, de régulation et de protection des citoyens.
Dans cet entretien, Benoit Rottembourg rappelle que l'IA ne constitue pas un ensemble homogène de technologies et distingue plusieurs familles de systèmes, des algorithmes à base de règles aux modèles génératifs. Il montre comment l'évolution des méthodes de développement, la mondialisation des activités informatiques et le recours massif au travail de labellisation des données ont profondément modifié les conditions de travail de celles et ceux qui conçoivent et alimentent ces systèmes. L'entretien met également en lumière les risques associés à l'industrialisation des outils d'IA : éloignement du travail réel, opacité des décisions algorithmiques, biais, difficulté de contestation et développement du management algorithmique. Face à ces enjeux, Benoit Rottembourg invite à dépasser la seule question de la transparence des algorithmes pour construire de véritables mécanismes de régulation par des processus d'explicabilité des décisions algorithmiques. Il plaide notamment pour le développement d'une « algo-vigilance », l'archivage des décisions algorithmiques ou encore le renforcement des capacités des représentants des travailleurs à comprendre, discuter et contester les systèmes qui affectent leur activité.

Entretien réalisé par Marion Gilles et Anne-Marie Nicot (Anact)

— Quand on parle d'IA, de quoi parle-t-on ? Pouvez-vous décrire les « ingrédients » des systèmes à base d'IA ? De quoi ces dispositifs techniques sont-ils composés ?

Benoit Rottembourg : Lorsqu'il s'agit de définir l'IA, il est souvent mentionné la définition de l'OCDE[1]. Je préfère une approche plus descriptive pour définir l'IA. On peut distinguer quatre grandes familles de technologies utilisées dans les systèmes d'aide à la décision, dont seules les trois dernières sont communément qualifiées d'IA.

Dans la première famille, nous avons les technologies à base de règles qui ont été développées dans les années 1990. Dans ces cas-là, c'est un être humain qui invente et négocie des règles qui sont ensuite implémentées dans un algorithme. Par exemple, dans un outil de planification, pour construire les emplois du temps dans une équipe, on rentre la règle « La pause repas a lieu moins de quatre heures après l'arrivée » ou la règle « Chaque employé doit recevoir le même nombre de shifts nocturnes dans le mois ». Et les règles sont explicites, définies par des humains, négociables entre humains, avant d'être implémentées. Par exemple, on peut hésiter entre les règles « Manger moins de quatre heures après l'arrivée » et « Toute personne A arrivant avant une personne B dans le centre mangera avant B ». Ces systèmes ne sont généralement pas considérés comme de l'IA mais peuvent déjà créer des problèmes lorsque les règles s'enchevêtrent. La conséquence de l'application de règles pourtant connues et simples, quand elles font l'objet d'une optimisation par un algorithme, sur une grosse population d'employés, est parfois inattendue, et la règle rejetée ou amendée.

Deuxième famille : l'IA symbolique et les mathématiques appliquées. Ici, les algorithmes sont bien décrits par des mathématiciens ou des informaticiens et sont utilisés pour faire soit des diagnostics (par exemple : « Tu as un risque élevé d'avoir Alzheimer »), soit des prévisions (par exemple : « La maladie s'aggravera dans 10 mois »), soit des prescriptions (par exemple : « C'est toi qui viendras à huit heures »). Ces systèmes offrent l'illusion d'être plus contrôlables car ils sont conçus explicitement par des mathématiciens ou des informaticiens, qui peuvent décrire, définir les algorithmes et disposent en général de réflexes sur leurs biais ou effets indésirables potentiels. Ce sont en effet des systèmes relativement contrôlables et explicables. Mais ils peuvent aussi générer des biais. Par exemple, dans un logiciel de génération de plannings que j'avais réalisé à la fin des années 1990, les personnes en fin d'alphabet avaient des emplois du temps plus dégradés que ceux du début d'alphabet, et je n'aurais pas pu le deviner si on ne me l'avait pas signalé. L'algorithme de Parcoursup est un bon exemple d'algorithme de cette catégorie. Intentionnellement simple pour être négociable, il est utilisé dans un système complexe (avec des universités disposant de règles différentes et de capacités distinctes) et dans une dynamique temporelle multi-acteurs (les 980 000 élèves qui décident au fil de l'été). Des effets indésirables et difficilement anticipables peuvent se produire malgré la simplicité de l'algorithme (dit algorithme de « mariage stable » et inventé en 1962).

Troisième famille : l'IA connexionniste, qu'on appelle fréquemment par analogie les « réseaux de neurones », déployés dans l'industrie depuis les années 2010. Cette technologie exploite beaucoup plus de données, très riches et très fines (individuelles et souvent comportementales), et de puissance de calcul. C'est une IA qui n'est plus explicite, puisque les règles ne sont pas écrites par l'homme : l'algorithme « apprend » sur des exemples et la « labellisation » des exemples (« ceci est une tumeur maligne, ceci n'est pas une tumeur maligne »), au lieu de suivre des règles ou des programmations explicites. Plutôt que d'écrire explicitement des règles, on demande à l'algorithme d'apprendre à partir d'exemples, ce qui provoque une certaine perte de contrôle et pose la question du juste choix des exemples d'entrainement et de leur labellisation. Cette approche s'accompagne aussi d'une culture du développement différente, et d'une sous-traitance généralisée – j'y reviendrai – qui a créé une distance plus grande entre ceux qui développent les algorithmes et ceux qui les utilisent ou les subissent.

1 https://www.oecd.org/fr/publications/2019/06/artificial-intelligence-in-society_c0054fa1.html

Enfin, la quatrième famille, c'est ce qu'on appelle les LLM (pour Large Language Models), ou modèles génératifs, qui sont de grands modèles de langage, comme ChatGPT, ou des variantes multimodales (générateurs de sons, d'images ou de vidéos). C'est une extension des réseaux de neurones, avec une dimension supplémentaire d'interface humaine, mais sans changer fondamentalement la nature des problèmes posés par l'IA connexionniste. Ils ajoutent un « confort humain », parfois une « illusion humaine », dans le dialogue et l'échange avec l'utilisateur, permettant à des personnes moins compétentes d'accéder aux outils, ce qui représente une extension des problématiques précédentes mais sans nécessairement poser des problèmes fondamentalement différents : on retrouve la problématique de la finesse des données collectées sur les comportements et la puissance de la boucle de feed-back (pour alimenter l'apprentissage automatique), qui tend à évacuer l'humain du contrôle du système. Le caractère hautement persuasif de ces outils (on parle de « compagnons IA » et de « machines à faire croire ») augmente toutefois la confiance qu'on accorde à leur réponse face à un système plus austère dans son interface. Et donc les conséquences d'erreurs, d'hallucinations ou de biais sont en croissance.

Par ailleurs, les algorithmes modernes fonctionnent souvent en parallèle ou en emboitement avec d'autres algorithmes. C'est pourquoi on ne devrait jamais parler d'un algorithme au singulier, car il s'agit toujours de systèmes complexes comprenant de nombreux algorithmes qui interagissent entre eux tels des agents assemblés. En outre, les entreprises modernes pratiquent constamment des « A/B tests », pour lesquels plusieurs variantes d'algorithmes sont testées simultanément sur différents groupes d'utilisateurs afin de comparer leurs performances. Par exemple, un outil de calcul de prix sur internet peut proposer, au même moment, deux prix différents à deux personnes qui veulent un même voyage. Au bout de quelques heures ou jours, le plus efficace des deux algorithmes sera retenu et généralisé à tous les clients.

— Quelles ont été les principales évolutions dans les méthodes de développement, et donc dans le travail de celles et ceux qui conçoivent, développent ces dispositifs ?

BR : Les méthodes de travail ont en effet beaucoup évolué dans le développement de ces systèmes, et je vous parle avec un regard vieux de plus de 35 ans de développement logiciel. Tout d'abord, on est passé d'une approche de spécification détaillée puis de développement (cycle en V), avec des cycles de plusieurs mois, à des méthodes dites « agiles », avec des « sprints » de courte durée et des « tickets » pour réaliser des corrections en cas de problème. Le sprint, c'est un bloc de travail de 15 jours, avec une logique de « prototypage rapide » et itératif : on démarre avec une trottinette, qui permet le déplacement de A à B, puis on lui met un moteur pour la vitesse, une carrosserie pour la pluie, et à la fin, ça produit une voiture. Les équipes se réunissent régulièrement pour planifier le travail et suivre les « tickets de développement ». Il y a tout un cérémonial autour de ces méthodes agiles, avec réunions quotidiennes, tickets à résoudre, sprints à compléter. Cela peut donner une illusion de contrôle, mais en fait cela permet parfois, par tâtonnement, d'éviter les effets tunnels, en forçant le client (le « métier ») à exprimer ses retours le plus tôt possible. La majorité des causes de « tickets » sont des problèmes de fonctionnement, de design, de maintenance, etc.

Dans ces boucles, on retrouve très rarement des syndicats ou des représentants du personnel, mais plutôt des utilisateurs « types » retenus par la direction du projet. Il y a un cadre, l'objectif économique, qui est en général bien défini (livrer une pizza le plus rapidement possible entre A et B), et des contraintes dures (respecter les sens interdits). Et puis, il y a toutes les qualifications de performance et d'externalité. Par exemple, quand un livreur de repas est victime d'un accident et qu'il est dans le coma sur le bord de la route, la seule chose que sait l'algorithme de planification, c'est que la pizza refroidit (il y

a retard par rapport au prévu et au prescrit), alors que les données de géolocalisation captées en permanence permettraient de savoir que la personne ne bouge pas. Pour la plateforme, la livraison dans les temps est une question de performance alors que la santé du livreur est une externalité.

Par ailleurs, pour faire des économies, on a opéré une internationalisation et une délocalisation du développement, avec une répartition des tâches selon la proximité avec le client. Il y a une division internationale du travail selon les niveaux d'intervention. Schématiquement, les travailleurs européens ont la charge des interfaces métier et de la compréhension des besoins (les activités qualifiées et couteuses) ; les développeurs délocalisés dans les pays européens (par exemple, les développeurs d'origine malgache qui viennent travailler en France, ou les travailleurs indiens en Angleterre) s'attachent aux interfaces techniques, dans des conditions de travail souvent plus précaires comme celles des ESN (entreprises de services numériques, ex-SSII) ; enfin, les équipes à l'étranger, « offshore » (en Inde) ou « nearshore » (dans les pays du Maghreb, par exemple), assurent la maintenance et le développement technique, à des couts réduits et dans des conditions de travail et d'emploi dégradées par rapport à l'Europe. Leur rôle est de résoudre les « tickets » le plus rapidement possible et à plus bas couts. Pour cela, il y a parfois des équipes dans tous les fuseaux horaires. Certaines entreprises ont jusqu'à 60 % de leur personnel de développement en Inde, ce qui n'est pas sans poser des défis en matière de qualité, notamment du fait de la distance culturelle et géographique avec les utilisateurs finaux des systèmes. Cela peut poser des problèmes de maintenabilité, augmentés, semble-t-il, par le recours à l'intelligence artificielle générative chez les développeurs (comme l'utilisation de Copilot).

Un autre aspect important du développement de l'IA est le travail de labellisation-qualification des données en masse, d'où le rôle crucial des « travailleurs du clic[2] » qui effectuent des micro-tâches de labellisation pour entrainer les algorithmes. Ces travailleurs, souvent dans des pays à bas couts, qualifient manuellement des données afin d'entrainer les algorithmes. Par exemple, dans les cantines, se développent de plus en plus des systèmes de reconnaissance automatique des éléments des plateaux pour calculer le prix à payer et automatiser l'ensemble de la prestation. Afin de développer cette technologie, des personnes, par exemple en Afrique, réalisent des micro-tâches de labellisation d'images pour faire la différence entre différents types de pains ou de boissons. Ce travail nourrit un système qui va pouvoir reconnaitre le plateau-repas (pain, yaourt, toutes les sortes de boissons, etc.) et le facturer automatiquement, sans recours à une caissière dans les périodes de pic. Ces personnes peuvent même intervenir en temps réel si le système échoue à reconnaitre un élément, en envoyant l'image du plateau-repas aux micro-travailleurs. Tout cela entraine les algorithmes. Les conditions de travail et d'emploi des micro-travailleurs sont très dégradées : au Kenya, certains travailleurs sont payés à peine 2 euros par jour pour réaliser des tâches de correction des réponses de ChatGPT. Ce travail peut être continu, même après le déploiement initial, pour améliorer les algorithmes. Certaines tâches peuvent être complexes et traumatisantes. Par exemple, la modération de contenus (modération de contenus violents, pédopornographiques, etc.) est également délocalisée dans des conditions de réalisation proches. En France, quand du personnel est employé pour de la qualification d'images présumées pédocriminelles – cela arrive malheureusement (on parle de plusieurs dizaines de milliers de vidéos « qualifiées » en France tous les ans) –, celui-ci est accompagné par une équipe psy à intervalles réguliers.

2 Référence citée : recherche sur les Microtaskers et leur rôle dans l'entrainement des algorithmes (Antonio Casilli, *En attendant les robots, Enquête sur le travail du clic*, Paris, Éditions du Seuil, 2019, 400 p.).

— L'un des neuf principes de prévention inscrits dans le Code du travail – Adapter le travail à l'homme – est-il compatible avec ce type d'outils qui ont d'abord pour objectif l'efficacité et dont le développement se fait sur des bases de plus en plus éloignées du travail réel ? Ce principe de prévention peut-il avoir une place dans ces processus de fabrication et de développement des SIA, les systèmes d'intelligence artificielle ?

BR : Certaines questions relèvent du développement des SIA, et d'autres des usages.

Tout d'abord, peu d'entreprises développent de l'IA pure, et elles sont rarement en Europe, comme nous l'avons vu. En revanche, des éditeurs proposent de plus en plus fréquemment un SIA déjà développé pour toutes sortes de fonctions à automatiser, des décisions à « augmenter ». Avec ces solutions sur étagère, il n'y a pas d'association des travailleurs de l'entreprise-cible à la conception de l'outil. Cette industrialisation de la fabrication des outils n'est pas sans poser des problèmes : l'écart entre le prescrit et le réel peut être considérable quand les équipes de développement n'ont jamais vécu la situation qu'elles planifient. C'est le syndrome des « trois couches par jour » chez Orpea, où il faut n'avoir jamais vécu une épidémie de gastro en Ehpad pour imaginer que la règle des trois couches est pertinente ; c'est une moyenne, mais les cas hors moyenne sont graves. Quand, au-delà des règles de gestion malheureuses, c'est un outil opaque qui produit la commande de couches, et qu'il n'a pas été entrainé sur des données fines (gastro incluse), le SIA est victime du syndrome « hors-sol » ou « pensé depuis le siège », et les effets indésirables ne seront découverts que très tardivement. C'est d'ailleurs une des raisons du rejet lors de déploiements d'IA, et des faibles taux de succès de l'IA générative en entreprise aujourd'hui.

Pour ce qui concerne l'usage des SIA, les entreprises ne sont pas à la hauteur de la révolution en cours. Il existe un vrai sujet d'acculturation non seulement sous-estimé dans l'entreprise, mais noyé dans l'effet magique des IA génératives. J'ai eu l'exemple récent d'une direction RH de grande entreprise, recevant des dizaines de milliers de CV par an à trier, et qui souhaitait mettre en place un outil de tri automatique (un scoring) assisté par IA. Plusieurs start-up venaient de leur proposer des solutions clé en main de tri de CV. En interviewant les chefs de projet sur les métriques de performance retenues pour valider la pertinence de la solution, indépendamment des risques encourus au sens du règlement européen sur l'IA, nous avons été stupéfiés par les critères énoncés. La notion de faux positif, la notion d'intervalle d'incertitude, la notion de proximités entre scores humains et machines, n'étaient pas acquises. En l'état culturel des acheteurs et des pilotes du projet, la solution ne pouvait qu'être déceptive. Et du point de vue des biais éventuels dans le tri induit, aucune démarche n'avait été mise en place malgré le caractère reconnu comme à « haut risque », au sens du règlement européen sur l'IA, de telles applications. Il y a des injonctions managériales fortes, donc des budgets et une invasion du domaine quotidien du travail, sans réflexion sur l'usage et les risques des outils qui vont transformer le travail, sans protocole pour anticiper les problèmes. En l'absence de véritable préparation et réflexion sur les choix des outils et l'adaptation du process qui en découle, la chaine de sous-traitance et l'éloignement géographique et culturel entre développeurs et utilisateurs créent des problèmes supplémentaires. Ce n'est pas sans lien avec la perception actuelle d'une bulle de l'IA.

Et puis le management algorithmique comporte ses propres effets. Le management algorithmique ne se limite pas à avoir un algorithme comme patron mais englobe l'évaluation, la prévision et la prescription. Il s'agit d'un « tourbillon algorithmique » dans lequel de nombreux travailleurs sont, de fait, impliqués. Par exemple, les notifications et injonctions algorithmiques engendrent une pression constante, même pour des utilisations ordinaires : chaque fois qu'on reçoit une notification de notre agenda, comme « réunion dans 4 minutes », c'est une forme de management algorithmique. Un mélange de surveillance et de prescription. Il y a aussi du management algorithmique sans IA, comme les systèmes de distribution d'appels dans les centres d'appels, ou le « voice picking » en logistique dans les années 2000. L'IA introduit une échelle supplémentaire, cependant, du fait de la collecte plus

fine des données (accéléromètres, données comportementales), d'une plus grande puissance de calcul, de boucles de rétroaction qui peuvent créer des prophéties auto-réalisatrices, d'absence de possibilité de discussion ou de contestation des décisions algorithmiques – malgré l'apparence des interfaces plus « humaines » (voix de synthèse perfectionnées, LLM) –, d'opacité des règles.

Mais l'IA ne fait que le dernier mètre : on lui a déjà enveloppé le monde[3]. Le problème, c'est que de nombreuses dimensions importantes pour les travailleurs sont absentes de ce monde, comme dans l'exemple d'un livreur de repas qui serait dans le coma sur le trottoir et recevrait toujours des notifications lui indiquant qu'il est en retard pour sa livraison… Cette question renvoie aux objectifs et critères de performance qui sont assignés aux SIA – et là, ce sont des décisions humaines, souvent faites à bas bruit par quelques personnes qui n'utilisent leur vélo que de manière récréative.

L'évolution vers des algorithmes plus sophistiqués a permis une efficacité accrue. Il est incontestable que gagner six mois sur une détection d'Alzheimer ou de certains cancers est porteur d'espoir pour la société. Mais cette nouvelle puissance d'aide à la décision a également créé de nouveaux dangers, notamment la perte de contrôle et les biais inhérents aux données comportementales. Saviez-vous que, quand on emploie un LLM pour prévoir le niveau d'urgence d'un patient (basé sur des données métaboliques et des verbatim du personnel aux urgences en France), on sous-estime l'urgence pour les femmes par rapport aux hommes, à situation de santé comparable ? Avec l'IA, les systèmes deviennent plus opaques : les règles ne sont plus explicites mais apprises à partir des données, et les décisions prises par les algorithmes sont, de fait, difficiles à expliquer et… très rapidement prises – d'autant plus que les algorithmes évoluent avec les données collectées. Même la personne qui l'a développé doit consacrer un temps important à éclairer ce qui se passe, avec des règles globales et l'explication locale d'une décision. Car, si le basketteur français star de la NBA, Victor Wembanyama, qui a une taille exceptionnelle de 2 m 24, était une femme, quelle taille lui donneriez-vous ? Probablement pas 2 m 24, car une femme aussi exceptionnellement grande que Victor, à six écarts-types de la moyenne des femmes, mesurerait en fait 2 m 05. Quel rythme cardiaque, quelle température corporelle ? Quel risque face à telle maladie aurait-elle, « toutes choses égales par ailleurs » ? Cette opacité provoque une absence de contradictoire : il est devenu impossible, pour les travailleurs, de contester ou discuter les décisions algorithmiques, contrairement aux situations avec un manager humain avec qui on peut s'expliquer. L'algorithme ne permet pas de dialogue. Et si les algorithmes peuvent reproduire ou amplifier des biais existants, parfois de façon invisible, qui surveille l'algorithme ?

— Transparence, discutabilité, auditabilité, explicabilité : quelles sont les marges possibles ? Est-ce que la transparence des algorithmes d'IA est possible ? Qu'est-ce qui peut être rendu discutable ? Et quels espaces possibles de transparence et de discutabilité avec des algorithmes qui évoluent par eux-mêmes ? Peut-on envisager qu'en développant transparence et discutabilité, les travailleurs gardent la main sur leur travail ?

BR : De mon point de vue, la transparence n'est ni nécessaire ni suffisante pour réguler les algorithmes. C'est vertueux dans l'absolu, bien sûr, mais trompeur. Comprendre comment un algorithme fonctionne n'assure pas qu'il ne causera pas de préjudice. L'explicabilité des décisions spécifiques est plus importante : pour chaque plainte, il faut une explication. Mais cette explicabilité est difficile à obtenir. En effet, les systèmes évoluant constamment, l'analyse rétrospective des décisions algorithmiques est impossible sans un archivage des données et des décisions algorithmiques. C'est particulièrement vrai pour les algorithmes qui évoluent rapidement et dépendent de données en temps réel. Cela crée une opacité des algorithmes qui se traduit notamment par la difficulté à comprendre comment les décisions sont prises, surtout avec les systèmes d'apprentissage automatique (*machine learning*). Il

3 Référence citée : Luciano Floridi, *L'Éthique de l'intelligence artificielle. Principes, défis et opportunités,* Sesto San Giovanni, Éditions Mimesis, 2023, 320 p.

n'en reste pas moins qu'on peut contrôler l'IA dans les métiers où le prix à payer de l'erreur est élevé, comme dans les domaines médical ou juridique, ou dans la fonction publique pour des décisions touchant le citoyen.

Or, contrairement au domaine médical avec la pharmacovigilance, il n'existe pas d'« algo-vigilance » systématique pour détecter les effets négatifs des algorithmes. L'idée serait donc de créer l'équivalent de la pharmacovigilance pour les algorithmes, avec une obligation de collecter et d'analyser les problèmes rencontrés. Ce serait une forme de régulation. Cela nécessiterait plusieurs conditions.

Tout d'abord, un tel dispositif ne peut se faire sans imposer un archivage obligatoire, ce qui signifie d'imposer aux entreprises de conserver les données et décisions algorithmiques pour autoriser des analyses rétrospectives, comme c'est déjà le cas dans le domaine médical. Cet archivage coute cher, et il faudrait définir un cadre pour imposer des obligations d'archivage aux entreprises utilisant des algorithmes qui impactent les travailleurs. Cette traçabilité n'est pas dans la culture des éditeurs de SIA.

Il faudrait aussi une mobilisation, que des gens puissent se plaindre (on a vu récemment la difficulté à dénoncer la présence d'une poupée sexuelle juvénile sur un site d'e-commerce), et qu'il y ait un « juge de paix » pour faire poids face à la direction et à l'éditeur. Comme dans le cas de la class-action des utilisateurs de Tinder sur les différences de prix de l'application selon l'âge[4]. Je parle de class-action (ou « action de groupe », en France), car un tel dispositif ne peut se faire sans collecte statistique : il faut rassembler suffisamment de données pour identifier les biais et problèmes. On ne peut pas attraper un algorithme sur un cas, il faut l'attraper sur des statistiques, puisque lui-même parle statistiques. Les syndicats pourraient jouer un rôle important dans ce type d'actions collectives. Avec 100 livreurs à vélo, pendant 100 jours de travail, sur une même ville ou un même quartier, on peut savoir si l'algorithme de dispatch est biaisé. Sinon, c'est la légende urbaine qui domine. Et l'entreprise oppose aujourd'hui son secret des affaires pour ne pas être observée de la sorte.

Enfin, pour pouvoir traiter réellement et pratiquement ces problèmes, il est nécessaire de mobiliser des collectifs pluridisciplinaires, associant experts techniques, experts juridiques et représentants des travailleurs, des métiers. De véritables mécanismes d'explicabilité ne peuvent résulter que d'un travail pluridisciplinaire. Cela veut dire qu'il faut former les personnes engagées dans ce travail, en particulier en renforçant la capacité des syndicats et représentants du personnel à comprendre et contester les systèmes algorithmiques.

C'est donc plus un enjeu de régulation, par des processus d'explicabilité des décisions algorithmiques, que de seule transparence du code. D'ailleurs c'est bien pour cela que j'ai renommé mon équipe et le projet associé Regalia (pour régulation des algorithmes[5]) à la place de Transalgo (pour transparence des algorithmes), car la transparence seule, outre la difficulté à l'atteindre, ne résout pas les problèmes. On a besoin de développer des mécanismes de régulation adaptés à la nature évolutive des algorithmes, et de renforcer la capacité des travailleurs (et des citoyens en général) à comprendre et contester les décisions algorithmiques qui les affectent.

4 https://slatervecchio.com/fr/class-action/tinder-class-action-canada/
5 https://www.inria.fr/fr/regulation-algorithmes-projet-regalia

UNE CONVENTION DE L'OIT POUR RÉGULER LE TRAVAIL DES PLATEFORMES

Entretien avec Cyril Cosme et Philippe Marcadent

Cyril Cosme est directeur du Bureau international du travail (BIT) de l'Organisation internationale du travail (OIT) pour la France.

Philippe Marcadent est ancien chef du bureau des conditions de travail et de l'emploi au BIT à Genève.

L'adoption par l'Organisation internationale du travail (OIT) de la Convention concernant le travail décent dans l'économie des plateformes, en juin 2026, constitue une étape importante dans la construction d'un cadre international pour appréhender les transformations du travail liées aux technologies numériques. Dans cet entretien, Cyril Cosme et Philippe Marcadent reviennent sur les négociations qui ont conduit à son adoption, sur les principes qui la structurent et les évolutions réglementaires qu'elle peut impulser au niveau des États. L'entretien souligne également le rôle que l'OIT entend jouer dans les discussions internationales sur la gouvernance de l'IA, en faisant du travail et de ses acteurs un objet à part entière du débat.

Entretien réalisé par Vincent Mandinaud (Anact)

— L'OIT vient d'adopter une nouvelle convention sur le travail des plateformes[1]. Avant de préciser le contenu de ce nouveau texte pouvez-vous revenir sur le processus qui a abouti à son adoption ?

Philippe Marcadent : La négociation de cette nouvelle norme s'est inscrite dans le cadre de la production d'une norme internationale du travail *(Cf encadré)*. Ainsi son périmètre ne portait pas sur la régulation des activités des plateformes en général, mais couvrait le champ des opérations de ces plateformes qui ont un impact sur le travail, notamment sur l'accès au travail, à la qualification et sur les conditions de travail. Le processus de production des normes associe le Bureau international du travail (BIT) et les mandants de l'Organisation internationale du travail (OIT), c'est-à-dire les organisations d'employeurs, de travailleurs et les gouvernements. Dans le cas de cette négociation sur le travail des plateformes, la procédure mise en œuvre est appelée « double discussion », signifiant notamment que le contenu de la norme est à l'agenda de deux discussions de la Conférence internationale du travail (CIT).

Le processus d'élaboration des normes comporte plusieurs étapes. La première est la production d'un rapport qui dresse un état de la réglementation et des pratiques dans le monde sur le sujet. Ce rapport comprend un questionnaire destiné aux États membres pour identifier leurs besoins, les lacunes à combler et les orientations souhaitées. Et c'est sur la base des réponses à ce questionnaire que le BIT publie un second rapport qui est transmis à la première session de la CIT. Concernant les plateformes, la première discussion de la CIT a eu lieu en juin 2025 ; la seconde en juin 2026. À l'issue de la première session de la CIT, le Bureau produit un troisième rapport de synthèse qui fait le point sur les débats et les positions des mandants. Ce rapport est soumis aux mandants pour observation. Le quatrième rapport inclut des propositions de conclusions, c'est-à-dire des avant-projets de textes normatifs. Ce processus original vise à s'assurer de la pertinence des normes négociées, à garantir qu'elles répondent bien aux besoins tels qu'ils sont analysés et exprimés par les mandants. Il est à noter que le questionnaire envoyé aux États membres sur l'économie des plateformes a obtenu un nombre de réponses très élevé, probablement le plus élevé de toute l'histoire de l'OIT. Ceci témoigne à la fois de l'intérêt énorme qu'il y a sur le sujet, et de la nécessité pour beaucoup de pays de disposer, à travers les normes de l'OIT, d'une forme d'orientation sur les régulations possibles au niveau national de ces plateformes digitales et leur impact sur le travail. Dans les discussions de la CIT en juin 2025, il y a eu beaucoup de débats mais peu d'éléments véritablement actés. L'accord intervenu lors de la dernière CIT en est d'autant plus remarquable.

Sur la forme de l'instrument, la décision a été prise lors de la première discussion, après de longs débats, de produire une convention complétée par une recommandation, la convention pouvant avoir des implications légales dans les pays *(sur la distinction entre convention et recommandation, lire l'encadré)*. Essentiellement faute de temps, la dernière CIT n'a pu se pencher sur le projet de recommandation, finalement abandonnée.

Pour ce qui concerne la question des plateformes couvertes, ce fut évidemment un des sujets-clés des négociations. L'économie des plateformes est un ensemble très vaste et très varié. Cette diversité réside notamment dans la nature des services qu'elles proposent, les modèles économiques qu'elles adoptent, la nature de leurs relations avec ceux qui travaillent via ou pour ces plateformes, le degré d'utilisation et la sophistication de la technologie qu'elles utilisent, le caractère national ou international de leurs activités, etc. C'est, on le voit, un ensemble extrêmement divers. Comment capturer cette diversité dans une norme unique ? C'était l'un des points que les employeurs ont soulevés pour s'opposer initialement à la production normative, position minoritaire qu'ils ont finalement dû

1 ILC.114/Convention no 193 : Convention concernant le travail décent dans l'économie des plateformes (12 juin 2026) - https://www.ilo.org/sites/default/files/2026-06/ILC114-Convention%20193-%5BRELMEETINGS-260522-003%5D-Web-FR.pdf

abandonner. Définir, dans cette diversité, quelles plateformes couvrir représentait en tout état de cause un défi à la fois technique et politique. Une définition a été adoptée en 2025 après de longs débats très tendus, mais elle était très large et son interprétation variait significativement entre les mandants. Cette interprétation a été clarifiée lors de la CIT de 2026. Dans cette définition, l'expression « plateforme de travail numérique » renvoie à une personne morale ou, lorsque la législation nationale le prévoit, à une personne physique qui, au moyen de technologies numériques, utilisant des systèmes automatisés de prise de décision, organise et/ou facilite un travail exécuté par des personnes. Et indépendamment du fait que le travail soit exécuté en ligne ou en un lieu géographique précis.

Une autre question-clé en relation avec les systèmes couverts est celle des travailleurs couverts, et plus encore celle des protections à donner à ces travailleurs. La très grande majorité de ceux qui travaillent via les plateformes sont classés comme indépendants dans la plupart des pays du monde. Certains gouvernements, principalement en Europe, considèrent la requalification des travailleurs des plateformes numériques en salariés comme *la* solution pour leur accorder une protection. Cependant, la majorité des pays et territoires hors d'Europe ont opté pour une approche différente en étendant certaines protections du travail aux travailleurs des plateformes numériques indépendamment de leur statut d'emploi, c'est-à-dire qu'ils soient travailleurs indépendants ou travailleurs salariés. Ces pays protègent les indépendants – le plus souvent, il ne s'agit pas de tous les indépendants, mais des indépendants qui présentent un bas niveau d'autonomie ou un faible pouvoir de négociation dans la conduite de leur activité. Ceci est aussi un débat central qui est très fortement attaché à la définition des plateformes. La question est abordée dans la directive européenne de 2024 sur les conditions de travail dans le cadre du travail via une plateforme qui précise que ce ne sont pas toutes les plateformes, mais celles qui ont un rôle important dans l'organisation du travail, ce qui réduit d'une certaine manière l'autonomie des travailleurs. La convention couvre finalement tous les travailleurs, indépendants ou salariés, et garantit aux premiers l'accès aux droits fondamentaux internationalement reconnus, y compris en matière de santé et sécurité, l'accès à la protection sociale, ainsi qu'à un certain nombre de droits en matière de rémunération.

— Quelles sont les principales dispositions et enjeux auxquels répond cette Convention 193 finalement adoptée ?

Cyril Cosme : Cette nouvelle économie des plateformes concerne de nombreux emplois même si, pour être honnête, nous ne savons pas exactement combien. Des chiffres circulent néanmoins : 40 millions dans l'Union européenne, 85 millions en Chine par exemple et à l'échelle mondiale les chiffres varient entre 150 millions à 400 millions de personnes qui seraient des travailleurs sur ces plateformes. L'approche générale de la Convention n'est pas de freiner le développement des nouvelles technologies, ni celui de nouveaux modèles économiques. Elle vise à promouvoir, dans cette nouvelle économie, de meilleures conditions de travail et de vie pour les travailleurs. Donc on trouve cet équilibre, entre les intérêts des employeurs et des travailleurs, mais aussi l'intérêt général consistant à promouvoir le travail décent dans les plateformes qui génèrent de l'emploi, mais en même temps de la protection.

Commençons par rappeler le champ d'application de cette convention. Elle ne concerne qu'un type particulier de plateformes numériques : celles qui utilisent les technologies numériques et les systèmes automatisés pour organiser et ou faciliter le travail. Mais ce qu'il faut surtout retenir, c'est qu'elle couvre l'ensemble des travailleurs en lien avec ces plateformes, quel que soit leur statut, salariés, indépendants, formels ou informels, qu'ils travaillent chez eux en ligne, dans un lieu déterminé, dans tous les secteurs où ces plateformes se sont développées : livraison de repas, transport de passagers, service des soins, travail domestique. Une grande partie de ces travailleurs des plateformes travaillent en ligne depuis leur domicile et sont donc largement invisibles. Ils réalisent des tâches comme l'étiquetage de données par exemple, ou encore la modération de contenu et, pour beaucoup d'entre eux, ils sont

indépendants, freelancers et exécutent souvent des micro-tâches – « les travailleurs du clic », comme on les appelle parfois. Ensuite, il faut souligner que la Convention est basée sur des principes fermes mais relativement flexibles quant aux modalités de mise en œuvre à déterminer au niveau national. Prenons l'exemple de la classification du statut des travailleurs : salarié ou indépendant. La Convention n'impose pas que tous les travailleurs des plateformes numériques soient qualifiés de salariés. Elle laisse la marge de manœuvre et la liberté d'avoir différents modèles qui peuvent impliquer une plus ou moins grande proportion de salariés ou d'indépendants. Mais, là où la Convention établit quand même un principe, c'est que la classification doit se faire principalement sur la base des faits et donc des faits qui se rapportent à une relation d'emploi ou d'absence de relation d'emploi. Enfin, la convention apporte un socle de droits et garanties pour les travailleurs, à charge pour les États d'en assurer l'effectivité et l'exercice. Cela concerne les principes et les droits fondamentaux, y compris la sécurité et la santé au travail, la protection contre la violence et le harcèlement. Cela concerne également la rémunération, la protection sociale, la protection des données personnelles des travailleurs qui est très importante aussi en ce qu'elle protège les travailleurs contre une suspension ou désactivation de compte qui serait effectuée sur une base discriminatoire.

Comme pour l'ensemble des conventions de l'OIT, celle-ci prévoit que les États-membres qui la ratifient prennent les mesures nécessaires à sa mise œuvre effective, à travers la législation nationale, les conventions collectives et le contrôle de leur application par le juge. La convention prévoit également que les travailleurs des plateformes puissent avoir accès à des mécanismes de règlement des différends qui soient sûrs, équitables et efficaces et aussi à des moyens de recours et de réparation appropriés le cas d'échéant.

— Pouvez-vous préciser en quoi ce travail de plateforme a partie liée avec le développement de l'intelligence artificielle ? Et plus précisément, dans quels cadres « les petites mains » de l'IA, les personnes qui annotent les contenus, les nettoyeurs du Net et autres crowdworkers sont concernés ?

PM : Nombre de plateformes digitales recourent à l'intelligence artificielle notamment pour interagir avec les utilisateurs via les chatbots ou, en interne, pour faciliter le développement et les opérations. Comme dans beaucoup de secteurs, on observe un accroissement de l'utilisation de l'intelligence artificielle dans le fonctionnement des plateformes. La définition des plateformes couvertes qui a été retenue par la conférence fait référence à l'emploi de technologies numériques et de systèmes automatisés. Elle ne se réfère pas au type de technologie numérique utilisée. Les sujets traités portent sur l'effet des systèmes automatisés sur l'accès au travail, y compris la perte de travail, et sur les conditions de travail, quelle que soit la technologie qui est derrière. Ce qui importe dans le cas des plateformes est que des machines, en l'occurrence les systèmes automatisés, et non des humains, effectuent le suivi ou l'évaluation du travail ou génèrent des décisions qui impactent l'accès au travail et les conditions de travail – et ce, quel que soit le type de technologie utilisé par ces systèmes automatisés. Je souligne d'ailleurs que l'OIT n'a ni le mandat ni les compétences pour réguler la technologie en soi. Par contre, elle a mandat pour intervenir sur l'impact des évolutions technologiques sur le travail. Par exemple, la recommandation 169 énonce que les États membres devront s'efforcer d'assurer, par des mesures appropriées, que les effets négatifs des changements technologiques sur le travail soient éliminés. Mais la régulation de la technologie en soi ne relève pas de la responsabilité de l'OIT.

Quant à la question sur les travailleurs couverts, on revient à ce que j'ai dit précédemment à propos des définitions. Dans tous les cas, seuls ceux qui travaillent via une plateforme pourront être couverts. En l'état actuel, les crowdworkers, ceux qui exécutent des micro-tâches, y compris compiler ou coder des données utilisées par l'IA, sont concernés s'ils effectuent leur travail via une plateforme digitale.

— Pouvez-vous préciser qui participe à ces négociations ? S'agit-il seulement des représentants des États ou également des représentants des travailleurs et des employeurs ? De plus, cette économie dite « de plateforme » est dominée par de grands fournisseurs de technologies. Comment sont-ils eux-mêmes impliqués dans les négociations qui sont ouvertes ? Et s'ils sont impliqués, le sont-ils seulement en tant que fournisseurs de technologies, ou en tant que quasi-employeurs ? Quelle est la place des représentants des travailleurs, qu'ils soient salariés ou indépendants ?

CC : Le tour de table de ces négociations à l'OIT est tout à fait original dans le système multilatéral, parce que l'OIT est une organisation tripartite. Les gouvernements des États membres sont mobilisés – ils sont 187 aujourd'hui – mais aussi les représentants des organisations syndicales de travailleurs et les organisations d'employeurs de chacun de ces pays membres. Pour les partenaires sociaux, deux organisations jouent un rôle majeur. Il s'agit de la Confédération syndicale internationale, la CSI, pour les travailleurs, qui supervise et qui coordonne la délégation des travailleurs à la CIT, tandis que l'OIE, l'Organisation internationale des employeurs, joue ce rôle pour les employeurs. D'ailleurs, l'OIE a été créée en même temps que l'OIT, voici plus d'un siècle, en 1919, en lien avec la nécessité institutionnelle de s'organiser pour assurer la représentation des employeurs dans les organes de gouvernance de l'OIT. Car il n'existe pas dans tous les pays une organisation représentative des employeurs sur le plan interprofessionnel, dotée d'un mandat de négociation sociale. Dans beaucoup de pays, il n'y a pas l'équivalent du MEDEF. Aux États-Unis, par exemple, les négociations se passent dans l'entreprise. Il existe une instance nationale de coordination des employeurs américains, l'US Council for International Business, qui fait l'interface entre l'OIE et les structures patronales américaines, en particulier le réseau des chambres de commerce. Mais au-delà de l'entreprise, il n'y a pas de négociation collective en tant que telle, donc il n'y a pas de structure patronale représentative dotée d'un mandat pour négocier avec les syndicats.

Par ailleurs, au sein même du groupe des employeurs, comme au sein de celui des travailleurs, des divergences s'expriment. Et l'on constate que, dans le groupe des gouvernements, traversé par des clivages politiques et géopolitiques, les divergences sont encore plus importantes. S'agissant des divergences parmi les employeurs et parmi les travailleurs, elles sont traitées et arbitrées dans chaque groupe, lesquels groupes parlent d'une seule voix dans toutes les négociations à la CIT. Dans les rangs du groupe des employeurs, certains secteurs estiment qu'il existe des distorsions de concurrence du fait de l'essor de l'économie de plateforme, à l'origine de nouveaux modèles économiques. Du côté des travailleurs, on est loin d'avoir une homogénéité totale des statuts ; le salariat ne concerne que la moitié des travailleurs dans le monde. Une grande partie de ces travailleurs relèvent de l'informalité, c'est-à-dire qu'ils ne travaillent pas dans le cadre d'un statut formellement établi par la loi. Les travailleurs informels, ce sont en général des micro-entrepreneurs, des travailleurs indépendants. Ils sont sous-représentés de façon structurelle par les organisations représentatives, même si aujourd'hui, les organisations syndicales s'attachent davantage à leur faire une place. En tout état de cause, les normes internationales qui sont négociées selon ce processus ne sont pas limitées aux travailleurs salariés. Elles ont une portée universelle. C'est particulièrement vrai des droits fondamentaux. Un des enjeux de cette négociation sur le travail de plateforme était précisément de faciliter l'accès des travailleurs non salariés à un certain nombre de droits, tels que les droits fondamentaux, la protection sociale et les conditions de travail, ce que garantit effectivement le texte de la convention.

— L'année dernière, par une résolution adoptée par consensus à l'Assemblée générale de l'ONU, il a été décidé de constituer un groupe d'experts internationaux indépendants sur l'intelligence artificielle et d'établir un dialogue mondial sur la gouvernance de l'IA. Comment l'OIT est-elle associée ? Cela a-t-il alimenté les discussions liées à l'adoption de la convention sur le travail des plateformes ? Et dans le sens inverse, dans quelle mesure cette nouvelle convention est-elle susceptible de nourrir les discussions sur la gouvernance de l'IA et d'alimenter les travaux du groupe d'experts ?

CC : Beaucoup d'initiatives intergouvernementales sont menées autour de la bonne gouvernance de l'intelligence artificielle. La France compte parmi les pays les plus engagés. En témoigne la tenue à Paris du sommet sur l'intelligence artificielle en février 2025. Ou encore son rôle dans le lancement en 2020 du partenariat mondial sur l'intelligence artificielle avec le Canada, qui est hébergé aujourd'hui à l'OCDE. S'agissant de l'initiative qui a été prise au niveau des Nations unies et pour en comprendre la portée, il faut remonter à 2024, lors du Sommet du futur organisé en marge de l'Assemblée générale des Nations unies en septembre, au cours duquel a été adopté le Pacte pour l'avenir, qui visait à poser les jalons de la réforme de l'ONU à l'horizon 2030-2040. Ce Pacte posait notamment les principes devant encadrer le développement de l'intelligence artificielle, en particulier la promotion de systèmes d'intelligence artificielle, selon la formule consacrée « sûrs, sécurisés et dignes de confiance ». Et dans la droite ligne de l'adoption de ce Pacte numérique mondial, la résolution organise une plateforme de dialogue mondial sur la gouvernance de l'IA. Deux sessions de ce dialogue se sont d'ores et déjà tenues, la dernière à Genève, sous les auspices de l'Union internationale des télécommunications (UIT). D'un point de vue institutionnel, ce dialogue mondial a vocation à coordonner les institutions et agences des Nations unies – comme l'UIT, l'Organisation des Nations unies pour l'éducation, la science et la culture (UNESCO) ou encore l'OIT – afin de promouvoir la coopération internationale, de mettre en commun les meilleures pratiques, de favoriser des discussions transparentes, d'identifier la contribution des systèmes d'intelligence artificielle à l'atteinte des objectifs de développement durable (sous l'angle notamment du développement des infrastructures nécessaires dans les pays en développement, pour éviter un fossé digital accru à l'occasion du développement des systèmes d'intelligence artificielle). Le rôle de l'OIT dans ces espaces de dialogue et de concertation est de mettre en avant les enjeux qui touchent le monde du travail et ses acteurs. Notre enjeu est de faire en sorte que le travail ne constitue pas un angle mort dans ce débat. La gouvernance de l'IA ne doit pas être réduite à une pure entreprise de normalisation technique ou technologique, ou à la seule opportunité économique. On ne tirera profit des nouvelles technologies et de l'intelligence artificielle qu'en prévenant les risques qu'elles comportent, en se penchant sur les usages, y compris les usages au sein du monde du travail. Un groupe d'experts indépendants doit nourrir par ses travaux ce dialogue. On peut d'ores et déjà souligner l'importance de développer une expertise internationale indépendante, aussi bien vis-à-vis des gouvernements que des industriels. C'est important, un peu à l'image de ce qu'a pu faire le GIEC sur le climat. L'indépendance et l'excellence de l'expertise, de même que le leadership, sont les forces motrices des organisations internationales. Et c'est important de les préserver à l'heure où, justement, le multilatéralisme est critiqué et parfois menacé.

La convention 193 constitue incontestablement une étape importante dans ce processus normatif international autour des enjeux de l'économie des plateformes et de l'IA. À en croire les nombreuses sollicitations dont le BIT est l'objet depuis son adoption, cette convention répond à des besoins et des attentes parmi les mandants. Au fil des ratifications, elle servira de base aux réglementations que les pays sont amenés à adopter dans le but de préciser le statut et le régime du travail, préserver les droits fondamentaux, qu'il s'agisse de la liberté syndicale, de la santé, la sécurité au travail ou de l'accès à la protection sociale, d'orienter le management algorithmique. La convention constituera également une référence pour tous les travaux à venir au niveau multilatéral. Cette norme, potentiellement, imposera un objet « travail » en lien avec l'économie de plateforme, un objet à construire et à faire progresser au fur et à mesure que cette technologie évolue. En revanche, un lien pourrait être fait aussi entre ce groupe indépendant et le réseau des observatoires sur l'intelligence artificielle, réseau que coordonnent l'OIT et l'OCDE. C'est d'ailleurs l'une des conclusions du sommet de Paris de février 2025.

— L'IA connait des développements très rapides. Comment cette négociation mondiale intègre-t-elle cette dynamique de développement ? S'accorder aujourd'hui n'est sans doute pas facile : un éventuel accord à date ne risque-t-il pas de se retrouver très vite dépassé ? Comment gérez-vous cette problématique ?

PM : Le fait que la convention 193 ne précise pas le type de technologie digitale utilisé, mais fait seulement référence aux systèmes automatisés, quelle que soit la technologie adoptée, est un choix volontaire pour réduire l'impact des évolutions technologiques. Et c'était un choix à peu près unanime. À partir du moment où vous allez faire référence à une technologie particulière, si une autre technologie fait son apparition, la norme devient obsolète. Certaines personnes peuvent aussi se débrouiller pour échapper à la norme en utilisant une autre technologie que celle mentionnée. Nous avions néanmoins proposé la possibilité d'inclure dans les normes une clause concernant une procédure accélérée et simplifiée de leur révision pour prendre en compte les évolutions très rapides des technologies digitales. La très grande majorité des pays du monde et les travailleurs étaient favorables à cette inclusion. Les employeurs, en revanche, ont exprimé une très forte opposition à cela. La conférence en 2025 a décidé finalement de ne pas introduire cette clause. Dans le texte final, il n'y a pas de clause spécifique pour faciliter et accélérer la révision de la norme, notamment pour prendre en compte les évolutions technologiques.

— Vincent Mandinaud (Anact) : Pouvez-vous préciser la spécificité des instruments conçus à l'échelle mondiale à travers ce type de négociation ? Vous avez indiqué que c'est une négociation à vocation normative : en quoi cela se distingue-t-il d'une directive européenne, par exemple, tant dans la conception de ces instruments que dans leur mise en œuvre ? Comment ces instruments peuvent-ils constituer *in fine* des points d'appui pour les employeurs, les travailleurs, les représentants des travailleurs et des employeurs et les États ?

Cyril Cosme (OIT France) : Les normes de l'OIT relèvent de deux grandes catégories. Il y a celles qui découlent des conventions ; ces normes sont contraignantes. Les conventions, ce sont des traités au sens du droit international public. Cela signifie qu'elles doivent être ratifiées pour avoir un effet juridique en droit interne. Dès la ratification des conventions par le parlement, dans un pays comme la France par exemple, la convention intègre littéralement le droit national. La seconde catégorie, ce sont les recommandations, qui sont des instruments non contraignants. Ces recommandations peuvent être adoptées indépendamment des conventions, sur des sujets qui se prêtent moins à des réglementations opposables, contraignantes, comme l'emploi et le travail décent pour la paix et la résilience, qui font l'objet de recommandations. Souvent, elles accompagnent l'adoption de conventions, et peuvent aider et orienter l'application de ces dernières. Comme elles ne sont pas contraignantes, les recommandations ne sont pas opposables et ne font pas l'objet d'une ratification.

Il existe des différences notables avec le droit européen. D'abord, le positionnement du Bureau est loin d'être comparable à celui de la Commission européenne. Celle-ci a beaucoup plus de pouvoir à travers le monopole d'initiative, le pouvoir de refuser des amendements ou de retirer des propositions mises sur la table, par exemple, si l'évolution de la négociation ne lui convient pas. Pour le Bureau international du travail, qui assure le secrétariat de l'OIT, c'est un positionnement délicat parce que le pouvoir législatif appartient aux mandants. Le bureau doit se conformer à une stricte neutralité, même s'il assure le secrétariat de ces négociations pendant la conférence et peut formuler des propositions.

Ensuite, une directive européenne ne s'applique pas directement en droit national. Il faut une loi ou des règlements de transposition. La convention de l'OIT, elle, s'applique littéralement : dès la ratification, elle s'applique dans la forme que lui a donnée la négociation et telle qu'elle est écrite à l'issue de la négociation. Il n'y a pas de réécriture de la convention en droit interne. En France comme dans d'autres pays, la Constitution lui donne une valeur supra-légale, ce qui signifie qu'elle s'impose à la loi en cas de non-conformité avec cette dernière. Concrètement, le juge peut écarter la disposition législative nationale pour faire prévaloir la disposition d'une convention de l'OIT. Les règlements européens, à la différence des directives, s'appliquent littéralement en droit interne, un peu comme les conventions internationales, mais à la différence des conventions de l'OIT, ils n'ont pas besoin d'être ratifiés. Ces instruments sont donc très différents. En droit international public, les conventions, comme dans tous les traités internationaux, ne comportent que des obligations entre États, et ne créent pas d'effet juridique dans les rapports entre les particuliers. S'agissant des conventions de l'OIT, c'est un peu différent. En effet, ces conventions visent *in fine* à soutenir un ensemble de droits et d'obligations entre les travailleurs, les employeurs et l'État. C'est une logique tout à fait originale dans l'univers du droit international public et des normes internationales. Par conséquent, les normes de l'OIT constituent, et c'est leur spécificité, autant de points d'appui possibles pour les partenaires sociaux, dans les négociations collectives au niveau des branches ou des entreprises. On l'a vu notamment pour des conventions récentes, la convention 190, par exemple, sur le harcèlement et les violences au travail, qui a servi de référence dans différentes formes de dialogue social, s'agissant d'un sujet émergent sur lequel le droit national pouvait faire l'objet de certains compléments conventionnels.

QUELLES MÉTHODES POUR MESURER LES IMPACTS DE L'IA ?

Entretien avec Éric Fourboul et Gauthier Roussilhe

Éric Fourboul vient du monde du numérique, avec plus de 20 ans passés chez des éditeurs de logiciels fournissant des services numériques. À partir de 2019, il s'intéresse aux enjeux numériques et environnementaux, constatant un manque de données fiables. Pour travailler sur ces sujets, il co-fonde Boavizta, une association, et Hubblo, un bureau d'études, visant notamment l'approfondissement de deux grandes questions de recherche : quelle est l'empreinte écologique du secteur numérique ? Et la numérisation du monde sert-elle la transition écologique ?

Gauthier Roussilhe, issu de l'univers du design, a dirigé une agence de conception de produits numériques jusqu'en 2017. Il s'est ensuite spécialisé dans le numérique et l'environnement. Depuis 2017, il travaille sur ces questions, remontant la chaine de valeur et intégrant des connaissances d'autres secteurs. Également membre de Hubblo, il s'est attaché plus particulièrement à étudier les impacts environnementaux amont (extraction, fabrication) du secteur numérique, et depuis 2021, l'évaluation environnementale d'impact net de solutions énergétiques. Il a soutenu une thèse au RMIT de Melbourne, autour de l'intégration des enjeux environnementaux dans les politiques de numérisation, en adoptant une approche interdisciplinaire.

À rebours des discours qui présentent l'intelligence artificielle comme une solution naturelle à la transition écologique, cet entretien interroge les modes d'évaluation de ses impacts, notamment environnementaux. Éric Fourboul et Gauthier Roussilhe montrent les limites des approches centrées sur la seule empreinte carbone et plaident pour des analyses multicritères et conséquentielles intégrant les effets rebonds, les transformations des infrastructures et les dynamiques d'usage. Ils soulignent que les impacts de l'IA ne peuvent être compris indépendamment des modèles économiques qui soutiennent son développement, des infrastructures énergétiques qu'elle mobilise et des rapports de pouvoir qui orientent sa diffusion. L'entretien met également en lumière les enjeux que ces transformations soulèvent pour le travail, le dialogue social et la capacité des acteurs à orienter les choix technologiques. Il appelle ainsi à renouveler les cadres d'évaluation et de régulation de l'IA en articulant les dimensions écologiques, économiques, sociales et politiques.

Entretien réalisé par Caroline Datchary (Université Toulouse - LISST), Julien Gros (Université Toulouse - LISST) et Vincent Mandinaud (Anact)

— Caroline Datchary : Dans vos travaux sur les mesures d'impacts écologiques du secteur numérique, vous insistez sur la nécessité de distinguer les types d'IA, les chaines techniques et les modes d'usage. Vous critiquez par ailleurs les approches qui se limitent à évaluer l'empreinte carbone, et celles qui évaluent l'impact de manière individuelle ou attributionnelle de l'IA générative et omettent, ce faisant, les conséquences ultérieures ou effets rebonds de leur usage. Pourquoi ces distinctions sont-elles cruciales pour comprendre les véritables enjeux écologiques ?

Éric Fourboul : L'enjeu essentiel est de développer des approches multicritères. C'est la différence entre l'ACV (analyse du cycle de vie), qui est multicritère, et l'empreinte carbone, qui est monocritère. Les deux méthodologies traitent de l'ensemble du cycle de vie, de la fabrication, voire de l'extraction des ressources, jusqu'à la fin de vie. Pour autant, l'empreinte carbone, qui pilote toutes les décisions environnementales, ne mesure que le CO_2, en lien avec le potentiel de changement climatique. Or, si l'on se focalise sur un seul indicateur, on passe à côté des transferts de pollution. Prenons l'exemple des voitures et de la comparaison entre diesel, essence et électrique. Sur le critère du carbone, l'essence est la plus mauvaise ; sur l'indicateur des particules fines, qui est vraiment l'élément le plus toxique pour l'humain à court terme, c'est le diesel ; pour tout ce qui est minéraux, ressources abiotiques, c'est l'électrique. Chacun a son versant toxique. L'ACV propose une vue sur l'ensemble de ces indicateurs. Mais l'ACV, de type attributionnel, a aussi des limites. L'ACV définit une sorte de quote-part : quand on évalue l'ACV d'un service numérique, on évalue une responsabilité, une part de l'ensemble des pollutions du monde. C'est intéressant pour évaluer l'efficacité sur les impacts environnementaux. Mais cela ne prend pas en compte les effets rebonds. Or, on sait qu'ils n'ont jamais été aussi nombreux. La limite de l'ACV attributionnelle, c'est donc de ne pas traiter des conséquences et d'en rester à cette dimension de l'efficacité. La numérisation d'une activité la rendra la plupart du temps plus efficace à usage constant, mais elle développera davantage d'usages qui auront des effets induits. Par exemple, une publicité numérique aura moins d'impact potentiel de changement climatique qu'une publicité papier, mais la numérisation du secteur entraine une augmentation du volume de publicité global, facilité par la multiplicité des espaces de diffusion potentiels. D'où l'intérêt de travailler sur le conséquentiel ou les impacts nets afin d'appréhender les conséquences de la numérisation, notamment.

Gauthier Roussilhe : Qu'on soit dans une approche attributionnelle ou conséquentielle, l'enjeu central, c'est la construction des indicateurs d'impact. Si des indicateurs tels que les équivalents CO_2 pour le potentiel de réchauffement climatique, ou l'ADPE (Abiotic Depletion Potential-Elements) pour l'épuisement des ressources minérales et métalliques, sont largement utilisés, leur élaboration, leur contenu et ce qu'ils mesurent réellement sont rarement interrogés. L'ACV est un outil intéressant, mais elle ne saurait être considérée comme une mesure exhaustive ou définitive des impacts. Elle comporte des limites structurelles liées à ses hypothèses et à ses conventions d'usage. Par exemple, l'indicateur de potentiel de réchauffement climatique repose généralement sur un horizon de 100 ans (GWP 100) qui inclut un temps de voyage de la molécule jusque dans l'atmosphère (qui peut prendre jusqu'à plusieurs dizaines d'années) où elle atteint son potentiel de réchauffement maximum. Quand un lanceur spatial émet dans l'atmosphère, il n'y a plus de temps de voyage. Par exemple, dans le cas du lancement de satellites en orbite basse, les émissions de CO2 sont directement injectées dans l'atmosphère, ce qui invalide l'indicateur d'impact de réchauffement sur 100 ans, qui ne fonctionne plus pour du satellitaire. Cela montre bien que les indicateurs ne sont pas universellement pertinents et doivent être complétés, voire remplacés par d'autres représentations, quantitatives ou pas. Par ailleurs, l'approche attributionnelle repose sur un système supposé stable, dans lequel les impacts sont répartis entre différents acteurs. Or, dans la réalité, les systèmes sont rarement figés, ce qui limite l'usage prospectif de l'ACV attributionnelle. À l'inverse, l'approche conséquentielle relève d'une logique fondamentalement politique et prospective : elle vise à anticiper les effets futurs de décisions prises aujourd'hui. Elle s'apparente à la planification publique, qui cherche à atteindre un objectif donné à un horizon déterminé en activant plusieurs leviers d'action. Cette approche est aussi plus complexe et remet en cause l'illusion de la quantification neutre et purement scientifique. Elle oblige

à intégrer des dimensions plus difficiles à modéliser, comme les comportements, les changements macroéconomiques, les changements structurels. En revanche, elle est plus adaptée à l'analyse des choix collectifs et des trajectoires de transition. Et elle contraint aussi à l'interdisciplinarité pour croiser différents domaines (énergie, sciences du comportement, sciences sociales, macroéconomie) afin de modéliser des événements ou des chaines causales complexes. Mais justement, l'intérêt du conséquentiel est de réintégrer de la complexité au sein de modèles qui, par construction, tendent à simplifier le réel pour rester activables. Cette approche s'appuie sur des catégories d'effets bien identifiées : effets d'optimisation de l'efficacité, effets de substitution ou de complémentarité, effets rebonds directs, indirects et macroéconomiques, ainsi que changements structurels. L'enjeu est d'évaluer si une innovation remplace effectivement une pratique existante ou s'y ajoute, si les gains d'efficacité conduisent à une augmentation des usages, ou encore si les ressources libérées sont réinvesties ailleurs. Cela donne une vision plus riche et plus dynamique des systèmes étudiés. Plus spécifiquement, l'analyse conséquentielle de l'impact net oblige à dépasser le périmètre initial du système étudié. En examinant les effets indirects et systémiques, elle révèle fréquemment des conséquences inattendues, y compris pour les acteurs directement concernés. Des travaux que nous avons menés sur le télétravail, les intrants agricoles, le covoiturage, les réseaux énergétiques, le transport, ont ainsi mis au jour des effets jusque-là invisibles, parfois connus seulement de certains acteurs isolés[1]. Cette démarche relève avant tout d'un protocole d'enquête, dans lequel la quantification constitue un support d'exploration plutôt qu'une fin en soi.

Mais la grande différence entre l'ACV – qu'elle soit attributionnelle ou conséquentielle – et l'analyse d'impact net, c'est que l'ACV répond plutôt à des enjeux de régulation, notamment juridiques, liés aux obligations des entreprises (sur leurs produits, par exemple) et à la conformité de leurs engagements environnementaux. Les analyses d'impact net, en revanche, visent davantage à éclairer les décisions d'investissement, dans une logique financière plutôt que réglementaire. Par exemple, pour la Commission européenne, elles servent à identifier les activités numériques compatibles avec la taxonomie verte et à lever des fonds ESG[2], en partenariat avec les banques pour orienter leurs investissements. À la Banque mondiale, l'enjeu est similaire : investir dans des solutions climat à rendement financier et décarbonant – ce qui est complexe pour le numérique. Ces analyses se font *ex ante*, dans le temps de l'investissement, contrairement aux évaluations postérieures traditionnelles, comme l'ACV, qui est généralement plutôt a posteriori. L'objectif est de fournir des critères rapides pour déterminer si une solution numérique est a priori positive ou négative pour le climat.

— Julien Gros : Vous évoquez également une forme d'immaturité du sujet, en particulier parmi les acteurs politiques, syndicaux, entrepreneuriaux. Comment cette immaturité se manifeste-t-elle ? Est-elle liée au fait que les repères disponibles sont encore trop instables ? Ou est-ce que la question « l'IA aide-t-elle ou nuit-elle à la transition écologique ? » est une question mal posée ?

Gauthier Roussilhe : En effet, la question « est-ce que la numérisation sert à la transition ? » est mal posée. Tout d'abord, elle repose sur des notions insuffisamment définies. La « numérisation » recouvre une pluralité de processus très hétérogènes. Or, dans le débat public, le terme sert souvent de raccourci pour désigner un modèle particulier : celui du développement des services numériques fondé sur une forte logique d'investissement capitalistique, à l'image des grandes plateformes issues de la Silicon Valley. D'autres formes de numérique, comme Wikipédia et OpenStreetMap, relèvent pourtant de dynamiques très différentes. De la même manière, la notion de « transition » demeure floue : transition climatique, réduction des émissions de CO2, préservation de la biodiversité, engagements territoriaux ou nationaux ? L'absence de définition précise explique le succès de cette question, qui engage peu

1 https://librairie.ademe.fr/#0379/embedded/m=and&q=IT4green

2 ESG : environnemental, social et de gouvernance. Les critères ESG sont des critères extra-financiers qui permettent d'évaluer un acteur économique en dehors des critères financiers habituels que sont la rentabilité, le prix de l'action et les perspectives de croissance.

et autorise des prises de position générales sans réel contenu opératoire. Ce flou est renforcé par la faiblesse de la pensée critique des techniques. Le débat oscille entre deux pôles caricaturaux : un techno-optimisme naïf qui fait des technologies – notamment de l'IA – des solutions miracles (« l'IA va nous sauver du changement climatique ») et une technocritique d'arrière-garde qui, dans ses pires formes, peut être réactionnaire, et qui va défendre une position morale surplombante de rejet technologique. Faute d'une pensée critique féconde, capable d'analyser concrètement les effets sociaux, environnementaux et politiques, le débat demeure généralement stérile, même si de nouvelles formes de technocritiques se reconstituent en France (Hubert Guillaud, le Mouton Numérique, etc.). Par ailleurs, la sacralisation de notions comme « le progrès » hier ou « l'innovation » aujourd'hui a contribué à verrouiller la discussion. Même dans les régimes démocratiques, il n'existe pas de véritable démocratie technique : les choix technologiques structurants échappent largement à la décision collective. Face à ces impasses, une autre démarche consiste à suspendre les présupposés, à ne pas postuler a priori les effets positifs d'une solution technique mais à les évaluer empiriquement. Cette approche par l'enquête révèle des impacts inattendus et fait émerger de nouvelles questions. Elle souligne surtout le caractère situé des savoirs : les effets d'une technologie, comme l'implémentation d'un data center sur un territoire, ne peuvent être analysés uniquement à l'échelle macro, sans prendre en compte les contextes territoriaux et locaux dans lesquels ils s'inscrivent.

Éric Fourboul : En effet il y a un enjeu à développer notre capacité collective à déconstruire l'innovation comme positive a priori. C'est vrai des institutions publiques, mais c'est vrai aussi de tout un chacun. Cela dépasse la question du numérique, si ce n'est que le numérique, aujourd'hui, représente le symbole absolu d'innovation, de croissance, de modernité, etc. Par ailleurs, le lien entre le numérique et l'environnement est marqué par une forte immaturité des connaissances, sur le plan tant académique que pratique. Contrairement à d'autres secteurs étudiés de longue date, le champ « numérique et environnement » constitue encore un angle mort de la recherche et de l'action publique. Cette lacune se traduit concrètement par une méconnaissance des infrastructures elles-mêmes : en France, il n'existe pas de recensement fiable du nombre de data centers, y compris chez les autorités publiques et les gestionnaires de réseaux énergétiques. Cette situation s'explique notamment par la reconnaissance récente de cette activité, dont la classification administrative (code NAF) n'a émergé qu'à la fin des années 2010. Finalement, il y a peu d'activités de type industriel où l'on a autant d'inconnues. L'autre élément autour de l'immaturité est lié au discours largement promu par les acteurs du numérique, présentant la numérisation comme intrinsèquement bénéfique pour la transition environnementale. Ce récit, appuyé par des stratégies de communication et des méthodologies d'évaluation discutables, a largement imprégné les représentations collectives. Dans un contexte où le numérique concentre une puissance économique et financière considérable, cette asymétrie entre la faiblesse des connaissances disponibles et la force des discours diffusés favorise des raccourcis analytiques, y compris au sein des sphères décisionnelles publiques et privées.

Gauthier Roussilhe : Le concept de dématérialisation en économie, désormais associé au numérique, apparait en réalité dès les années 1960, où il désigne le passage d'une économie industrielle à une économie de services. Ce n'est que secondairement que cette notion est transposée au numérique, celui-ci produisant principalement des services, notamment dans le contexte de l'économie états-unienne. Dès les années 1970, la question des effets environnementaux de ces transformations est posée, notamment à travers les premiers travaux sur le télétravail, sans que des réponses définitives n'aient émergé. Un tournant majeur s'opère au début des années 2000, quand l'industrie s'empare de ce débat et impose progressivement un cadrage sémantique présentant le numérique comme

une solution plutôt que comme un problème[3]. Ce cadrage façonne durablement les imaginaires sociotechniques, y compris dans le monde du travail. Les acteurs publics ont perdu la main sur la construction des imaginaires et se contentent largement aujourd'hui de les instituer, de les stabiliser sans exercer de véritable regard critique. Bien que nombre d'acteurs institutionnels expriment des doutes quant aux bénéfices environnementaux du numérique, ils continuent à le financer faute d'alternatives perçues, maintenant ainsi une croyance en son caractère inévitable.

— Vincent Mandinaud : Pouvez-vous expliciter les liens entre le déploiement de l'IA, notamment générative, et la question des infrastructures, notamment énergétiques ?

Éric Fourboul : Concernant la dimension énergétique, je pense qu'il y a une vraie bascule. L'IA, et même l'IA générative, ne sont pas des technologies nouvelles. En revanche, cette bascule a lieu parce que les capacités de calcul et le volume de données sont incomparables à ce qui existait par le passé. Dans les data centers, les serveurs sont organisés en baies[4]. Jusqu'en 2020, la puissance moyenne par baie était de l'ordre de 6 à 8 kW. Aujourd'hui, avec les infrastructures de dernière génération, notamment celles intégrant des GPU NVIDIA[5], on atteint des puissances pouvant aller jusqu'à 600 kW. Alors qu'avant 2020, la puissance doublait tous les cinq ou six ans, on observe dorénavant des croissances beaucoup plus importantes, proches d'un facteur 100 en quelques années. Une telle densité électronique transforme profondément le design des data centers, en particulier les systèmes de refroidissement. Ces évolutions placent l'énergie au cœur des enjeux du numérique, dans des proportions inédites par rapport aux débats antérieurs sur la consommation des data centers, avant l'essor de l'IA générative. Le data center « Colossus » de Musk, situé dans le Tennessee, illustre cette bascule : avec une puissance visée de plus de 1,5 gigawatt, il dépasse la consommation électrique d'une ville comme Marseille et représente environ 60 % de la puissance des data centers en France. L'énergie devient ainsi un enjeu central du numérique, au point que les grands acteurs se dotent désormais de « chief energy officers ». Autre exemple : en Irlande, où la majorité des data centers sont concentrés autour de Dublin, du fait de politiques de défiscalisation, les projections de l'IEA[6] indiquent qu'ils pourraient représenter 30 % de la consommation électrique du pays en 2030. Face à la saturation du réseau, certains projets ont été conditionnés à l'autoproduction d'énergie, ouvrant la perspective que des acteurs comme Google deviennent également des acteurs du secteur énergétique. Cette évolution soulève la question d'une possible dérégulation et d'une prise de contrôle partielle des infrastructures énergétiques par les acteurs du numérique. L'IA générative marque donc un point de bascule dans la relation « énergie et numérique », posant des enjeux majeurs de dépendance, de concentration des acteurs et de vassalisation des infrastructures énergétiques.

Gauthier Roussilhe : Jusqu'à l'arrivée de l'IA générative, les acteurs du numérique se conformaient à la temporalité des opérateurs d'infrastructures énergétiques. Les délais de raccordement, parfois de cinq à sept ans, constituaient le nœud bloquant de l'ensemble de l'infrastructure, et les opérateurs du numérique s'y adaptaient. Maintenant, c'est terminé avec l'essor de l'IA générative, dans un contexte de bulle spéculative visant à déployer le plus rapidement possible des capacités de calcul pour entrainer des modèles et fournir des services à une masse grandissante d'usagers. Cette dynamique impose la création d'« AI factories », fondées sur des infrastructures très denses en puissance de calcul, donc en puissance électrique. Ne pouvant plus attendre les calendriers des énergéticiens, certains acteurs adoptent des stratégies d'autoproduction, à l'image de projets reposant sur des turbines à gaz, des génératrices, des batteries, constituant ainsi des silos énergétiques. Cette énergie

3 Voir Roussilhe, G. (2025), "A Tale of Twin Transitions: How Did We Come to Think That Digitization and Sustainability Go Hand in Hand", in *The Palgrave Handbook of Cybersecurity, Technologies and Energy Transitions* (p. 1-14), Palgrave Macmillan, Cham.

4 Structure métallique conçue pour héberger divers équipements électroniques (serveurs, routeurs, etc.).

5 Carte graphique permettant un rendu et un traitement graphique puissants, notamment pour le montage vidéo, les jeux vidéo et d'autres opérations informatiques complexes.

6 International Energy Agency, ou Agence internationale de l'énergie.

est majoritairement carbonée, le recours massif et rapide aux énergies renouvelables étant irréaliste à cette échelle et dans ces délais. C'est donc principalement du gaz. L'industrie numérique exerce une pression croissante sur l'industrie de l'énergie pour augmenter les investissements, multiplier les raccordements, les lignes à haute tension, accroitre les capacités, d'autant qu'il y a une pénurie de transformateurs (déjà très largement mobilisés par les infrastructures d'IA). C'est l'une des hypothèses sur lesquelles je travaille en ce moment : si la demande électrique liée à l'IA n'atteint pas les niveaux anticipés, une partie de ces infrastructures pourraient rester sous-utilisées, laissant les opérateurs de réseaux fortement endettés sans retour sur investissement. Cette situation pourrait entrainer une hausse des tarifs de l'électricité afin de compenser ces investissements. Le risque est alors une propagation de la bulle financière et foncière de l'IA vers l'économie réelle par le vecteur énergétique. C'est dans ce mécanisme que se joue une possible vassalisation des infrastructures énergétiques, dont les effets toucheraient l'ensemble des usagers.

— Vincent Mandinaud : Quelles sont les implications de ces dynamiques pour les employeurs et les travailleurs de France ? Peut-on imaginer que l'ampleur des investissements dans les infrastructures serve à légitimer, voire à encourager une généralisation accélérée des usages de l'IA, notamment dans le monde du travail, à travers des discours promotionnels destinés à justifier ces choix d'équipement ?

Éric Fourboul : De mon point de vue, il y a deux scénarios. Le premier est celui d'une bulle spéculative : l'IA générative consomme beaucoup plus d'énergie que d'autres services numériques. Le cout d'un échange avec ChatGPT est supérieur à celui d'une requête sur un moteur de recherche classique. Il existe donc un véritable cout énergétique et infrastructurel, dans un contexte où les services sont largement gratuits et où personne ne gagne encore réellement d'argent, selon une logique déjà observée dans l'histoire du numérique. Le second scénario est celui d'une dépendance progressive : malgré l'absence actuelle d'un socle d'études académiques solides, de nombreux utilisateurs ont le sentiment de gagner en efficacité, notamment dans le monde professionnel. L'hypothèse est alors qu'un modèle payant finira par s'imposer une fois les usages installés et l'addiction créée. Les enjeux sont d'autant plus importants que les plus grosses capitalisations boursières mondiales investissent massivement dans l'IA, dans une logique de type « too big to fail ». Les investissements dans les data centers des GAFAM ont ainsi été multipliés par quatre entre 2020 et 2024, alors qu'ils étaient déjà très élevés en 2020. Faute de recul et d'évaluations rigoureuses des gains d'efficacité réels, l'issue reste incertaine entre bulle spéculative et installation durable de l'IA perçue comme socle indispensable des entreprises.

Gauthier Roussilhe : Ces deux scénarios ne sont pas mutuellement excluants. Il peut y avoir à la fois une bulle et une propagation via l'IA générative. Le paradoxe, c'est que, malgré des entreprises numériques aux niveaux très élevés de capitalisation mais aussi de burn rate, l'usage de l'IA reste trop faible pour atteindre la rentabilité attendue par les actionnaires. Même gratuite et déjà largement déployée, l'IA n'est pas assez utilisée, d'où le forçage pour augmenter le nombre d'usagers et l'intensité d'usage, par exemple avec des prompts ou la vidéo générée par IA. Il faudra encore des milliards d'utilisateurs pour que l'IA devienne financièrement rentable. C'est le paradoxe de l'IA : on ne l'utilise pas assez pour qu'elle soit rentable. Et dans le modèle des capital-risqueurs siliconiens, c'est de moins en moins la rentabilité qui compte, mais plutôt la croissance. Une entreprise qui croît augmente sa valeur. Elle peut alors être revendue plus cher et doubler ou tripler la mise des investisseurs initiaux. En cas de bulle, seuls quelques acteurs avec des marchés captifs, comme Microsoft, Google et Meta, survivront, tandis que beaucoup d'argent et d'entreprises seront perdus en route. Mais les technologies de base de l'IA générative – les GPU, les transformers en tant que réseaux neuronaux et PyTorch comme langage – resteront utilisées, et elles peuvent l'être indépendamment du modèle économique actuel qui est particulier. Il est possible d'imaginer des usages nettement plus sains. Par

exemple, je peux entrainer un modèle sur le corpus documentaire de mon université uniquement, et imaginer que chaque étudiant, quand il arrive, a un nombre de tokens à dépenser pour réaliser des recherches documentaires compatible avec le budget de l'université en énergie.

Ceci étant dit, on regarde beaucoup ce que l'IA fait au monde du travail, mais assez peu ce que le monde du travail fait à l'IA. Or, il est intéressant d'observer comment le monde du travail va s'adapter à l'IA et comment l'IA s'adapte au monde du travail actuel. L'IA générative constitue un outil puissant qui intervient dans un monde du travail fragile et qui porte le risque d'amplifier cette tendance. Par exemple, elle peut être un formidable outil de négociation des salaires à la baisse, s'il est considéré qu'un agent d'IA peut remplacer un employé humain ou qu'un stagiaire peut suffire pour corriger les résultats produits par l'agent d'IA. Et cela vient alors déséquilibrer un peu plus le dialogue social.

— Caroline Datchary : Pourquoi les acteurs du monde du travail apparaissent-ils si faibles par rapport aux acteurs technologiques ? Comment se joue ou se rejoue le forçage de l'usage de l'IA dans le monde du travail ? Et en quoi vos travaux sur les critères d'évaluation pourraient-ils les aider à affronter les enjeux soulevés par le déploiement de ces systèmes ?

Gauthier Roussilhe : Lors d'une conférence aux Assises de la traduction littéraire, je suis intervenu sur les effets de l'IA sur le métier de la traduction. En partant de questionnements – quels étaient déjà les problèmes prédominants dans votre monde du travail ? Est-ce que vous étiez valorisé par les éditeurs ? Comment les éditeurs négociaient-ils votre travail de traducteur ? Quels étaient vos leviers dans la négociation ? Comment les auteurs percevaient-ils votre travail ? –, on constate que l'IA ne fait qu'exacerber des difficultés déjà anciennes du métier de traducteur (faible valorisation par les éditeurs, marges réduites de négociation, par exemple). L'arrivée de l'IA pose alors une question centrale : pourquoi passer par un traducteur si un modèle peut travailler à sa place ? Cela renvoie directement à la valeur de la traduction et du traducteur, une valeur bien connue des professionnels mais de moins en moins monétisable. Il s'agit d'une renégociation générale de la valeur des tâches et des activités humaines. La question devient alors : comment redéfinir et augmenter la valeur perçue du travail humain à l'ère de l'IA, et quelles nouvelles clés de mise en valeur peuvent encore être défendues ? Et par rapport à nos travaux, finalement, c'est la même question : est-ce qu'une analyse environnementale du numérique a de la valeur ? Le numérique, et plus encore l'IA, se déploient dans un contexte de quasi-absence de régulation, contrairement à la plupart des autres industries. Cette propagation rapide de l'IA a en outre bénéficié d'un timing particulier, marqué par une forte vulnérabilité mentale des populations, notamment après le Covid.

Éric Fourboul : Le problème dépasse le numérique et relève presque d'une dimension philosophique. Le monde du travail est depuis longtemps structuré autour de l'ultra-performance, avec des effets de violence bien documentés. L'IA apparait comme une couche supplémentaire, voire suprême, qui accentue cette logique de performance au détriment du bien-être humain. Elle s'inscrit dans un continuum : on n'a pas attendu l'IA pour constater que ces gains d'efficacité ne produisent pas que de la « zénitude ». Mais son déploiement intervient aussi dans un monde plus instable, où des contraintes géopolitiques et matérielles pourraient, à terme, freiner cette fuite en avant.

Gauthier Roussilhe : On suppose souvent que l'IA se substitue simplement à des tâches que l'on faisait déjà soi-même, ce qui est une hypothèse discutable. Certes, elle permet des gains de temps par substitution, mais elle entraine surtout la création de nombreuses tâches nouvelles qui n'existaient pas auparavant. Elle rend possibles des activités autrefois externalisées ou impossibles, comme produire soi-même un logo sans recourir à un graphiste. Le résultat n'est pas forcément une hausse nette de productivité, mais une intensification du travail et une charge accrue pour les individus, car les personnes font et vont faire beaucoup plus de tâches qu'avant et, au final, certainement travailler beaucoup plus. Mais c'est l'histoire du numérique et du progrès de manière générale quand il ne

s'accompagne pas de progrès sociaux. Alors, dans ce contexte, quel est l'enjeu de l'évaluation ? Bien sûr, cela sert à y voir plus clair ; si on est en aveugle, on ne prend pas les bonnes décisions. Mais on peut aussi considérer que ce n'est pas la voie systématique. Prenez Airbnb : il serait extrêmement complexe de mener une analyse environnementale conséquentielle de son déploiement et de ses usages. Mais en a-t-on besoin ? On sait déjà que c'est problématique d'un point de vue social et d'un point de vue foncier. On n'a pas besoin de l'analyse environnementale pour justifier le fait qu'il faut réguler.

— Julien Gros : Vous soulignez que la régulation des technologies numériques, et en particulier des IA génératives, intervient trop tard, une fois la dynamique technologico-économique déjà engagée. Faut-il dès lors inverser la logique en intégrant, dès la conception et l'orientation des usages, les enjeux écologiques et sociaux, compte tenu du caractère durable des infrastructures mises en place ?

Gauthier Roussilhe : Le débat autour de l'IA oppose innovation par régulation et innovation par laisser-faire, mais nos institutions, surtout gouvernementales, sont de moins en moins opérantes pour encadrer ces questions. Parallèlement, le changement climatique est largement ignoré dans les discours sur l'IA, avec des projections optimistes souvent peu fiables. L'IA générative s'impose dans un contexte de vulnérabilité sociale et d'hiver des politiques environnementales, amplifiant un biais collectif vers l'inaction. Les effets financiers et énergétiques risquent d'être considérables avant toute correction de marché. En résumé, le timing et l'absence d'institutions crédibles rendent la régulation de l'IA difficile avant que la totalité de ses impacts ne se manifestent.

LA NÉCESSITÉ D'UNE RÉGULATION DE L'IA AU TRAVAIL

Entretien avec Franca Salis-Madinier

Franca Salis-Madinier est secrétaire nationale CFDT CADRES et membre, depuis 2015, du Comité Économique et Social Européen (CESE), vice-présidente du Groupe Travailleurs. Elle a été rapporteure au sein du CESE de plusieurs avis concernant les transitions numériques et l'intelligence artificielle. Franca Salis-Madinier a aussi été membre de la commission gouvernementale sur l'IA qui a livré un rapport au Président de la république en mars 2024[1].
Dans cet entretien, Franca Salis-Madinier, s'appuyant sur son expérience au sein de la Commission IA et au sein du Comité Économique et Social Européen, plaide pour une régulation européenne spécifique de l'IA au travail. Elle insiste sur plusieurs recommandations clés : renforcement de l'information et de la consultation des représentants du personnel en amont de l'introduction d'outils IA, développement de la formation à l'IA, garantie d'un contrôle humain effectif sur les décisions automatisées et explicabilité de celles-ci, prévention des risques liés au management algorithmique. Elle met également en évidence les limites du cadre réglementaire actuel, notamment du règlement européen sur l'intelligence artificielle (AI Act), qui encadre principalement les systèmes d'IA comme produits sans traiter suffisamment leurs effets sur les relations de travail.

Entretien réalisé par Vincent Mandinaud (Anact)

1 « IA : Notre ambition pour la France », Rapport, Commission de l'Intelligence Artificielle, Mars 2024.

— Vous avez participé à la commission IA qui a rendu un rapport au président de la République voici plus de deux ans. Vous étiez la seule représentante des salariés dans cette commission. Pouvez-vous nous rappeler les principales avancées de cette commission sur les liens entre le développement de l'intelligence artificielle et les conditions de travail ?

Franca Salis-Madinier : J'étais effectivement la seule représentante des salariés au sein de cette commission, constituée de 15 membres et qui a rendu un rapport au président de la République en mars 2024. Il me semble que ce rapport est important pour les partenaires sociaux, car il apporte des recommandations sur la manière de déployer cette technologie qui est, comme le reconnait ce rapport, une technologie de rupture.

L'une des recommandations pose le dialogue social comme la pierre angulaire du déploiement de l'IA dans les entreprises et les administrations. Cette recommandation de principe est ensuite précisée. Notamment elle précise que les dialogues social et professionnel doivent porter sur les usages par la co-construction avec les travailleurs. Les recommandations identifient aussi les risques associés au déploiement de ces technologies et les moyens pour les réduire. D'autres mesures prioritaires sont préconisées dans le rapport. Par exemple : la nécessité de décliner en France l'accord européen sur la transformation numérique, signé en 2020. La France est très en retard, car cet accord aurait dû être transposé en deux ans, timing habituel de la transposition des accords européens. Dans ce texte adapté au niveau national, il sera important d'intégrer un volet sur la formation et la sensibilisation des travailleurs à l'intelligence artificielle.

Autre recommandation-clé que l'on retrouve dans les règlements européens : la garantie de respecter les codes du travail et le règlement général sur la protection des données (RGPD) sur les systèmes de supervision et de ceux qui mesurent la performance des travailleurs. En associant à cela l'obligation d'un droit effectif à la décision humaine. Cette commission recommande en effet que les voies de recours à un humain soient définies en cas de décision prise par la machine qui comporte des conséquences importantes pour les travailleurs comme par exemple un licenciement ou le refus d'une promotion, garantissant ainsi à l'humain sa capacité à décider et avoir le dernier mot vis-à-vis de la machine.

Plusieurs recommandations portent sur la prévention des risques pour la santé et la sécurité liés aux usages des IA, sur la nécessité de renforcer l'information, la consultation et la transparence lors de l'introduction des outils d'IA. Cela renvoie explicitement au renforcement du code du travail, notamment à l'article L2312-38, avec la nécessité d'inclure une clause explicite sur cette obligation en cas de déploiement des systèmes d'IA.

La formation est aussi une recommandation essentielle du rapport, avec la question associée de son financement. La formation, sensibilisation et acculturation large des citoyens, des travailleurs et des employeurs est un enjeu fort. C'est urgent, parce que cette technologie se déploie très rapidement : la rapidité de ses déploiements est l'une de ses spécificités, donc la formation doit être massive et accélérée. Ensuite, on peut citer le besoin d'un investissement spécifique pour la formation dans le champ professionnel, c'est-à-dire pour les partenaires sociaux, autant les syndicats de travailleurs que des employeurs. Là aussi, le besoin est important et urgent.

— Vous avez évoqué l'accord-cadre européen de 2020 sur la transformation numérique des organisations. Il y a aussi eu un accord-cadre européen en 2022 spécifique aux administrations centrales et fédérales. Ces deux accords intègrent des préoccupations autour de l'intelligence artificielle. Vous avez par ailleurs évoqué la commission IA du gouvernement français, à laquelle vous avez participé, et dont a émergé un rapport. Ces trois documents existent. Comment expliquer les difficultés à progresser concrètement sur le sujet, à s'approprier les accords européens sur le territoire national et à traduire concrètement les recommandations issues de la commission IA ?

FSM : J'étais membre de la délégation qui a négocié cet accord européen de 2020. Je représentais les syndicats français. Je dois dire que cet accord s'est conclu de manière assez inattendue en plein contexte de Covid-19. La finalité principale de cet accord qui porte sur les transitions numériques était d'avoir un droit européen à la déconnexion. C'était une attente claire des représentants des travailleurs. L'intelligence artificielle est abordée mais n'est pas le thème central. Il existe néanmoins une méthode, très bien définie dans cet accord, qui mobilise les dialogues social et professionnel, de manière itérative, cyclique. Cette méthode est parfaitement adaptée à une technologie comme celle de l'intelligence artificielle, qui justement est une technologie qui s'adapte et évolue, et donc nécessite en permanence de mobiliser le dialogue social et professionnel, pour évaluer, ajuster, réguler. J'ai travaillé à un bilan de la déclinaison, dans les pays européens, de cet accord, et un avis a été rendu au niveau européen en janvier 2025.

Dans la plupart des pays, cet accord européen sur les transitions numériques a donné lieu à des précisions et parfois à des accords sur le télétravail et à des précisions sur la déconnexion. Par exemple en Espagne, mais aussi en Autriche, en Allemagne, au Luxembourg. Si l'on ne rencontre pas véritablement des accords qui déclinent cet accord européen, des éléments en ont été traduits par certains pays dans des lois ou règlements. La traduction porte beaucoup sur le télétravail, régime qui a proliféré dans le contexte du Covid et dans le monde entier. C'était encore très marginal dans certains pays d'Europe, en Italie par exemple, et cela a explosé dans ce contexte. D'où les besoins d'encadrement et de régulation.

Les accords ont surtout été des accords de télétravail qui impliquaient aussi une régulation de la connexion et de la déconnexion. La plupart des déclinaisons de cet accord n'ont pas porté sur l'IA. On a noté quelques ouvertures sur l'IA, par exemple en Allemagne où cet accord a suscité des discussions entre partenaires sociaux sur ce sujet, mais pas de négociation. Concernant la France, rien de spécifique n'a été accompli. Il y avait déjà un accord sur le télétravail et une loi en matière de droit à la déconnexion, mais la France n'a pas saisi cette occasion pour développer d'autres aspects, par exemple sur le dialogue social et professionnel ou encore sur la formation sur l'IA.

— Vous êtes vice-présidente du Comité économique et social européen (CESE), où vous représentez le groupe des travailleurs. Vous venez d'évoquer l'avis de janvier 2025 que vous avez porté auprès de la Commission européenne pour créer les conditions d'un déploiement équilibré de l'IA dans les milieux professionnels. Pouvez-vous nous indiquer les principales recommandations portées par cet avis, et peut-être aussi les principales difficultés rencontrées pour l'établir ?

FSM : Tout d'abord, cet avis est un avis à l'initiative du groupe des travailleurs du Comité économique et social européen qui représente la société civile organisée européenne. Le Comité accueille les représentants des entreprises européens, les syndicats, et les organisations non gouvernementales et associations diverses. Au sein du groupe des travailleurs, nous avons pris l'initiative d'un avis avec un titre qui était assez explicite : « Pour une intelligence artificielle au service des travailleurs ». Pour nous, cela signifie « au service de tous » car, lorsque cela se passe bien pour les travailleurs dans le travail, les entreprises la société tout entière sont aussi gagnantes.

Pourquoi avoir pris cette initiative ? Parce qu'il y avait, dans les tuyaux du Parlement européen, la remise d'un rapport sur l'intelligence artificielle et sur le management algorithmique en particulier, ceci dans la perspective d'une initiative législative au niveau européen. Le management algorithmique consiste à prendre des décisions ayant des effets directs sur les humains, de manière automatisée et autonome, grâce aux systèmes d'intelligence artificielle. Dans ce contexte, il nous est apparu essentiel que notre comité donne son avis sur une question qui allait être débattue au Parlement dans les prochains mois. Pour élaborer cet avis, nous avons utilisé la méthode de prospective, avec la participation d'experts, dont l'Anact a fait partie. Une trentaine d'experts ont été invités, experts européens surtout mais aussi au-delà de l'Union Européenne. La méthode était intéressante, elle nous a permis de débattre, de dessiner des scénarios concernant une technologie qui représente une évolution majeure mais qui comporte aussi des incertitudes. Dans cette projection, les scénarios ont montré tout d'abord que ce n'est pas trop tard pour influencer le déploiement de cette technologie. Dès lors il faut se donner les moyens d'agir. D'autre part, la question du dialogue social est revenue de manière centrale avec l'idée que cette voie du dialogue est la plus appropriée, pour déployer une intelligence artificielle dans l'intérêt de toutes et tous. Le dialogue social a été identifié comme l'instrument adapté pour élaborer des utilisations fiables de l'IA au travail. Par ailleurs, le scénario idéal rappelle qu'il existe certainement des ajustements réglementaires à faire car les législations actuelles n'abordent pas d'une manière spécifique les opportunités et les risques induits lorsque ces technologies d'IA sont déployées dans le monde du travail.

S'agissant des recommandations principales formulées dans cet avis, évoquons d'abord la définition : de quoi parle-t-on, quand on parle d'intelligence artificielle ? Nous avons repris la définition de l'OCDE et celle du règlement européen sur l'IA : l'intelligence artificielle est un système automatisé qui prend de manière autonome des décisions et qui s'adapte après son déploiement. Il présente des caractéristiques essentielles : la possibilité de prendre une décision automatisée avec un degré d'autonomie plus ou moins important, et la capacité d'interagir et d'influencer les environnements du travail. Nous avons également analysé les règlements et directives existants. Les règlements, comme le RGPD, sont assez inadaptés au monde du travail, notamment parce qu'il n'est question que des données individuelles et pas du tout des données collectives qui peuvent être exploitées dans le monde du travail.

Quant à l'*AI ACT,* cette loi représente certes une avancée significative dans la réglementation des systèmes d'intelligence artificielle à haut risque, mais elle demeure insuffisante pour relever pleinement les défis posés par la gestion algorithmique sur le lieu de travail, entre autres. Bien qu'elle classe les outils d'IA liés au travail comme présentant un risque élevé, elle se concentre principalement sur la mise sur le marché, la sécurité des produits et les obligations de conformité des fournisseurs et des utilisateurs, et non sur la relation employeur-employé.

D'autres directives, comme celle sur le travail via une plateforme, ne concernent que les travailleurs de plateformes comme Uber et Deliveroo. Or, le management algorithmique, technologie sur laquelle repose le modèle des plateformes, est de plus en plus étendu à toutes les réalités du travail. D'après une étude de la Commission européenne de décembre 2023, plus d'un quart des entreprises en Europe recourent à des systèmes de management algorithmique. L'utilisation la plus fréquente est finalisée à l'évaluation des performances des travailleurs. Ce type de management – sans intervention humaine – des hommes et des femmes s'étend aux entreprises traditionnelles, il faut donc que certaines dispositions de la directive sur les plateformes, qui étaient réservées aux travailleurs des plateformes, puissent s'étendre à tous les travailleurs.

Nous demandons notamment que le chapitre 3 de cette directive s'étende à tous les travailleurs : ce chapitre porte sur l'obligation de transparence algorithmique des décisions automatisées et sur le droit des travailleurs à l'information et à l'expertise. Ce point est très important : l'exigence d'intelligibilité des informations données aux représentants des travailleurs sur les systèmes à base d'IA permettrait un dialogue social plus transparent et qualitatif, notamment sur la prévention des risques. Dans la directive il y a aussi une obligation sur les résultats de l'analyse d'impact relative à la protection des données. Les articles 10 et 11 de la directive sur les plateformes abordent le contrôle des systèmes de surveillance ou des prises de décision en introduisant un droit à une interface humaine. Le recours humain en cas de décision automatisée, est d'ailleurs précisé, comme évoqué, dans une des recommandations de la commission IA. Il s'agit de s'inspirer de cette directive, et c'est l'une de nos recommandations centrales. Notre point de vue, énoncé dans cet avis, est que pour améliorer la législation existante, voire l'élargir, il y a nécessité de produire un texte européen ad hoc sur l'intelligence artificielle au travail.

— Pouvez-vous préciser et nous expliquer pourquoi, de votre point de vue, ce règlement européen sur l'intelligence artificielle (AI Act) vous paraît inadapté ou en tout cas limité pour prendre en charge les sujets du travail ?

FSM : Comme évoqué, la finalité de ce règlement est le déploiement des produits ayant des systèmes d'IA, avec des règles cohérentes, en Europe, pour qu'il y ait une harmonisation sur les marchés intérieurs. Il s'adresse avant tout aux fournisseurs et aux concepteurs de ces systèmes, et aux usagers, c'est-à-dire aux utilisateurs c'est-à-dire les entreprises ou les administrations qui adoptent ces systèmes. Ce règlement n'est pas fait pour réguler toutes les questions qui peuvent surgir au niveau du travail. Même si l'identification des niveaux de risque des systèmes l'aborde en partie. Par exemple, dans les systèmes définis à « haut risque », certains des éléments évalués concernent la gestion humaine, l'introduction du management algorithmique, par exemple, dans les relations de travail. Mais le règlement indique lui-même que, même si un système peut être conforme aux règles qui sont prévues dans ce règlement, il peut y avoir, selon les usages, des effets imprévus. Par exemple, imaginons des effets discriminatoires sur des critères de santé : les étapes pour réguler ces risques sont très longues et, en attendant, de multiples situations de discrimination peuvent exister. Donc, pour la Confédération européenne des syndicats, ce règlement ne répond pas aux enjeux de protection des travailleurs ; il faut que ces questions soient traitées dans un instrument *ad hoc*.

Bien sûr, tous les partenaires n'ont pas ce point de vue. Les employeurs considèrent au contraire qu'il y a trop de législation. Dans la mouvance actuelle qui est celle de la dérégulation, d'une simplification réglementaire, beaucoup jugent que tous les maux de l'Europe viennent de l'excès des normes. Les représentants des employeurs se sont engouffrés dans cette posture. Or, je vois surtout une idéologie dominante anti-régulation qui nous vient des États-Unis, parmi d'autres. Les Américains et leurs compagnies de la Big Tech ont tout intérêt, quand ils agissent en Europe, à l'absence de règles. Je pense au contraire que l'Europe devrait trouver sa voie, une troisième voie, qui n'est pas celle de la dérégulation ni celle de la dictature chinoise et du contrôle, mais une voie européenne qui puisse déployer des systèmes d'IA de manière durable et fiable, ce qui passe par des normes qui respectent un certain nombre de principes, par exemple celui du contrôle humain comme dernière voix pour prendre des décisions.

Un label européen pourrait garantir que le système déployé de cette façon ne portera pas atteinte, ne nuira pas aux hommes et aux femmes qui travaillent, ou aux citoyens. À mes yeux, l'Europe devrait poursuivre sur cette voie qui n'entravera pas l'innovation. Car ce ne sont pas les normes qui nous empêchent d'innover. Le problème du retard de l'Europe est plutôt lié à cette incapacité à investir de manière robuste, à s'unir entre pays autour d'une stratégie commune, à retenir et attirer des talents. Les règles peuvent à l'inverse favoriser et soutenir l'innovation en clarifiant le cadre de l'utilisation d'une

technologie comme l'IA et en instaurant le même niveau d'exigence pour les entreprises européennes et extra-européennes qui opèrent en Europe. Rappelons que le marché européen, avec ses 500 millions de citoyens, est un des plus convoité au monde.

— Vous avez évoqué à plusieurs reprises la directive européenne sur le travail via une plateforme. Cette directive génère du droit pour ces travailleurs, et peut-être aussi pour d'autres types de travailleurs, comme vous l'avez souligné et comme vous le défendez. Savez-vous comment cette directive européenne est transposée en France ? Comment cette transposition française pourrait-elle étendre les dispositions à d'autres catégories de travailleurs, qui n'étaient pas initialement visées par la directive en question mais qui sont quand même concernées par des formes de management algorithmique ?

FSM : Cette directive a été adoptée pendant la présidence portugaise du Conseil européen, qui a été une présidence très active dans le domaine du social. Non sans difficulté, elle a été adoptée en novembre 2024. L'objectif était une transposition début décembre 2026. Qu'a fait la Commission après l'adoption de cette directive ? Un groupe d'étude a été mis en place, réunissant des experts de chaque pays membre et des partenaires sociaux. Ces travaux ont débuté en 2024 et se sont poursuivis en 2025. Par exemple, des réflexions ont été menées sur les possibilités d'étendre le chapitre 3, que j'évoquais précédemment, à toutes les organisations de travail, au-delà des plateformes. La France a avancé avec le travail de la commission IA et la formalisation de recommandations ; il reste à opérer des avancées concrètes. Le ministère du Travail anime actuellement un groupe sur la transposition de cette directive, mais la France est très figée sur le sujet, et pose notamment comme préalable la question du statut des travailleurs, qui est assez flou dans la directive. La position française est de dire que, parmi les travailleurs des plateformes, on retrouve de vrais indépendants et de faux indépendants, et elle renvoie au juge la responsabilité de trancher. Dans d'autres pays, le sujet a été pris autrement. Par exemple, l'Espagne affirme que, par défaut, les travailleurs sont des salariés.

Chaque État devra donc décider si un travailleur est dépendant ou réellement indépendant. La position de la France est de maintenir cette distinction. Mais en réalité, nous devons nous concentrer sur le fait que, désormais, la méthode de management basée sur un management sans l'intervention humaine se répand. Il me parait essentiel d'étendre à toutes les réalités du travail les exigences de transparence sur la technologie employée, ainsi que sur la collecte et l'utilisation des données personnelles et collectives, et sur l'évaluation d'impact une fois la technologie déployée : est-elle discriminante, oublie-t-elle l'intervention humaine, porte-t-elle atteinte à la vie privée, à la dignité humaine, etc. ? Et dans ce contexte, nous devons redonner toute sa place au dialogue social et dialoguer sur la base d'informations approfondies, plus explicites, plus transparentes et intelligibles s'agissant des systèmes d'IA. Il reste du chemin à parcourir en France pour que cette directive sur le travail des plateformes soit déclinée et pour que certaines obligations soient élargies à tous les salariés et travailleurs.

ENCADRER L'IA AU TRAVAIL PAR UNE DIRECTIVE EUROPÉENNE

Entretien avec Nayla Glaise

Nayla Glaise est présidente d'Eurocadres, organisation syndicale européenne représentant plus de 5 millions de cadres. Elle est également membre de l'Union générale des ingés, cadres et techs (Ugict-CGT).
Dans cet entretien, Nayla Glaise analyse les enjeux soulevés par le déploiement de l'intelligence artificielle pour les cadres et, plus largement, pour l'ensemble des travailleurs européens. À partir de l'expérience d'Eurocadres dans les négociations européennes, elle met en évidence les limites des dispositifs actuels de régulation et plaide pour l'adoption d'une directive européenne spécifique sur le management algorithmique. Celle-ci permettrait de garantir des droits renforcés en matière de transparence, d'explicabilité, de contrôle humain des décisions automatisées, d'expertise et de négociation collective lors de l'introduction de systèmes d'IA dans les organisations. L'entretien souligne également l'importance d'interroger les besoins auxquels répondent réellement les systèmes d'IA, d'évaluer leurs effets sur le travail et de préserver les marges de manœuvre des travailleurs comme condition d'une utilisation responsable de ces technologies.

Entretien réalisé par Marion Gilles et Vincent Mandinaud (Anact)

— Vous représentez les cadres au sein de la Confédération européenne des syndicats (CES), environ 6 millions de cadres, dans les secteurs privés comme public. Après avoir présenté succinctement votre organisation, pouvez-vous nous indiquer quels sont les principaux enjeux soulevés par votre organisation à propos des conditions de travail des managers dans un contexte de déploiement rapide de l'IA dans les entreprises et les administrations ?

Nayla Glaise : Eurocadres est une organisation qui regroupe les syndicats qui syndiquent les cadres dans 24 pays européens. Nos membres ne sont donc pas les travailleurs directement, mais les syndicats qui syndiquent les travailleurs cadres et professions intermédiaires. En France, six organisations syndicales sont membres d'Eurocadres – l'Ugict pour la CGT, par exemple. Tout en faisant partie de la CES, Eurocadres est reconnue par la Commission européenne comme un partenaire social européen. Cela signifie que nous avons notre place à la table des négociations européenne avec la CES, au côté des trois organisations d'employeurs européens (Business Europe pour les grandes entreprises, SME United pour les PME et SGI Europe pour les employeurs de la fonction publique). Ainsi, en cas de consultation européenne, nous sommes consultés à l'instar des autres partenaires sociaux. De même, Eurocadres participe aux négociations. C'est important, car cela signifie que nous disposons d'une certaine autonomie pour les sujets qui concernent principalement les cadres.

Nous sommes basés à Bruxelles. Parmi nos activités, nous pratiquons le lobbying, c'est-à-dire que nous travaillons auprès des institutions européennes afin de soutenir des amendements, convaincre les parlementaires et la Commission d'inscrire certains sujets et certaines orientations à l'ordre du jour. Ce travail de lobbying vise à influencer la réglementation européenne. Nous sommes actifs aussi en dehors des processus de réglementation inscrits à l'agenda de la Commission. Par exemple, Eurocadres était à l'origine de la directive sur les lanceurs d'alerte. Et à l'époque, quand nous avons commencé, il n'y avait pas de texte de la Commission ni de volonté de légiférer. Nous avons débuté par une campagne sur les lanceurs d'alerte en créant une plateforme avec des ONG, et ce travail a abouti à une directive.

La grande partie de notre travail, cependant, se mène auprès de nos membres pour faciliter l'appropriation des textes européens et accompagner la transposition des directives dans les États membres. Par exemple, actuellement, nous travaillons sur la directive sur la transparence des salaires qui doit être transposée avant juin 2026 dans tous les États membres. En atelier de travail, nous explicitons chaque article de cette directive et formons les syndicats afin qu'ils la transposent au mieux et que les élus s'approprient cette loi. Cela passe également par la publication de rapports et de guides pratiques. Notre travail est aussi d'informer tous nos membres sur l'actualité de la législation européenne, les nouveaux textes, ceux en cours de négociation, les votes du Parlement, les enjeux associés.

Il est enfin des sujets sur lesquels nous sommes mobilisés en continu. Les risques psychosociaux sont une priorité forte. Notre campagne Endstress.eu appelle à une directive, même s'il est très compliqué, avec cette Commission et ce Parlement, d'obtenir quoi que ce soit en termes de législation. Mais on continue cette campagne parce que c'est une priorité pour nos membres, c'est une priorité pour tous les travailleurs, et particulièrement les cadres. Bien sûr, une autre thématique qui nous occupe massivement, c'est l'intelligence artificielle. Beaucoup de textes ont été adoptés sur les questions liées à l'IA durant le dernier mandat de la Commission.

— Concernant l'IA, quelles sont les principales problématiques que vous identifiez pour les cadres ?

NG : Plusieurs textes législatifs sont liés à l'IA, mais il n'y a pas grand-chose sur la façon de prévenir les risques liés à l'IA au travail. Or, les travailleurs cadres sont directement confrontés à ces nouvelles technologies, dans leur métier mais aussi en tant qu'intervenants dans la conception des organisations du travail. Tout comme les développeurs, membres également d'Eurocadres, qui conçoivent des systèmes d'IA. L'idée selon laquelle il s'agit d'un simple outil qui n'a pas d'impact sur l'organisation du travail est très répandue.

Il existe un fort enjeu à informer les cadres, pas seulement sur l'utilisation de ces systèmes, mais aussi sur l'impact que ces systèmes peuvent avoir sur l'organisation du travail et sur leur travail. À informer aussi sur les risques de discrimination, de surveillance, etc. La législation actuelle ne suffit pas, même si la tendance est à la « simplification » – ce qui, en fait, veut dire déréglementer –, notre enjeu est d'obtenir une directive européenne sur le management algorithmique. La demande est forte, pas uniquement des syndicats d'ailleurs, mais aussi de plusieurs partis politiques qui mesurent le besoin d'encadrer.

En décembre 2025, un rapport du Parlement a également appelé pour un texte législatif encadrant le management algorithmique.

— En 2020, Eurocadres a participé à la négociation d'un accord-cadre européen sur la transformation numérique qui intègre un chapitre sur l'IA. En 2022, un accord spécifique a été trouvé pour les administrations centrales et fédérales, avec également un chapitre sur l'IA, beaucoup plus étoffé, avec une douzaine d'articles assez précis. Quel bilan tirez-vous de ces deux accords, de leur appropriation en Europe, en France ? Concernant l'accord de 2020, il était prévu une évaluation de sa mise en œuvre dans les États membres. Cette évaluation a-t-elle eu lieu ? Quelles ont été les suites données à cette évaluation ?

NG : Je ne vais parler que de l'accord-cadre sur le numérique de 2020, car celui de 2022 concerne uniquement la fonction publique, et c'est la Fédération européenne des services publics qui a négocié ce texte pour les syndicats (EPSU) et SGI Europe pour les employeurs. Eurocadres faisait partie des négociateurs de l'accord-cadre sur le numérique, qui est un accord vraiment important. Tout d'abord, il touche tous les secteurs, public et privé. D'autre part, c'est un accord normatif qui pose l'objectif de négociation dans l'entreprise pour toute introduction de nouveaux outils numériques. La France est un très mauvais élève pour ce qui est de la mise en œuvre de cet accord. Des évaluations régulières associent les États membres qui doivent réaliser des rapports pour donner à voir leurs avancées. La France est l'un des seuls pays qui ne produisent pas ce rapport. Et l'on sait que la déclinaison de cet accord est quasi nulle.

— Quels sont les obstacles à l'implémentation en France ?

NG : Dans les discussions sur l'implémentation française en 2022, le MEDEF a clairement bloqué le processus, considérant que, sur toutes les parties de l'accord-cadre, la France avait déjà les outils et n'avait pas besoin de nouveaux accords. Par exemple, sur la partie formation, le MEDEF estimait que le CPF suffisait. L'accord-cadre orientait vers la formation des salariés aux nouveaux outils numériques avant leur mise en place – le CPF ne peut clairement pas être vu comme une réponse satisfaisante à cela. De même pour le droit à la déconnexion. Le MEDEF a jugé qu'il existe un droit à la déconnexion depuis 2017 en France et qu'il n'ait pas besoin d'introduire de nouvelles mesures. Alors même que l'accord-cadre insistait sur la nécessité de discuter et de produire des mesures relatives aux modalités de connexion et de déconnexion. Le seul droit à la déconnexion ne suffit pas, ce sont les modalités à mettre en place qui sont essentielles.

Il n'y a pas eu de rapport français parce qu'une partie du patronat français n'avait aucune intention de négocier un accord national sur la transformation numérique, et notamment sur tout ce qui est prévu dans l'accord-cadre. Quant à l'implémentation de cet accord de 2020 sur le numérique, cela ne s'est pas passé de la même manière dans d'autres pays. En Belgique, plusieurs articles de loi font référence à l'accord-cadre, par exemple pour ce qui touche à la communication des télétravailleurs avec leurs représentants, au respect de la vie privée, à la non-surveillance, etc. Au Danemark, un accord tripartite a été conclu, à la suite de l'accord-cadre, pour renforcer l'accès à la formation continue et améliorer sa qualité. Les partenaires sociaux se sont notamment accordés pour élargir le programme de formation sur l'utilisation des technologies numériques, avec un fonds de financement de 50 millions

d'euros pour faciliter cette transition sur le marché du travail et inciter les salariés à participer à des formations sur les outils numériques et leurs usages.

En Espagne, un accord sur l'intelligence artificielle a été signé avec les partenaires sociaux et le gouvernement pour garantir les droits du travail des personnes qui travaillent dans le secteur des plateformes numériques. Et cet accord a été concrétisé par un règlement (la loi pour les livreurs). Il donne par ailleurs le droit au comité d'entreprise d'être informé par l'entreprise des paramètres, règles et instructions sur lesquels reposent les algorithmes ou les systèmes d'intelligence artificielle et qui ont un impact sur la prise de décision susceptible d'affecter les conditions de travail, l'accès à l'emploi et son maintien, y compris le profilage. Toutes ces mesures ont été mises en place après la négociation de cet accord.

Cela montre la limite des accords-cadres lorsque certains partenaires ne veulent pas négocier, et la nécessité d'aboutir à une directive sur des sujets. Cela ne signifie toutefois pas qu'on ne veut pas négocier avec les employeurs. Le programme de travail de 2026-2029 est en cours de négociation avec le patronat européen. À ce jour le sujet de négociation pour les deux prochaines années serait l'actualisation de l'accord autonome européen sur « active ageing and intergenerational approach » signé en 2017. Ce programme de travail devrait être signé à la rentrée 2026. Mais on voit bien que, dans certains pays, s'il n'y a pas une contrainte plus forte, la négociation seule ne produit rien. S'agissant du management algorithmique, notre objectif est d'aboutir à une directive, car les risques sont importants pour les travailleurs.

— Est-ce que l'IA équivaut à un management algorithmique ? Ou est-ce parce qu'il s'agit d'une traduction opérationnelle et saisissable dans l'entreprise ? Une directive sur le travail via une plateforme a été adoptée. Si elle porte sur le management algorithmique, elle est restreinte aux travailleurs des plateformes. Elle aboutit à un certain nombre d'éléments intéressants pour ces derniers, éléments qui pourraient être étendus à d'autres travailleurs. Quand vous dites que vous cherchez à obtenir une directive sur le management algorithmique au travail, l'objectif est-il d'étendre la directive sur les plateformes au champ salarié et public ?

NG : Eurocadres a beaucoup travaillé sur la directive concernant les plateformes, et surtout sur la partie « management algorithmique ». Avec pour résultat une avancée notable, notamment sur l'interdiction de traiter certaines données, comme l'état psychologique des travailleurs, l'appartenance religieuse, la sexualité, mais aussi les données qui concernent les conversations privées et les informations extérieures à l'activité de travailleurs. C'est d'autant plus notable que cette interdiction intervient dès le processus d'embauche, ce qui était essentiel pour nous. Outre la question de la transparence, nous posons aussi celle de l'explicabilité. Cela permet d'avoir accès à des expertises : l'employeur lui-même n'étant pas toujours capable d'expliquer, il a besoin d'une expertise un peu plus neutre, externe. C'est un point très important de cette directive ; nous y voyons une brèche qui s'est ouverte, grâce à laquelle il est possible de poursuivre le travail pour aller plus largement vers une directive sur le management algorithmique qui concernerait tous les travailleurs.

Un processus est en cours au sein du Parlement européen. Nous défendons l'idée que reprendre le chapitre sur le management algorithmique développé dans la directive sur les plateformes opérerait une véritable simplification : cela permettrait de traiter tous les travailleurs de la même façon. Il n'y aurait plus de question sur la distinction entre travailleurs, salariés et indépendants quand il s'agit de management algorithmique. Par ailleurs, nous souhaitons qu'il y ait une obligation de négocier. C'est-à-dire, en amont de la mise en place d'un système d'IA, qu'il y ait évidemment un processus d'information, de consultation des comités d'entreprise ou des représentants du personnel avec recours possible à un expert, mais également qu'une négociation soit lancée à cette occasion. Dans une directive européenne, tout ne peut pas être légiféré ni anticipé, notamment en termes de risques. Cette évaluation des risques et des mesures adaptées pour les prévenir doit se faire au niveau de chaque

entreprise, dans une discussion avec les syndicats, car cela dépend de l'activité de l'entreprise, du secteur d'activité, de la population au travail, etc.

Parallèlement, la commission européenne a lancé en Janvier 2026 le processus de consultation des partenaires sociaux sur le règlement pour des emplois de qualité. Les sujets évoqués incluent les RPS, le droit à la déconnexion, le télétravail, le management algorithmique, la sous-traitance, etc… À ce jour nous attendons toujours la 2ème phase de consultation des partenaires sociaux sur ce règlement. Selon le programme de travail de la commission européenne de 2026, ce règlement devrait apparaitre fin 2026.

Par ailleurs, un autre aspect essentiel pour nous est la formation sur les systèmes d'IA, sur leurs impacts sur l'organisation du travail et sur les métiers, sur leurs risques et la façon d'anticiper ceux-ci. La formation pour tous les travailleurs, mais aussi pour les cadres et les dirigeants. Ce qui figure dans la directive sur les plateformes en matière de contrôle humain est tout aussi essentiel : l'interdiction que la gestion algorithmique soit complètement autonome, sans contrôle humain, notamment pour les systèmes d'IA qui ont un impact sur les conditions de travail.

La responsabilité claire des employeurs en matière de santé et sécurité au travail est un autre point majeur qu'il faut rappeler. De notre point de vue, la discussion entre employeurs et syndicats sur la mise en place des systèmes d'IA est un plus pour tout le monde, pour les travailleurs comme pour les employeurs. Par exemple, on observe beaucoup de pratiques appelées « *shadow AI* », quand les travailleurs recourent à des systèmes d'IA sans l'autorisation de l'employeur. Ils mettent en place des systèmes d'IA sur leur ordinateur et les utilisent parce qu'ils perçoivent cela comme une aide dans l'accomplissement de leurs tâches. Mais la question est : quand il y a un problème, qui est responsable ? Prenez le cas d'une banque où des fuites de données confidentielles sont causées par des systèmes d'IA utilisés par des travailleurs sans le consentement de l'employeur. Qui est responsable ? Il y a un vrai intérêt à ouvrir cette négociation, à avoir un dialogue transparent, notamment pour clarifier la responsabilité des employeurs et les besoins des travailleurs.

— Vous venez de dire que les gens utilisent l'IA comme une aide pour accomplir leurs tâches. Cela peut-il aussi générer des difficultés en termes de conditions de travail, un abandon de son jugement professionnel par exemple ? Quelle lecture critique faites-vous de l'introduction et de l'usage de l'IA ? Est-ce simplement pour vous un enjeu d'encadrement, à travers la négociation ou le droit ? Ou y voyez-vous des enjeux en termes de contenu du travail, de développement professionnel, et pas simplement en termes de formation et de développement des compétences en digital ?

NG : C'est très délicat ! Il n'est pas évident de répondre de façon générale, car chaque situation est particulière. C'est la raison pour laquelle la négociation est essentielle. La première question à se poser dans la négociation est celle-ci : quels sont nos besoins ? Beaucoup d'employeurs mettent en place des systèmes d'IA juste parce que c'est une tendance, une mode. Alors que des fonctionnalités sont probablement déjà accessibles avec un logiciel existant, et qu'il n'y a pas besoin d'investir et de créer des risques non maitrisés.

Cette question du besoin n'est pas suffisamment posée dans les entreprises, par les employeurs comme par les syndicats. La première question à poser sur la table de la négociation est celle du pourquoi : pourquoi mettre en place ce système d'IA ? Qu'est-ce qu'il va accomplir que nos outils actuels sont incapables de réaliser ? Quels impacts l'introduction de ce système peut-elle avoir sur le travail ? Quant au lien avec la question du sens, il me semble que, lorsqu'un salarié décide d'utiliser un système d'IA, c'est parce que c'est une aide dans son travail ; non parce qu'il ne voit plus le sens de son travail, mais parce qu'il veut aller plus loin dans la réalisation de son travail. En revanche, il

peut perdre du sens si l'outil lui est imposé et qu'il n'en voit pas l'intérêt ou qu'il perd des marges de manœuvre pour contrôler, maitriser ce qu'il fait et comment. D'où l'importance de commencer à établir les besoins et voir en quoi un nouvel outil y répondrait. La négociation est l'occasion, pour les représentants des travailleurs, de discuter avec les travailleurs, de les entendre parler de leur travail et de leurs besoins. Les situations à risque sont celles où les travailleurs perdent le contrôle et donc le sens de leur travail.

— Nous avons évoqué la question de l'IA dans le travail des salariés français et européens. Mais on sait aussi que les systèmes à base d'IA relèvent d'un agencement compliqué, avec un ensemble de travailleurs dont certains sont actifs à l'autre bout du monde, dans des conditions d'emploi très dégradées. Comment raisonner en termes d'IA éthique au travail sur l'ensemble de la chaine de valeur, en intégrant le salarié européen ou français qui utilise des IA mais aussi les travailleurs – du Sud souvent – grâce à qui ces systèmes se développent mais qui n'ont pas les mêmes dispositions à la négociation et à construire des conditions de travail qui leur soient favorables ?

NG : En effet, en matière de droits humains, les travailleurs qui alimentent les systèmes en données le font dans des conditions extrêmement dégradées. Par ailleurs, l'impact sur l'environnement est particulièrement négatif. La question de l'impact environnemental, qu'il faut évaluer, doit faire partie des discussions sur les lieux de travail. Et là encore, cela nécessite de se poser la question du besoin : est-ce qu'on a vraiment besoin de ce système d'IA ? On sait aussi l'impact très négatif des data centers sur l'environnement. Dans le cadre d'une directive sur le management algorithmique, on doit cerner ce périmètre et le limiter. Le périmètre est celui de l'utilisation d'une IA éthique au travail et pas de sa conception. Nous sommes contraints d'avancer par petits pas, car nous sommes loin, au niveau européen, de poser la question des « travailleurs du clic », et très loin aussi de considérer les impacts environnementaux du développement de « giga centers ». La directive sur le devoir de vigilance est néanmoins une porte pour atteindre ces travailleurs sur la chaine de sous-traitance. Malheureusement, l'omnibus proposé par la Commission européenne va à l'encontre du progrès et limite le périmètre de cette directive. C'est une discussion politique qu'il faut toutefois mettre en place ; les annotateurs et modérateurs de contenu dont les droits fondamentaux ne sont pas respectés, ne sont pas uniquement en Afrique, mais aussi en Europe.

— Concernant le règlement européen sur l'IA (AI Act), quel est le point de vue d'Eurocadres ? Pensez-vous qu'il couvre les enjeux du déploiement de l'IA pour les travailleurs en général et les cadres en particulier ? Quels sont ses points d'intérêt, ses points faibles et ses limites ?

NG : Ce règlement a le mérite d'exister pour sécuriser un marché de l'IA en Europe. Son gros point faible, c'est qu'il n'est pas bâti pour traiter la protection des travailleurs. Bien que nous nous soyons beaucoup mobilisés à Bruxelles, c'est finalement un texte basé sur le produit et pas sur le travail. Il est axé sur les produits des systèmes d'IA qui vont circuler dans le marché européen, et c'est malgré tout un bon début pour la réglementation. Un point intéressant : il classe les systèmes d'IA par niveau de risque et exige certaines mesures selon les risques ; si le risque est évalué comme « inacceptable », il est interdit sur le marché européen. Par exemple pour des systèmes d'IA utilisés pour le recrutement, le licenciement ou l'évolution de carrière, il existe des exigences en matière de transparence, d'évaluation des risques, de traçabilité, de qualité des données, de contrôle humain. Mais c'est lié au produit et cela encadre sa conception qui doit répondre à ces exigences-là pour qu'il puisse être estampillé CE. Mais une fois le produit vendu, il n'y a plus d'exigence, notamment en matière d'évaluation des risques. Or, on sait très bien qu'un système d'IA qui est mis en place au travail évolue selon les données qui y sont injectées ; il change de comportement. Des usages peuvent aussi détourner le système. S'il n'y a pas une évaluation sur l'utilisation de ces systèmes d'IA au travail, on peut retrouver des risques de discrimination, de surveillance, etc., parce qu'on a utilisé un système d'IA pour autre chose.

Je vous donne un exemple : Midjourney, un système d'IA qui crée des images à partir de descriptions textuelles. J'utilise beaucoup cet outil pour nos communications. C'est très pratique pour cette activité. Mais on a voulu le tester. Si vous tapez en anglais « Doctor » et « Nurses », les images proposées pour le premier sont des médecins, tous des hommes, d'un certain âge et de couleur de peau blanche, tandis que, pour « Nurses », ce sont des femmes jolies, en général, jeunes, etc. Ce n'est pas un outil classé à haut risque, donc il n'est concerné par aucune exigence de règlement sur l'IA. Mais si je l'utilise au travail pour mener une campagne de recrutement, par exemple, je suis déjà dans les biais de cet outil-là et cela peut avoir au final un impact discriminatoire. Voilà la limite du règlement : les exigences s'imposent aux vendeurs mais pas aux utilisateurs, ni aux acheteurs, par exemple un employeur. C'est pour cela qu'une directive sur l'utilisation de l'IA au travail est nécessaire.

FAIRE DE L'IA, UN LEVIER D'INNOVATION POUR LES ENTREPRISES

Entretien avec Yahya Fallah

Yahya Fallah est entrepreneur et membre du comité exécutif des Entrepreneurs (anciennement CPME), président de la commission IA des Entrepreneurs.

À partir de son point de vue d'entrepreneur dans le développement de solutions numériques à base d'intelligence artificielle pour les ressources humaines, Yahya Fallah décrit, dans cet entretien, les nouvelles possibilités offertes par ces technologies pour analyser les compétences, personnaliser les parcours de formation, accompagner les trajectoires professionnelles ou encore détecter des signaux faibles dans les organisations. Il insiste sur l'élargissement des capacités d'action que permet l'IA grâce à l'exploitation de données actualisées et contextualisées. L'entretien aborde également les enjeux portés par Les Entrepeneurs et notamment les conditions de diffusion de l'innovation dans les PME et les ETI ou encore les questions de formation et d'appropriation des technologies par ces entreprises.

Entretien[1] réalisé par Marion Gilles et Vincent Mandinaud (Anact)

1 Entretien réalisé le 24 octobre 2025

— En tant qu'entrepreneur, vous avez fondé plusieurs sociétés, et notamment une société de développement d'outils digitaux, dont à base d'IA, au service des RH. Pouvez-vous nous décrire ces outils ? Comment fonctionnent-ils et à quoi servent-ils ?

Yahya Fallah : Effectivement, j'ai cofondé plusieurs sociétés. Cela fait plusieurs années que je suis entrepreneur et je suis vraiment passionné par tout ce qui touche à l'entrepreneuriat. Je trouve que c'est vraiment un bon moyen pour avoir un impact significatif sur la société. J'ai une deuxième passion, l'éducation, et tout ce qui touche à la transmission de la connaissance. Pour moi, la connaissance, la formation, c'est ce qui permet de libérer les individus, pour qu'ils puissent avoir un esprit critique, une compréhension des enjeux, et qu'ils puissent être libres de prendre leurs propres décisions. Lorsque nous avons commencé voici un peu plus de 11 ans, mon associé et moi, nous avions cette idée que la technologie doit être un levier pour proposer des solutions qui offrent à un nombre croissant de personnes d'apprendre, et de la meilleure façon qui soit. Sur les dernières décennies, le numérique, la technologie, sans encore parler d'intelligence artificielle, ont permis de démocratiser la connaissance, de la faire mieux circuler. Alors, bien sûr, la médaille a son revers sur certains aspects, avec les réseaux sociaux. Mais on voit des aspects positifs.

C'est avec cette vision-là que nous avons créé un organisme de formation, Seven, et une plateforme, OpenAleph, qui associe gestion des ressources humaines, développement des compétences et formation personnalisée. L'objectif principal : offrir le parcours de formation le plus personnalisé possible. Et pour cela, il faut disposer de données sur les individus. Il faut connaitre l'apprenant, sa position initiale et celle souhaitée, afin de déterminer le chemin le plus rapide ou le plus adapté. La première étape est de savoir ce que la personne sait, et cibler ce qu'elle ne sait pas, par rapport à ce qu'elle vise. Deuxième étape : se demander qui elle est, comment elle apprend, quelle est sa meilleure façon d'apprendre. Sans aller vers des choses trop difficiles à identifier, il peut simplement s'agir de la langue, par exemple. Beaucoup d'institutions, d'entreprises ne savent pas quelle est la langue que les personnes maitrisent le mieux. Or, leur apprendre quelque chose en français, en anglais, en allemand ou dans d'autres langues n'a pas le même impact ! C'était cela, notre objectif initial.

Nous avons créé cette plateforme SIRH avec plusieurs applications qui fournissent des données sur les collaborateurs, des données propres, une data de qualité. Et grâce à l'IA, nous pouvons réaliser avec ces données des analyses assez avancées, produire par exemple des mappings de compétences, faire des suggestions de parcours de formation ou de gestion de carrière. Voilà le but de la plateforme. Elle propose plusieurs applications, par exemple pour mener des entretiens de fin d'année, de mi-année, pour rédiger des rapports d'étonnement, pour réaliser des feedbacks 360, des sondages, pour fixer des objectifs et suivre les objectifs des collaborateurs avec leur manager, etc. Les managers, la direction ou même les individus eux-mêmes peuvent utiliser ces applications. Un salarié peut consulter son profil et voir ce qu'il a fait, où il en est de son parcours, obtenir des analyses de ses compétences, son propre mapping de compétences. Grâce à un *chat* à base d'IA, il peut poser des questions, demander des requêtes supplémentaires. Les applications comprennent aussi des modules d'apprentissage, avec des propositions ciblées en fonction de vos compétences, de celles que vous voulez développer, et de vos modes d'apprentissage privilégiés.

— À qui vendez-vous ces applications ?

YF : Essentiellement à des entreprises en B2B (*business to business*). Mais mon rêve serait de réussir à développer la plateforme pour qu'elle puisse être utilisée dans une version premium par n'importe qui dans le monde. Qu'un individu, où qu'il soit, puisse y recourir afin d'apprendre de manière ciblée, libre et personnalisée, et qu'il profite ainsi de contenus de qualité et sur mesure. Bien sûr, beaucoup de contenus sont aujourd'hui en libre accès. Par exemple, si vous allez sur le site du MIT, vous bénéficiez quasiment de tout le corpus du MIT, de tous les cours. Mais est-ce que tout le monde est capable de

lire et de suivre les cours du MIT ? Ce n'est pas certain. L'idée, c'est de rendre accessibles des contenus de qualité, de façon personnalisée et adaptée à ce que vous savez faire et à la manière dont vous apprenez. Voilà notre ambition. Peut-être qu'un jour, on utilisera la plateforme dans les écoles !

— Avez-vous identifié des transformations dans les pratiques de GRH liées à l'introduction et à l'usage de cette plateforme et des applications associées ?

YF : Oui, il y a clairement des transformations. On parle souvent de gain de productivité, d'aller un peu plus vite, etc. C'est ce qu'on observe, mais ce n'est qu'une partie du tableau général. Il existe surtout des choses que l'on peut faire actuellement et que l'on ne pouvait pas faire auparavant, peu importent les moyens. Par exemple, réaliser un mapping de compétences précis, à jour et de manière rapide, c'était très difficile car il faut beaucoup de données sur la personne, puis des capacités d'analyse de ces données. Il faut collecter, compiler des données et obtenir des données qui ne soient pas obsolètes. Si l'on met du temps à collecter les données, le risque, c'est l'obsolescence. La plateforme permet de collecter des données dans la durée et à jour sur l'individu (les formations suivies, les synthèses d'entretien, les évaluations, les projets développés et objectifs atteints, etc.). Et avec ces données, il est possible, grâce à l'IA, d'obtenir instantanément des mappings extrêmement poussés. C'est un exemple, mais voici quelque chose qu'il était presque impossible d'accomplir de manière simple auparavant, et que vous pouvez désormais faire de manière incroyablement facile.

Bien sûr, il faut avoir le bon modèle avec le bon contexte. Mais techniquement, c'est largement à notre portée. On voit aussi cela pour l'analyse, grâce à la possibilité de traiter un grand volume. Prenons un exemple : j'ai en tête une ETI de 2 000 collaborateurs qui avait terminé ses entretiens de fin d'année. Dans ce cas, il y a des modèles d'entretien, avec des questions différentes en fonction du métier. Ce n'est donc pas un questionnaire unique pour les 2 000 personnes. Cela demande d'analyser 2 000 fois 15 questions et 15 réponses. C'est humainement impossible. Aujourd'hui, vous pouvez réaliser des synthèses de ces entretiens avec des questions extrêmement précises.

Autre exemple : je travaille avec une entreprise qui voulait identifier des sujets, des signaux faibles, notamment sur les questions de harcèlement ou de discrimination. Pas des faits avérés, bien sûr, mais des choses plus fines, des signaux faibles qui pourraient constituer des alertes. L'IA a permis cela. Cela ne veut pas dire que l'outil produit des certitudes : ce sont des hypothèses, des pistes ; ensuite, il faut rencontrer les gens, conduire des entretiens, discuter avec les managers. Mais cela a permis ce ciblage précoce. Auparavant, c'était assez difficile à identifier. Vous deviez vous appuyer sur ce que vos strates managériales intermédiaires faisaient remonter. Là, vous pouvez l'analyser de manière beaucoup plus fine sur la sémantique même.

— Dans cet exemple avec les 2 000 fois 15 réponses que la machine peut traiter, on repère des signaux faibles, on rentre dans une granularité très fine. Cela permet de repérer des éléments qu'on ne voyait pas, parce que c'était traité par des humains et par la strate managériale qui faisait remonter ces éléments. Une conséquence possible n'est-elle pas de court-circuiter la strate managériale, notamment dans la remontée d'informations ?

YF : Surtout pas ! En tout cas, ce n'est pas mon approche. L'idée est plutôt de remettre l'humain au centre. Par exemple, quand on gagne du temps grâce aux technologies, la question est : que fait-on de ce temps gagné ? Pour moi, l'objectif est de passer encore plus de temps avec les gens, du temps de qualité, entre humains, parce que vous êtes libéré de certaines tâches. Dans mon exemple, l'objectif n'est surtout pas de court-circuiter les managers. D'ailleurs, sur la plateforme, les managers eux-mêmes peuvent utiliser toutes les applications. Ce n'est pas « en haut, on l'utilise, et en bas, on n'y a pas accès ». Non, eux-mêmes l'utilisent pour traiter la donnée. Surtout, si des signaux faibles sont

détectés, il ne faut pas s'arrêter au signal reçu. Il faut aller voir les gens, toutes les strates, les personnes concernées, échanger, discuter, comprendre, affiner, décortiquer. C'est un signal. Est-ce que le signal est juste ? Pas forcément. Il faut affiner.

— Cela signifie-t-il que toutes les personnes concernées, le travailleur lui-même, l'ensemble de la chaine, peuvent disposer des mêmes données ? Comment le rapport à la strate managériale est-il géré quant à l'utilisation des données ?

FY : Exactement, et c'est majeur. Sur notre plateforme, la donnée qui vous concerne vous est forcément accessible. Il n'est pas possible de paramétrer la plateforme pour vous interdire cet accès, on n'a pas développé de fonctionnalité pour cela. En revanche, la donnée qui concernerait d'autres personnes ne vous est pas forcément accessible, cela dépend de votre fonction. Il existe une espèce de hiérarchie, de structure : vous avez plus ou moins accès à telle ou telle donnée selon les niveaux d'accès dont vous disposez. Un manager accède aux données des collaborateurs qu'il manage, par exemple.

— J'ai accès aux données qui me concernent, mais pas aux données identifiantes des autres. Pourrais-je avoir accès aux données anonymisées, ce qui permettrait par exemple de me comparer aux autres ? Ces données anonymisées sont-elles rendues accessibles aux représentants du personnel, pour étayer les discussions en instance de dialogue social sur les problématiques RH ?

YF : Je n'ai pas d'exemple d'usage en tête, mais les entreprises avec lesquelles nous travaillons, les grands comptes, au moins les ETI, qui ont des représentants du personnel, sont des entreprises où règle globalement la transparence. Dans ces organisations, par exemple, beaucoup de personnes ont accès au niveau le plus élevé, c'est-à-dire qu'elles ont accès aux données correspondant à leurs responsabilités. Il n'y a pas cette idée qu'il faut cacher des données. C'est peut-être pour cela qu'on travaille ensemble… Mais si je travaillais avec d'autres types d'organisation, je les encouragerais à ouvrir ces données-là le plus largement possible, donc aussi aux représentants du personnel. En fait, je ne vois même pas pourquoi ils ne le feraient pas. Je ne vois pas d'intérêt à ne pas ouvrir la donnée. Il y a un intérêt à partager largement les analyses collectives anonymisées, dans le respect de la confidentialité des données individuelles, car grâce à elles, tout le monde peut contribuer aux réflexions, par exemple sur les parcours ou l'évolution des métiers.

— Vous dites également qu'au-delà de la productivité, on observe finalement un élargissement des possibles : on peut faire des choses qu'on ne pouvait pas faire. Vous ajoutez que la condition pour cela, c'est d'avoir des datas de qualité. Pour bénéficier d'analyses performantes, il faut que la donnée soit ajustée, propre, fraiche et actuelle. Cela suppose une collecte, ainsi que des formes de repérage et de suivi. Quelles sont les conditions de ce suivi ? Quelle est la frontière entre le suivi, l'actualisation et la surveillance ? Cela pose aussi la question de la protection des données personnelles.

YF : C'est assez technique comme sujet. La réglementation, par exemple le RGPD, avant même l'intelligence artificielle, aborde ces questions. Pour moi, il existe une règle d'or : je dois pouvoir maitriser mes données. Et cela dépasse la question du consentement : il s'agit aussi de n'utiliser la donnée qu'à partir du moment où elle est vraiment utile. On entend beaucoup de débats en ce moment sur le développement des modèles, leur utilité, l'idée que toutes les données sont utiles. Je pense que ces débats sont sains et qu'il faut les avoir, car il faut mener une réflexion en profondeur pour éviter toute dérive. Donc oui, il faut avoir de la donnée propre et que chaque utilisateur ait le contrôle de ses données. Sur notre application, par exemple, l'utilisateur a accès à la totalité des données qui le concernent, et il peut exercer ses droits sur ces données, dans le cadre prévu par la réglementation. C'est la base de l'architecture de notre plateforme.

— Pour revenir sur la question de la productivité, vous avez mentionné l'augmentation de la productivité par le levier technologique. Mais à quelles conditions ces systèmes technologiques rendent-ils l'activité plus productive ? Et quelles sont les conséquences de cette augmentation de la productivité ? Comment gère-t-on les effets de l'augmentation de la productivité, si augmentation il y a ?

FY : Pour apprécier l'intérêt de l'investissement, il faut comparer les gains obtenus aux coûts engagés. Est-ce qu'il y a des gains avec l'IA ? Oui. Est-ce qu'il faut mettre de l'IA sur toutes les tâches ? Non. D'ailleurs, certaines tâches n'ont pas besoin d'être automatisées, même sans IA. La question se pose à chaque fois. Certaines personnes tombent dans le piège qui consiste à tout automatiser, à mettre de l'IA partout, en pensant qu'ils vont forcément y gagner. Non. Il y a toujours un investissement initial et des couts de maintenance dès que les outils sont un peu techniques. Mais sur certaines actions, c'est assez simple et cela augmente la productivité. Prenez la rédaction de synthèses de textes : que faire de ce gain de productivité ? Nous avons nous-mêmes enregistré des gains assez importants avec ce type d'action. Cela a souvent transformé en partie le métier, ou permis à des collaborateurs et collaboratrices de faire de nouvelles choses, ou des choses qu'ils faisaient peu, qu'ils auraient aimé faire davantage mais pour lesquelles ils avaient moins de temps. Les premiers gains de productivité qu'on a observés concernent les tâches typiquement chronophages, comme le classement, le tri, les réponses rapides, la génération de devis, etc. Ces tâches sont obligatoires mais d'une faible valeur ajoutée. En automatisant, vous gagnez du temps, avec, à la clé, des gains de productivité sur ces tâches que l'on peut qualifier de premier niveau.

En général, le collaborateur ou la collaboratrice est plutôt content(e) de pouvoir se concentrer sur des tâches à plus forte valeur ajoutée, plus intéressantes intellectuellement. J'ai constaté cela dans ma propre organisation. Et ce gain de productivité nous a offert la chance d'approfondir certaines activités, et même de recruter, paradoxalement. Dans notre cas, cela s'est traduit par des créations d'emplois. On entend fréquemment le contraire. Je suis un *tech aficionado*, c'est sûr ! Néanmoins, je suis convaincu que cela va créer de l'emploi. Parfois, ce n'est pas immédiat, mais cela se verra plus tard. Cela créera de l'emploi sur la technologie en elle-même. Même si des études montrent, et c'est assez ironique, que certains métiers de développeur sont touchés par l'intelligence artificielle. Par exemple, aux États-Unis, des statistiques indiquent une baisse du nombre d'offres d'emploi pour les développeurs juniors. Mais il me semble que c'est temporaire, le métier va traverser une transformation. On commence à voir de nouveaux métiers connexes émerger avec l'intelligence artificielle et le développement informatique. À mes yeux, il s'agit essentiellement d'une petite transition.

Que faire avec ces gains de productivité ? Pour moi, on continue surtout à travailler, à travailler fort, individuellement et collectivement, afin de relever tous les défis d'aujourd'hui et de demain. Nous faisons face à un triple choc, démographique, environnemental et technologique. Pour le choc démographique, je m'interroge sur la manière dont on prend soin de nos ainés et des plus jeunes. À l'école maternelle, on compte 25 ou 30 enfants pour une maitresse ou un maitre. Je trouve ça frappant. Je ne sais pas à quel moment la société s'est dit que c'était une bonne idée… S'il existe des gains de productivité, ce sont tous ces métiers-là qu'il faut valoriser. Nous devons réfléchir à la manière de mieux prendre soin de nos plus jeunes, aux meilleures conditions pour leur développement, car c'est essentiel pour le restant de leur vie. De même, comment mieux prendre soin de nos ainés ? La vraie solidarité se situe ici. Pour moi, les gains de productivité doivent aussi permettre de renforcer notre capacité collective à financer la solidarité. La vraie solidarité, c'est lorsque celui qui peut travailler travaille toujours pour celui qui ne peut pas travailler.

— Qu'en est-il du partage de la valeur pour celui qui continue à travailler ? Si je comprends bien, on augmente la productivité et cela permet de soutenir ceux qui ne peuvent pas travailler. Et pour celui qui travaille, qui travaille pour lui, mais aussi pour les autres, est-il mieux rémunéré ? Vous défendez l'idée que le développement de l'IA augmentera le nombre d'emplois. Va-t-il aussi augmenter les salaires ?

YF : Pour moi, les salaires augmentent quand l'économie va bien, quand l'entreprise va bien, quand il y a de la stimulation, quand il y a de l'emploi, quand il y a création d'emplois. Cela créé des conditions favorables à la progression des salaires. En tout cas, une hausse durable des salaires doit pouvoir s'appuyer sur la création de richesse. Ils s'accroissent par la création de richesse et par un marché de l'emploi qui fonctionne bien. Pour peu, encore une fois, que nous soyons dans un marché fonctionnel, dans un marché compétitif. Si les conditions d'un marché compétitif ne sont pas réunies, les salaires ne progressent pas. Par exemple, à Paris, les conditions de marché sont réunies, le recrutement fait l'objet d'une compétition, les salaires sont en hausse. Mais si l'on se rend dans des zones géographiques où les conditions de marché ne sont pas réunies, il faut collectivement se poser la question de la répartition du travail et de la valeur. Mais ces approches ne sont pas de moi, beaucoup d'auteurs les ont très bien développées.

— Quels travaux menez-vous au sein de la CPME, et notamment de la commission IA que vous présidez ? Quel regard portez-vous, dans le cadre de la CPME, sur le rapport de vos adhérents – et peut-être plus largement des PME – à l'IA ? Où en sont-ils ? Quelles sont les pratiques qui vous remontent ou que vous observez ? Et qu'est-ce que vous diffusez finalement auprès de votre cible autour de ces enjeux ?

YF : Je pense qu'un travail assez important a été mené – pas seulement par la CPME, mais aussi par les médias, le monde économique de manière plus large, des institutions, des agences comme vous – autour de la sensibilisation. L'importance de l'IA, le fait qu'il ne faut pas louper le virage, mais aussi les avertissements, les points d'attention à garder en tête. On a passé ce cap-là, selon moi. Le deuxième cap, c'est le passage à l'action. Ce que nous constatons auprès des TPE, PME et ETI – et cela ne concerne pas que l'IA, c'est plus général –, ce sont les difficultés du passage à l'action en matière d'innovation. D'ailleurs, je travaille aussi avec de grands groupes qui affrontent cette problématique. Ils ont souvent davantage de moyens, mais ils rencontrent quand même des difficultés sur cet aspect. Au sein de la commission IA de la CPME, nous travaillons largement sur cette dimension.

Nous revenons aux fondamentaux. Selon nous, pour innover, que ce soit sur votre chaine de production, dans vos process, vos outils, etc., il faut avoir des marges de manœuvre économiques. C'est aussi simple que cela. Et la réalité, depuis plusieurs décennies, c'est que les entreprises, en France comme en Europe – mais en France surtout –, ont peu de marges de manœuvre économiques. Sur ce point, on revient à des concepts très basiques qu'on porte depuis longtemps concernant les prélèvements obligatoires, les charges, les taxes, qui réduisent considérablement ces marges. Dans ce contexte, un chef d'entreprise peine énormément à se projeter, parce qu'il voit cela comme un risque. Même si on lui apporte une solution un peu plus préparée, un peu prémâchée, il n'est pas certain – et à juste titre – que cela va prendre, que cela va marcher. Il se dit : « Pour moi, c'est un investissement, c'est un risque. Si ça ne marche pas, j'aurai perdu de l'argent et du temps. » Les chefs d'entreprise hésitent, et la dynamique d'innovation est faible. Ce n'est pas nouveau, nous observions cela bien avant l'arrivée de l'intelligence artificielle.

Pour nous, il y a dès lors un enjeu majeur à faire confiance aux entreprises sur leur propre capacité à innover – pour peu qu'elles en soient conscientes, évidemment. Et dans ce cadre, il faut leur laisser des marges de manœuvre économiques. Nous portons beaucoup cette idée-là dans nos échanges avec les institutions publiques. Ces acteurs mentionnent fréquemment les programmes mis en place avec des systèmes de prêts ou de subventions. Oui, c'est vrai et c'est mieux que rien. Mais la véritable

innovation a besoin de stabilité et de vision. Or, les soutiens sont souvent fléchés, rarement reconduits. Ce modèle pose deux problèmes.Tout d'abord, on doit savoir dans quoi il faut investir. Ce n'est pas une question facile quand on parle d'innovation. Qui sait quelle sera l'innovation-clé de demain ? Moi, je n'en sais rien. Parler d'IA, c'est facile, c'est arrivé. Mais qui investissait ? Quel programme d'investissement ou de formation en intelligence artificielle existait en 2020, 2019, 2018 ? Je parle à grande échelle pour les TPE, PME, je ne parle pas des programmes de recherche. Ces dispositifs restaient encore trop peu accessibles à l'ensemble des TPE et PME. Une fois que c'est arrivé, oui, on y va. Mais il faut aussi permettre aux entreprises d'anticiper davantage.

L'innovation de demain, personne ne sait d'où elle viendra : des entrepreneurs, des TPE, des ETI, de quelqu'un quelque part qui a une idée. Cela peut être un salarié. Mais pour qu'il puisse développer cette idée, il lui faut des marges de manœuvre économiques. C'est un vrai sujet de société : comment stimule-t-on l'innovation ? Pour moi, il faut laisser davantage de marges de manœuvre économiques et laisser des personnes innover là où elles sont. Leur faire confiance.

Deuxième problème : ces programmes de financement pour les TPE et ETI s'inscrivent dans plusieurs circuits et sont toujours assez lourds. Vous devez, en tant que chef d'entreprise, d'abord identifier le besoin, ce qui n'est pas facile. Ensuite, vous devez identifier les systèmes de financement. Puis, vous devez cocher toutes les cases pour enfin, éventuellement, décrocher le financement. En fonction des financements, vous aurez des systèmes de contrôle derrière. Dans certains cas, vous devez payer vous-même un cabinet d'audit… On finit par se dire que ce système, en tout cas en matière d'innovation, fonctionne un peu à l'envers. Il ne s'agit pas de dire qu'on enlève d'un coup les subventions, mais plutôt de diminuer la pression en termes de charges qui pèsent sur les entreprises et les travailleurs. Quitte, en parallèle, à remplacer certains de ces systèmes de soutien par des allègements durables pour avoir juste un système qui soit plus sain, et donner de la visibilité aux entreprises et aux travailleurs qui ne se demanderont pas en permanence si la subvention est reconduite ou pas et à quelle hauteur, tout cela dépendant de choix politiques dont on sait l'instabilité.

— Avez-vous d'autres sujets de travail au sein de la commission IA ?

YF : Notre feuille de route comporte trois axes. Le premier, c'est la sensibilisation. Cela se fait de multiples façons : échanges, rencontres, webinaires, partage de contenus, etc. Cela passe aussi par l'identification d'entreprises qui ont mis des choses en place, puis par le partage de ces pratiques avec d'autres entreprises. La CPME participe également à des travaux avec des institutions – France Num, la CNIL, la DGE, etc. – pour tenter de promouvoir ou mettre des actions en place pour les TPE, PME et ETI.

Deuxième axe : travailler avec les pouvoirs publics, avec les élus, avec des institutions pour discuter, sensibiliser les parties prenantes aux enjeux tels qu'on les perçoit, et partager, diffuser nos propositions d'actions. Par exemple, ce que je viens de vous dire sur les charges qui pèsent sur les entreprises et qui les empêchent d'innover et d'investir. Le troisième axe – auquel on croit beaucoup mais qui risque d'être un peu plus long à développer – consiste à réfléchir à l'ensemble de la chaine de valeur de l'intelligence artificielle. Ne pas voir uniquement l'IA par le prisme du dernier maillon de la chaine que sont les LLM développés par des entreprises comme Mistral, OpenAI, etc., mais vraiment regarder l'ensemble de la chaine de valeur et l'ensemble de l'industrie. Si l'on remonte un tout petit peu, on a les data centers, le cloud, les puces et les machines qui fonctionnent dans ces data centers. Si l'on remonte encore, il y a les fonderies, et ainsi de suite. Aujourd'hui, TSMC joue un rôle central dans la fabrication des puces les plus avancées, notamment celles conçues par Nvidia. On retrouve ici les machines lithographiques pour graver ces puces fabriquées par le Néerlandais ASML, acteur industriel européen majeur. C'est un bon exemple, d'ailleurs, parce qu'ASML était loin de régner sur les marchés boursiers européens voici quelques années. Grâce à des investissements de long terme, ils

ont développé des machines devenues essentielles à la fabrication des puces les plus avancées. On a de la chance en Europe d'avoir l'un des premiers maillons de cette chaine ! Sans oublier Mistral, qui arrive au dernier maillon.

L'idée, c'est de réfléchir à la manière de travailler collectivement, en France et dans les autres pays d'Europe, avec les grands groupes, les TPE, PME, les institutions publiques, pour reconstruire l'ensemble de cette chaine de valeur qui porterait la quasi-totalité des gains économiques, des emplois, etc. Nous pouvons adopter des modèles comme ceux de Mistral ou autres, mais cela ne suffit pas, il ne faut pas se focaliser sur ce bout de la chaine. Il faut porter une attention forte sur toutes les entreprises en amont de la chaine de valeur si l'on veut créer de l'emploi et stimuler l'économie. Pour ce troisième axe de ma feuille de route, nous essayons de travailler avec les institutions en jeu pour voir comment on peut innover sur ces sujets, rapprocher les entreprises, rapprocher les activités. Il existe énormément de pistes à explorer pour tenter de réapprendre à faire cela. Fut un temps, nous avions un programme nucléaire ambitieux, puis nous avons eu un programme ambitieux dans l'aéronautique. Aujourd'hui, il faut déployer un programme tout aussi vaste et ambitieux dans l'industrie de l'intelligence artificielle.

— Si l'on parle de chaine de valeur, on peut remonter un peu plus loin encore, jusqu'aux minerais, aux annotateurs et aux travailleurs invisibles. Se pose ici la question de l'IA éthique. Une IA éthique est-elle possible dans une organisation ou une chaine de valeur qui ne le serait pas ?

YF : L'éthique est évidemment un sujet majeur que nous devons porter collectivement à l'échelle de la France, de l'Europe et du monde. Chacun doit regarder ce qu'il fait et pourquoi il le fait. À titre personnel, le jour où j'ai décidé de quitter mon emploi, que j'adorais et pour lequel j'avais fait cinq ans d'études, afin de devenir entrepreneur, j'ai franchi ce pas pour avoir un impact positif. Les conditions de travail dans les minerais constituent un enjeu essentiel. Beaucoup d'autres situations sont tout aussi essentielles. Nous devons nous poser des questions en tant que citoyen, en tant que nation, en tant qu'Europe, en tant que citoyen du monde. Je vous rejoins, il existe des enjeux gigantesques qu'il nous faut affronter, et toute innovation n'a de sens que si elle nous permet d'améliorer nos conditions de travail à tous, notre qualité de vie, et pas le contraire. La technologie doit toujours être au service de l'humain. Les défis sont gigantesques. Et certains pensent que l'IA va résoudre tous les problèmes et nous transporter dans un monde prospère. Non. Pour moi, je le redis, chaque gramme de gain de productivité doit être mis collectivement au service de l'ensemble des individus pour essayer d'améliorer nos réponses et de relever tous ces défis.

Directrice de la publication : Caroline Gadou
Rédactrice en chef : Marion Gilles

Comité éditorial
Amandine Brugière (Anact), Nicolas Fraix (Anact)
Marion Gilles (Anact), Anne-Marie Nicot (Anact)
Clément Ruffier (Anact), Béatrice Sarazin (Anact)

Comité scientifique et technique
Cédric Dalmasso (Ecole des Mines Paris)
Algava Elizabeth (Dares)
Balas Stéphane (Cnam)
Blandin Olivier (Atemis-Lir)
Boboc Anca (Orange)
Bourguignon Rémi (Institut de recherche en gestion, Paris Est Créteil)
Datchary Caroline, (Lisst)
Folcher Viviane (Cnam)
Lhuillier Dominique (Cnam)
Parent-Thirion Agnès (Eurofound)

Recherche documentaire : Patricia Therry et Christine Veinhard - Mission Veille et Management de l'Information (VMI) de l'Anact
Secrétariat de rédaction : Vincent Degrez
Réalisation : Catherine Benini

© *La Revue des conditions de travail* est une marque déposée
Numéro national d'enregistrement INPI : 14/4119665

La Revue des conditions de travail

ANACT
192, avenue Thiers / CS 800 31 / 69457 LYON CEDEX 06
Tél. : 04 72 56 13 13
anact.fr

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

anact