

L'intelligence artificielle au service de la santé et de la sécurité au travail, enjeux et perspectives à l'horizon 2035

Journée prospective de l'INRS. Paris, 18 novembre 2022

AUTEURS:

C. Bigaillon, F. Ly, département Études et assistance médicales, INRS

EN RÉSUMÉ

Le 18 novembre 2022, l'INRS organisait une journée de restitution de ses travaux de prospective visant à explorer comment des systèmes mobilisant de l'intelligence artificielle (IA) pourraient être utilisés à des fins d'amélioration de la santé et de la sécurité au travail à l'horizon 2035. Cette journée a permis la présentation et la discussion des résultats d'une réflexion collective associant des experts de l'IA, divers acteurs de l'entreprise et des spécialistes de la prévention.

MOTS CLÉS

Technologie avancée / Organisation du travail / Santé au travail

En ouverture, *S. Pimbert, Directeur général de l'INRS*, a rappelé que ce 9^e exercice de prospective avait pour but d'explorer le futur des technologies d'intelligence artificielle (IA) sous l'angle des opportunités qu'elles pourraient représenter dans le domaine de la santé et la sécurité au travail (SST) à l'horizon 2035.

QU'EST-CE QUE L'IA ? DÉFINITIONS, HISTORIQUE ET PERSPECTIVES

J. Clerté (INRS) a commencé par rappeler que le nombre de publications scientifiques et de brevets sur le sujet de l'IA est exponentiel depuis les années 1980-1990. Le marché de l'IA s'est considérablement

développé avec des applications dans de multiples secteurs, dont la SST. Depuis la proposition de directive de la Commission Européenne, présentée le 21 avril 2021, qui vise à encadrer le développement de l'IA (*Artificial Intelligence Act*), les acteurs institutionnels, dont les partenaires sociaux, se sont emparés du sujet dans le but d'évaluer le risque éthique lié à son usage en environnement de travail et ont émis des recommandations. L'intérêt de l'INRS pour le sujet n'est pas nouveau car ses chercheurs utilisent déjà l'IA comme outil de travail, en particulier en toxicologie. La thématique de la prévention des risques professionnels liés à l'émergence de l'IA dans les machines y est également un objet d'études. Les objectifs de l'exercice de prospective étaient de se projeter dans le futur

L'intelligence artificielle au service de la santé et de la sécurité au travail, enjeux et perspectives à l'horizon 2035

et de réfléchir aux opportunités présentées par l'IA en SST. Les usages des systèmes d'IA spécifiquement dédiés à la prévention des risques professionnels ont donc été explorés en envisageant leurs atouts et les points de vigilance à retenir, et en proposant une liste de pistes d'actions pour permettre que son développement soit bénéfique à tous les acteurs de la prévention. Pour cela, 3 cas d'usage principaux ont été identifiés : le traitement intelligent de données massives, les technologies de sécurisation des environnements de travail et la robotique avancée recourant à l'IA. Les développements motivés par des objectifs autres que la préservation de la santé au travail n'étaient pas visés.

B. Braunshchweig (projet « Confiance. ai ») a rappelé que la première définition du concept d'IA remonte à la conférence de Dartmouth en 1956. L'IA y est définie par l'idée que l'on peut traduire sous forme d'algorithme tous les aspects de l'apprentissage ou toute caractéristique de l'intelligence et les simuler. Elle peut également être définie en extension par les disciplines qui la composent. Selon l'Académie des Technologies, il existe 5 grandes disciplines qui contribuent à l'IA : le raisonnement logique, la représentation des connaissances, la planification automatique, le traitement du langage naturel et la perception. Il existe, de plus, deux grandes tendances en IA : l'IA symbolique, faite « manuellement », explicable mais coûteuse et l'IA numérique par apprentissage automatique, qui génère des modèles à partir de données. Cette dernière est peu compréhensible car elle fonctionne comme une boîte noire. Un bref historique de l'IA montre que, même si son développement est croissant depuis le milieu du XX^e siècle, son histoire est émaillée de

périodes d'hiver (arrêts). Les progrès actuels remarquables de l'IA dans différents domaines (reconnaissance d'images, traitement de la parole, langue naturelle, jeux, aide à la décision...) peuvent, dans le futur, se heurter à nouveau à différents murs dont n'importe lequel peut entraîner un nouvel hiver : la confiance en l'IA, la consommation d'énergie, la sécurité, les interactions homme-machine et le besoin de compréhension mutuelle, le mur d'inhumanité c'est-à-dire l'absence de sens commun de la machine.

N. Fatès (Institut national de recherche en sciences et technologies du numérique/Laboratoire lorrain de recherche en informatique et ses applications – Inria/Loria) a expliqué que le groupe de travail n'a pas retenu une unique définition consensuelle de l'IA car ses contours sont imprécis. Entre autres, en 2018, le rapport Villani désignait l'IA « moins comme un champ de recherches bien défini, qu'un programme fondé autour d'un objectif ambitieux : comprendre comment fonctionne la cognition humaine et la reproduire ; créer des processus cognitifs comparables à ceux de l'être humain ». Bien que le terme IA puisse mettre mal à l'aise car il applique une faculté du vivant à la machine, il s'agit d'un mouvement en avant qui engage toute la société. Après un rappel étymologique, N. Fatès a résumé les bases fondamentales de l'IA à partir des travaux d'Alan Turing depuis 1936.

PRÉSENTATION DE L'EXERCICE DE PROSPECTIVE

LA MÉTHODOLOGIE

M. Héry (INRS) a rappelé que la prospective à l'INRS a pour objectif

d'identifier les enjeux qui auront des conséquences en SST dans les décennies à venir. Il s'agit d'un exercice pluridisciplinaire qui permet également de développer des contacts avec l'extérieur et de favoriser les collaborations internes entre disciplines et métiers. Après une première étape de définition du sujet et de son périmètre, la deuxième phase de cet exercice a consisté à élaborer des scénarios de futurs possibles des usages de l'IA d'ici à 2035. Différentes hypothèses d'évolution ont été envisagées pour 12 variables, dont la combinaison a permis d'aboutir à quatre scénarios contrastés. Ils ont permis de se projeter concrètement en 2035 dans différents contextes et ont servi d'outils de travail pour la troisième étape. Celle-ci a consisté à identifier les enjeux à partir des matériaux recueillis. Dans ce but, 3 ateliers consacrés aux cas d'usages présentés précédemment par **J. Clerté** ont été organisés afin d'approfondir la réflexion concernant les enjeux de SST et de prévention des risques professionnels. Cette dernière étape a permis d'aboutir à 22 messages clés, principaux enseignements issus de cette démarche, classés en 4 rubriques : un marché en plein développement, des promesses en SST, les limites et points de vigilance sur les usages de l'IA en SST et les pistes d'action.

LES SCÉNARIOS DE PROSPECTIVE

Les 4 scénarios ont été illustrés par des présentations (sous forme journalistique) de situations fictionnelles avant d'être explicités.

J. Clerté est intervenue sur le premier scénario « **Les géants du numérique imposent leurs solutions et leur vision** », illustré par l'exemple des failles d'un système technologique de prévention de

l'endormissement des chauffeurs de poids lourds. Dans ce scénario, l'exubérance technologique se poursuit dans un cadre de compétition pour la maîtrise de l'intelligence artificielle, mais les acteurs principaux en sont les géants du numérique, essentiellement les AMAMA¹ à l'Ouest et les BATX² en Chine. Leur puissance dépasse largement le champ du numérique et de fait ils contrôlent l'essentiel de l'innovation et dominent des pans entiers de l'économie mondiale. Dans un contexte de rivalité internationale systémique, les États doivent composer avec ces acteurs et s'appuyer sur eux pour maintenir leur puissance comme pour assurer les fonctions clés de leur souveraineté. La réglementation est morcelée entre les États et largement influencée par ces géants. Les utilisateurs acceptent ces normes par commodité et parce qu'elles sont devenues indispensables. L'automatisation progresse et la surveillance s'installe comme outil privilégié de la sécurité au travail. Celle-ci justifie un contrôle des activités de chacun dans un contexte de collaboration étroite entre les hommes et les machines.

M. Héry a présenté le deuxième scénario « **Les États garantissent un cadre pour l'intégration de l'IA** » à partir d'un exemple de suspicion de discrimination à l'embauche par une application officielle pour l'emploi étudiant. Le développement de l'IA sans maîtrise a poussé les États européens à travailler sur un règlement commun afin de pouvoir en encadrer l'écosystème et les principes éthiques. Au cours de la période, la multiplication des problèmes entraînant des préjudices pour les citoyens, les entreprises et les travailleurs les a poussés à durcir cet encadrement dans un contexte de montée

1. *Alphabet, Meta, Amazon, Microsoft, Apple*

2. *Baidu, Alibaba, Tencent, Xiaomi*

des préoccupations environnementales. Les États, afin d'orienter l'allocation des ressources, ont décidé de ne développer que des systèmes d'IA sobres, répondant à des critères élevés (notamment autour du concept d'intérêt général), sur des secteurs non critiques et sous supervision humaine. Cela passe par un meilleur contrôle des fabricants, par le développement de pôles d'expertises européens et nationaux, mais également par des expérimentations qui permettent de démontrer l'innocuité de l'IA et des mesures de contrôle et d'audit. L'IA au travail se développe donc dans un environnement assez encadré, permettant une implémentation, dès lors que l'intérêt et la non dangerosité des dispositifs ont été approuvés.

A. Olympio (CNP assurance) a commenté le troisième scénario « **Développement démocratique** » sur le thème de l'utilisation de robots autonomes pour l'exploitation d'une mine souterraine de lithium. Dans ce scénario, les années 2020 voient se mettre en place des processus de contrôle démocratique par les travailleurs et citoyens, nécessaires au bon développement des IA et à leur expansion encadrée dans la vie civile et professionnelle. Dans un contexte de croissance économique mondiale, générant de l'emploi dans l'industrie et les services, et permettant d'investir dans la formation, les conditions sont réunies pour que les systèmes d'IA se déploient largement dans le monde du travail, et concourent au façonnage progressif d'une maîtrise collective de ces projets technologiques. Le recours aux dispositifs d'IA est facilité *via* l'essor des outils *open source* et le développement de solutions très accessibles (*low code, nocode*). De plus, les recherches

en IA lancées depuis les années 2010 finissent par aboutir dans les années 2030 à la conception de systèmes d'IA hybrides combinant la puissance de l'apprentissage automatique à la transparence des systèmes de raisonnement logique. En restaurant le principe éthique d'explicabilité comme clef d'appropriation, ces résultats contribuent à construire la confiance collective dans l'IA et à la mettre au service de la performance, de la santé et de la sécurité dans les organisations du travail.

M. Malenfer (INRS) est intervenu sur le quatrième scénario « **Déclin de l'IA** », présentant un conflit juridique autour de l'usage d'équipements de brouillage individuels permettant de déjouer les systèmes de surveillance dans l'entreprise. En début de période, les usages des systèmes d'IA se développent dans tous les domaines professionnels. Poussée par les progrès technologiques, la numérisation généralisée de la société et les nouvelles organisations de travail, l'IA est majoritairement bien acceptée dans le monde du travail. Depuis 2022, elle est considérée comme un atout pour les employeurs (automatisation, productivité, qualité...) et pour les travailleurs (pénibilité, sécurité...). Cette considération est principalement basée sur les promesses que les systèmes d'IA apportent. Progressivement, la déception face aux applications de terrain, les failles de ces systèmes qui provoquent des incidents, accidents ou crises, génèrent un rejet des systèmes d'IA dans le monde du travail. À partir de 2030, ce rejet amène à un déclin de cette technologie et de ses usages professionnels.

LES DOMAINES D'USAGE

En complément des scénarios, des ateliers portant sur des domaines

L'intelligence artificielle au service de la santé et de la sécurité au travail, enjeux et perspectives à l'horizon 2035

d'usage possibles des systèmes d'IA pour la SST ont été organisés.

M. Bieri (Laboratoire d'innovation numérique de la Commission nationale de l'informatique et des libertés – CNIL/LINC) a présenté leur utilisation en épidémiologie grâce aux systèmes qui permettent les croisements de bases de données (ex. : étude de l'INRS sur la polyexposition afin de mettre en lien expositions et pathologies) et en accidentologie autour des systèmes de traitement automatique du langage pour exploiter des données mal structurées ou textuelles (ex. : usage par la Caisse régionale d'assurance maladie d'Île-de-France – CRAMIF – pour exploiter le contenu des déclarations d'accidents du travail). Les conditions nécessaires au développement de ces outils sont la disponibilité des données et la volonté de les partager, la mise en place d'un cadre réglementaire pour leur protection, la capacité de stockage et de traitement ainsi que la fiabilité des systèmes. Leur utilisation permettrait des avancées en faisant progresser l'évaluation et l'analyse des risques et en apportant un meilleur suivi de la santé des travailleurs. Il faudra cependant être vigilant quant à la sécurisation des données et à la transparence lors de leur utilisation. Il existe également un risque de focalisation sur des facteurs individuels de santé et de centralisation du pilotage de la SST. Pour les acteurs de la SST, l'utilisation de ces outils nécessitera une montée en compétence et une implication dans leur développement sans se focaliser uniquement sur leur utilisation au détriment d'autres missions.

T. Silvestre (Commissariat à l'énergie atomique et aux énergies alternatives – CEA) a présenté le

deuxième cas d'usage, sur les techniques de sécurisation des environnements de travail recourant à l'IA. L'objectif est d'utiliser cette dernière en temps réel afin d'alerter des dangers imminents (ex. : systèmes de détection de piétons ou de gaz toxiques) ou de prévenir des pathologies chroniques en informant les opérateurs par l'analyse de flux d'informations ou de données biométriques (ex. : équipements de protection individuelle connectés). Les conditions nécessaires au développement de ces outils sont l'évolution des puissances de calcul pour miniaturiser les équipements et réduire leur consommation d'énergie et les coûts ainsi que le déploiement de réseaux de communication performants. Pour la SST, ces nouveaux outils permettraient de diminuer la sinistralité, d'assurer une meilleure prévention pour les salariés encore peu sensibilisés et de développer des actions ciblées de formation. Ces systèmes ne doivent cependant pas faire baisser la vigilance en laissant de côté les facteurs organisationnels ou être détournés vers des actions de surveillance des salariés. Pour les acteurs de la SST, les bénéfices seraient un meilleur ciblage des facteurs de risque sur le terrain, le développement de la recherche et la formation des préventeurs. Ces outils ne se substituent ni à l'évaluation des risques par l'entreprise, ni aux actions de prévention.

M. Sarrey (INRS) a présenté deux cas d'usage de l'IA en robotique avancée. Le premier se situe dans le domaine de la télé-opération, qui permet au robot de réaliser des missions autonomes. L'opérateur est ainsi éloigné de la situation de travail dangereuse (ex. : test d'un robot pour inspecter les galeries d'un site d'enfouissement des déchets nucléaires de l'Agence nationale

pour la gestion des déchets radioactifs). Le second cas d'usage est la collaboration entre des opérateurs humains et des machines ayant des perceptions améliorées de l'environnement (ex. : détection de la présence de personnes dans l'environnement du robot pour une réaction adaptée afin d'éviter l'accident). Les conditions nécessaires au développement de ces équipements sont leur démocratisation, leur intégration dans les organisations de travail et leur acceptation par les opérateurs. L'IA permettrait ainsi de supprimer certains risques et de maintenir au poste des opérateurs qui ne seraient pas en capacité de réaliser les tâches physiques liées à la mission ou de les faire monter en compétence. Cependant, il faudrait veiller au risque d'intensification des tâches de l'opérateur, de perte des gestes métiers et à la capacité à gérer les situations anormales. Pour les acteurs de la SST, il sera nécessaire d'accompagner le déploiement de ces techniques pour s'assurer de la pertinence de leur emploi, de leur robustesse et de leur innocuité.

LES MESSAGES CLÉS ISSUS DE L'EXERCICE DE PROSPECTIVE

M. Sarrey a développé les points relatifs au développement du marché des systèmes d'IA pour la SST. Du point de vue de la prévention, l'IA offre la possibilité d'automatiser des tâches et de soustraire les travailleurs à certains risques. Les systèmes développés devront être compatibles avec les valeurs françaises et européennes de la SST. Cependant, ces technologies n'ayant pas encore fait leur preuve, toutes les avancées en SST ne doivent pas reposer uniquement

sur l'IA. Afin qu'ils tiennent toutes leurs promesses, il est nécessaire de favoriser le développement d'outils transparents et compréhensibles par tous.

M. Malenfer a présenté les promesses en SST liées aux progrès en matière d'IA. Il existe des domaines où l'IA est déjà utilisée ou expérimentée comme en toxicologie ou en épidémiologie. Au niveau des entreprises, les solutions de sécurisation seront d'autant plus intéressantes dans la prévention des accidents qu'elles s'intégreront dans une démarche d'amélioration de l'organisation du travail. Enfin, le développement de robots autonomes permettra de soustraire l'homme à certains risques. Ces développements ne devraient pas entraîner une destruction de l'emploi dans le domaine de la SST mais ils transformeront les métiers avec une montée en compétence des acteurs. Tous ces usages ne sont cependant pas anodins.

M. Héry a exposé les limites et points de vigilance sur les usages de l'IA en SST. Il a rappelé que la mise en place d'une démarche de prévention s'appuie sur les neuf principes généraux de prévention (L.4121-2 du Code du travail) et que le 4^e principe « *adapter le travail à l'homme dans le but de réduire les effets du travail sur la santé* » doit rester prioritaire. Ces technologies ne doivent pas avoir pour conséquence la mise au second plan du travail humain. L'absence de réflexion préalable à leur mise en place pourrait générer de nouveaux risques organisationnels. Cet exercice de prospective est l'occasion de rappeler qu'il faut rester critique face à la technologie: il ne serait pas acceptable qu'en réduisant certains risques, l'IA en génère d'autres (ex.: risques psychosociaux (RPS)). Les

situations atypiques ou dégradées durant lesquelles surviennent fréquemment les accidents du travail sont une limite à l'utilisation de l'IA. Une vigilance particulière doit être apportée au choix des données servant à l'entraînement des systèmes afin qu'ils correspondent au domaine d'applicabilité qui peut varier selon les activités et les situations de travail.

J. Clerté a présenté les pistes d'actions que le groupe a identifiées. Un effort d'investissement devra porter sur la formation des acteurs de la prévention et du dialogue social afin que les systèmes choisis et déployés dans l'entreprise répondent aux besoins et que leurs modes de fonctionnement soient compris. Il faudra également sensibiliser les développeurs et concepteurs au respect des principes de prévention dans les algorithmes. Une attention particulière devra être portée à la réglementation encadrant l'IA afin qu'elle prenne en compte les principes de SST pour le développement d'outils sûrs. La question de la sécurisation des données collectées, nécessaires au fonctionnement de ces systèmes, est capitale. Une réflexion collective experts-partenaires sociaux sur ce sujet semble nécessaire afin de définir des règles pour la constitution des jeux de données et l'encadrement de leur utilisation. Les organismes de prévention doivent dès aujourd'hui travailler à développer des outils pour guider les acteurs.

DISCUSSION SUR LES ENJEUX

Une première séquence de discussion a permis de donner la parole à **R. Chatila** (*Institut des systèmes*

intelligents et de robotique – ISIR) et **J.G. Ganascia** (*Centre national de la recherche scientifique – CNRS*) afin qu'ils puissent commenter les résultats présentés.

R. Chatila rappelle que, même si la tendance est de personnifier l'IA, celle-ci n'a pas de capacité de compréhension. C'est un outil perfectionné qui doit rester sous supervision humaine. Il est nécessaire d'être conscient de ses limites. Elle peut soit augmenter la charge cognitive de l'humain pour permettre sa compréhension soit la diminuer. Elle doit simplifier le travail de l'humain et être une aide, selon le concept américain des 3D, pour effectuer les tâches « *Dull* » (inintéressantes), « *Dirty* » (sales, insalubres), « *Dangerous* » (dangereuses).

J.G. Ganascia explique, au sujet des limites de l'IA, que ces technologies transforment l'organisation sociale et que le système doit être pensé au regard de l'ensemble de l'organisation sociotechnique. L'IA va-t-elle remplacer l'homme, induire une perte du sens du travail ou exclure certaines personnes dans une logique de rentabilité? L'expérience montre que le numérique n'a pas réduit l'homme à l'inaction mais qu'il a déplacé son activité, modifié les situations de travail ou le rythme bien qu'il puisse générer parfois d'autres problèmes. **R. Chatila** cite l'exemple de l'utilisation de robots dans des entrepôts de logistique qui ont vu augmenter les troubles musculosquelettiques (TMS). En effet, les salariés n'avaient plus à se déplacer mais effectuaient des gestes de conditionnement à une cadence accélérée du fait de l'arrivée en continu des produits déplacés par les robots. Le système doit être évalué dans sa globalité tout au long de son déploiement. Il sera

L'intelligence artificielle au service de la santé et de la sécurité au travail, enjeux et perspectives à l'horizon 2035

nécessaire de former et d'accompagner le changement. En SST, l'IA doit être un outil qui permettra de mieux appréhender les risques, pas un remplaçant.

J.G. Ganascia insiste sur le fait que l'IA est basée sur un système d'apprentissage continu et qu'il est nécessaire d'obtenir de façon systématique des retours d'usage de l'ensemble des utilisateurs. *R. Chatila* précise que l'on est dans l'expérimentation permanente et qu'il faut se réserver la possibilité de revenir en arrière. Il faut également mener ces expérimentations dans des conditions d'éthique selon les principes de non malversation, de justice/équité, d'intelligibilité, d'autonomie de l'être humain et de concertation lors de la mise en place. Certains principes peuvent cependant être difficiles à appliquer car abstraits. *J.G. Ganascia* explique que, du point de vue scientifique, les systèmes d'IA sont des logiciels capables de faire des erreurs et qu'il est difficile d'apporter la preuve de leur absence d'erreur. Leur expérimentation peut reposer sur la réalisation de tests, une infinité de tests ressemblant à une preuve. Il faut étudier les modes d'appropriation de ces technologies au sein de l'entreprise et se demander si les effets sont bénéfiques ou délétères sur l'organisation du travail.

R. Chatila rapporte le fait qu'il n'existe pas encore de réglementation propre à l'IA applicable à ce jour, mais qu'il existe une proposition d'un groupe d'expert de la Commission Européenne qui est en cours de discussion. Elle pourrait être disponible en 2025 et permettrait de définir le domaine de l'IA et la notion de risque. Elle devra cependant ne pas sur-réglementer pour ne pas bloquer l'innovation.

Dans ce but, la réglementation pourrait s'axer majoritairement sur les domaines d'application qui mettent en jeu la santé et la sécurité (ex. : domaine des transports, de la santé...). En ce qui concerne le respect de l'éthique, *J.G. Ganascia* fait référence au rapport du groupe de travail n° 3 du ministère chargé de la Santé³ sur la cellule éthique qui a émis des recommandations en 4 points : les données, la garantie humaine (la responsabilité de la décision revient à l'homme), la notion d'évaluation et la traçabilité. *R. Chatila* rappelle que le principe d'éthique *by design* consiste à mettre en place une démarche qui respecte les principes éthiques et tient compte des valeurs humaines, l'humain étant au centre de l'utilisation du système.

EXEMPLE FICTIF DE L'UTILISATION D'UNE IA EN SANTÉ ET SÉCURITÉ AU TRAVAIL

Lors de ce sketch, *M. Sarrey* présente Mi-OSH, un ordinateur ressemblant à un jouet soit disant doté d'une IA qui aurait la faculté de répondre à des questions d'assistance en matière de SST, comme le ferait un expert de l'INRS. Il est expérimenté en direct avec des questions posées, et ses réponses sont commentées par *N. Fatès*.

Mi-OSH a conquis le public qui dans sa grande majorité l'a trouvé crédible, les questions et leurs réponses ayant été rédigées à l'avance. L'exemple permet ainsi d'illustrer la vigilance qu'il faut avoir vis-à-vis de ces technologies qui peuvent tendre à invisibiliser le travail humain et conduire à certaines dérives, notamment en matière de protection des données.

3. https://esante.gouv.fr/sites/default/files/media_entity/documents/ethic_by_design_guide_vf.pdf

DISCUSSION DES RÉSULTATS DE L'EXERCICE DE PROSPECTIVE

Une première table ronde a réuni *L. Laurent, directeur des études et recherches de l'INRS* ; *Y. Ferguson, sociologue à l'Institut catholique d'Arts et métiers de Toulouse et coordonnateur du programme LaborIA – Icam/LaborIA* ; *A. Ponce del Castillo, avocate à l'Institut syndical européen – ETUI* ; *A. Parent-Thirion, responsable de l'enquête européenne sur les conditions de travail à la Fondation européenne pour l'amélioration des conditions de travail – Eurofound à Dublin*.

Les discussions ont notamment porté sur :

- la question de la protection des données dans le cadre professionnel où le rapport de subordination fait que l'employé aura une difficulté à donner un consentement éclairé face à son employeur qui voudrait imposer un système d'IA ;
- la nécessité de réfléchir en amont sur l'utilité et les conséquences possibles que pourrait engendrer l'utilisation de ces systèmes ;
- la nécessité de poursuivre les efforts à fournir en terme de prévention, 33 % des personnes au travail rapportant qu'un risque existe pour leur santé/sécurité ;
- l'intérêt à ajouter l'IA à la palette d'outils de prévention en identifiant les applications qui sont les plus susceptibles de rapporter des bénéfices face à une population de travailleurs et des situations de travail très différentes ;
- l'instauration d'un dialogue entre les différents acteurs pour être informés et intégrer l'IA afin qu'elle puisse être utilisée au mieux ;
- les études et recherches faites à l'INRS dans le but de se projeter sur les postes où seront utilisées les

machines embarquant de l'IA afin d'analyser tous les éléments que cela impliquera, allant de la technique de la machine à l'organisation, la perception et l'acceptabilité des nouveaux systèmes par les utilisateurs. Cela est important pour voir les risques que cela pourrait engendrer, puis de les évaluer dans les conditions réelles d'utilisation;

- la nécessité d'un environnement de travail bien pensé pour intégrer une IA et créer un collectif d'appropriation;

- l'intérêt des principes définis par l'Organisation de coopération et de développement économiques (OCDE) pour une IA : transparente, inclusive, robuste, sécurisée et qui favorise le bien être;

- les questions communes qui reviennent dans différents groupes de travail internationaux : la fiabilité, la responsabilité ou encore les limites en matière de collecte de données.

Toutes ces questions doivent être traitées et réfléchies en amont tout en continuant à être évaluées en situation réelle une fois les systèmes mis en place.

INITIATIVES POUR LE DIALOGUE SOCIAL SUR L'INTELLIGENCE ARTIFICIELLE

Cette séquence réunissant **V. Mandinaud (Agence nationale pour l'amélioration des conditions de travail – ANACT)**; **Y. Ferguson (Icam/LaborIA)** et **P. Balzer (Caisse d'assurance retraite et de la santé au travail – CARSAT – Midi-Pyrénées)** a porté sur les initiatives en matière de promotion du dialogue social sur l'IA par la formation des partenaires sociaux sur ces sujets.

La CARSAT Midi-Pyrénées a établi un partenariat avec l'ICAM/LaborIA pour développer un dispositif de formation/sensibilisation des partenaires sociaux aux conséquences de l'IA sur le travail. Un premier temps permet de faire de la pédagogie en développant les concepts de l'IA : comprendre, nommer, qualifier ce qu'est l'IA aujourd'hui, et ce qu'elle n'est pas. Dans un deuxième temps, sont développés les rapports entre l'IA et le travail, notamment en déconstruisant l'idée que l'IA mènerait à la fin du travail. Le dernier temps est celui de la simulation de réunion d'un Comité social et économique (CSE) portant sur un projet d'implantation d'un système d'IA. L'idée est de sortir d'une approche de peur pour aller vers une approche de problèmes à régler sous l'angle de la SST. Cette expérimentation a été très bien reçue et la CARSAT Midi-Pyrénées souhaite pouvoir démultiplier l'expérience.

Dans ce cadre, l'ICAM et la CARSAT ont répondu à l'appel à manifestation d'intérêt (AMI) de l'ANACT « Développer le dialogue social technologique au travail ». Il s'agit de permettre un dialogue entre les acteurs afin d'explorer les effets des transformations technologiques et organisationnelles sur le travail, pour que les opérateurs reprennent la main sur ces transformations. Cela passe par une articulation des dialogues : dialogues institués entre les représentants du personnel et de la direction mais aussi d'autres formes de dialogues comme des dialogues managériaux ou des dialogues entres professionnels qui voient leur métier changer.

DISCUSSION PAR DES PARTENAIRES SOCIAUX

Cette dernière table ronde a réuni **F. Salis-Madinier (CFDT et membre du Conseil économique et social européen – CESE)**; **N. Blanc (CFE-CGC et membre expert de l'observatoire de l'IA à l'OCDE)**, et **B. Frugier (Fédération des industries mécaniques – FIM)**.

Aujourd'hui, les organisations syndicales se rejoignent sur la nécessité de démystifier la technologie, de mieux la comprendre pour se l'approprier de façon collective, le plus en amont possible.

Au-delà de la question de l'amélioration des performances et des gains de productivité, les sujets du temps libéré, les enjeux sur l'emploi et la collecte des données doivent être discutés afin de permettre aux salariés d'être acteurs de ces transformations et de copiloter les projets d'intégration d'IA dans les entreprises.

L'usage de ces outils peut aussi être ambivalent puisqu'ils peuvent à la fois améliorer le travail mais aussi servir à surveiller l'opérateur. Aujourd'hui, il y a un enjeu d'appropriation collective sur le pourquoi de l'adoption d'un système d'IA. Or il y a souvent carence de dialogue social lors de l'introduction du numérique en entreprise.

CONCLUSION

B. Salengro (Président du Conseil d'administration de l'INRS) rappelle qu'il s'agit, pour les acteurs de la prévention, de construire ensemble un monde du travail soutenable et décent.

L'objectif de cet exercice de prospective consacré à l'IA a été d'essayer

L'intelligence artificielle au service de la santé et de la sécurité au travail, enjeux et perspectives à l'horizon 2035

de détecter les opportunités que peuvent apporter ces innovations pour améliorer la prévention. Les promesses dans ce domaine sont réelles tout comme les modifications structurelles générées. Les discours sont variés et vont du plus alarmiste au plus optimiste, d'où la nécessité de débattre. Les systèmes d'IA sont à regarder pour ce qu'ils sont et il ne faut pas leur prêter des qualités humaines qu'ils n'ont pas. Il faut avoir à l'esprit que ces systèmes sont développés par des humains à des fins particulières et qu'ils peuvent parfois dévier de l'objectif qui leur a été assigné ; d'où la nécessité de travaux sur la fiabilité. Ces outils peuvent aussi emmener vers une approche de plus en plus individuelle du travail, conduisant à une perte de sens et à l'isolement. Il ne faut pas que cela se développe au détriment des approches collectives de la prévention.

Des pistes pour agir existent :

- tirer parti des capacités offertes par certains systèmes d'IA pour faire progresser les connaissances ;
- promouvoir des outils pour accompagner les entreprises, notamment les plus petites, et savoir les guider dans l'appropriation de ces systèmes ;
- former pour faire comprendre comment fonctionnent ces outils.

L'exercice de prospective complet est présenté dans la brochure «L'Intelligence artificielle au service de la santé et sécurité au travail. Enjeux et perspectives à l'horizon 2035», consultable sur le site de l'INRS (<https://www.inrs.fr/media.html?refINRS=Pv%2019>).

Le *replay* de cette journée peut être visionné sur la chaîne Youtube de l'INRS (<https://www.youtube.com/playlist?list=PLqlw8IH6G3t27jPFT5DeJAtXkLMK2ZFiU>).