

L'INTELLIGENCE ARTIFICIELLE AU SERVICE DE LA SANTÉ ET SÉCURITÉ AU TRAVAIL : ENJEUX ET PERSPECTIVES À L'HORIZON 2035

L'introduction de systèmes d'intelligence artificielle en milieu professionnel est une réalité qui soulève des questions en matière de prévention des risques. L'INRS a conduit un exercice de prospective pour identifier les enjeux et les perspectives d'évolution de l'usage de ces technologies dans le champ de la santé et sécurité au travail à l'horizon 2035. Le périmètre de cette réflexion pluridisciplinaire est donc centré sur des domaines pouvant impacter la prévention des risques professionnels. Cet article propose un éclairage sur les travaux conduits, qui reposent sur l'exploration de différents scénarios d'évolution possibles et des projections dans trois domaines d'usages. L'ensemble des enseignements est rassemblé en une vingtaine de messages clés.

ARTIFICIAL INTELLIGENCE IN THE SERVICE OF OCCUPATIONAL SAFETY AND HEALTH: CHALLENGES AND PROSPECTS FOR 2035 – *The introduction of artificial intelligence systems in the professional environment is a reality that raises questions in terms of risk prevention. INRS conducted a prospective exercise to identify the challenges and future prospects for the use of these technologies in the field of occupational safety and health by 2035. The scope of this multidisciplinary reflection is therefore focused on the areas of use that can affect occupational risk prevention. This article sheds light on work conducted, which is based on the exploration of different possible development scenarios and projections in three areas of use. All of the lessons drawn are compiled in a number of key messages.*

MARC
MALENFER,
MICHEL
HÉRY,
JENNIFER
CLERTÉ
INRS, mission
Veille et
prospective

MICHAEL
SARREY
INRS,
département
Ingénierie des
équipements
de travail

Contexte et démarche de l'exercice de prospective

Parmi les neuf principes généraux de prévention qui guident l'action des acteurs de l'entreprise en matière de préservation de la santé et de la sécurité des travailleurs, figure en cinquième position celui-ci : «tenir compte de l'état d'évolution de la technique¹». Le rythme et le foisonnement des innovations rendent cette tâche particulièrement ardue, d'autant plus qu'il s'agit également de respecter les autres principes généraux, et notamment le deuxième : «évaluer les risques qui ne peuvent pas être évités». Une évaluation rigoureuse des risques nécessite du temps et place le décideur dans une position délicate face aux innovations technologiques. Il s'agit pour lui

d'arriver à les utiliser pour améliorer la productivité et les conditions de travail, tout en s'assurant qu'elles n'engendrent pas de nouveaux risques.

En matière de technologies de l'information et de la communication (TIC), le domaine de l'intelligence artificielle (IA) mobilise énormément de moyens et d'attention depuis plusieurs années. Que ce soit à des fins d'amélioration de la productivité [1], dans l'optique de diminuer la vulnérabilité des chaînes d'approvisionnement [2] ou pour des raisons géopolitiques de relocalisation d'activités dans des secteurs stratégiques [3], le fait de pouvoir disposer de systèmes d'IA est perçu comme un atout qu'il convient de développer, y compris pour la gestion des risques professionnels [4,5].



Les acteurs de la santé et sécurité au travail (SST) s'intéressent naturellement aux conséquences que peut avoir l'introduction de systèmes d'IA en milieu professionnel. Il s'agit généralement d'essayer d'évaluer les risques que peuvent faire peser ces dispositifs sur la santé physique et mentale des travailleurs et d'émettre des préconisations de prévention. Ces technologies sont d'ores et déjà mobilisées sur certains lieux de travail, à un niveau encore marginal, mais elles soulèvent déjà des questions. L'INRS prend logiquement part à ce travail et la mission Veille et prospective de l'Institut a déjà traité de sujets impliquant l'utilisation de nouvelles technologies en général [6,7] et des robots d'assistance physique en particulier [8]. Le parti pris de l'exercice de prospective, dont une synthèse des résultats est présentée dans cet article, est de se focaliser sur les usages possibles des dispositifs d'IA à des fins de protection de la SST, à un horizon d'une douzaine d'années. La démarche suivie reprend les fondamentaux de la pratique de la prospective à l'INRS : un travail collaboratif et pluridisciplinaire, impliquant de nombreux experts internes et externes à l'Institut ; l'application d'un cadre méthodologique permettant de donner à voir différents futurs possibles, tous cohérents, mais pas toujours souhaitables ; un approfondissement de certains enjeux et l'extraction de messages clés destinés aux acteurs de la prévention, au premier rang desquels figurent les partenaires sociaux. L'objectif est de leur fournir des éléments de réflexion et des pistes d'action pour les aider à préparer l'avenir.

Éléments de définition

Intelligence artificielle, de quoi parle-t-on ?

Il existe de très nombreuses définitions de l'IA, aucune ne réglant totalement la question car la discipline présente des contours flous. Pour les besoins de cet exercice, le groupe de travail² a retenu trois définitions, qui sont compatibles entre elles. Les deux premières sont des définitions par *intension*³. Celle proposée par la Commission d'enrichissement de la langue française et publiée au Journal officiel du 9 décembre 2018 décrit un « *Champ interdisciplinaire théorique et pratique qui a pour objet la compréhension de mécanismes de la cognition et de la réflexion, et leur imitation par un dispositif matériel et logiciel, à des fins d'assistance ou de substitution à des activités humaines.* »

La seconde est celle publiée par la Commission européenne, la même année : « *Même s'il n'y a pas de définition consensuelle, le terme "IA" fait référence à des systèmes qui exhibent des comportements intelligents en analysant leur environnement et en effectuant des actions avec un certain degré d'autonomie, pour atteindre des objectifs spécifiques. Les systèmes d'IA peuvent être purement logiciels, agissant sur le monde virtuel. Des exemples sont*

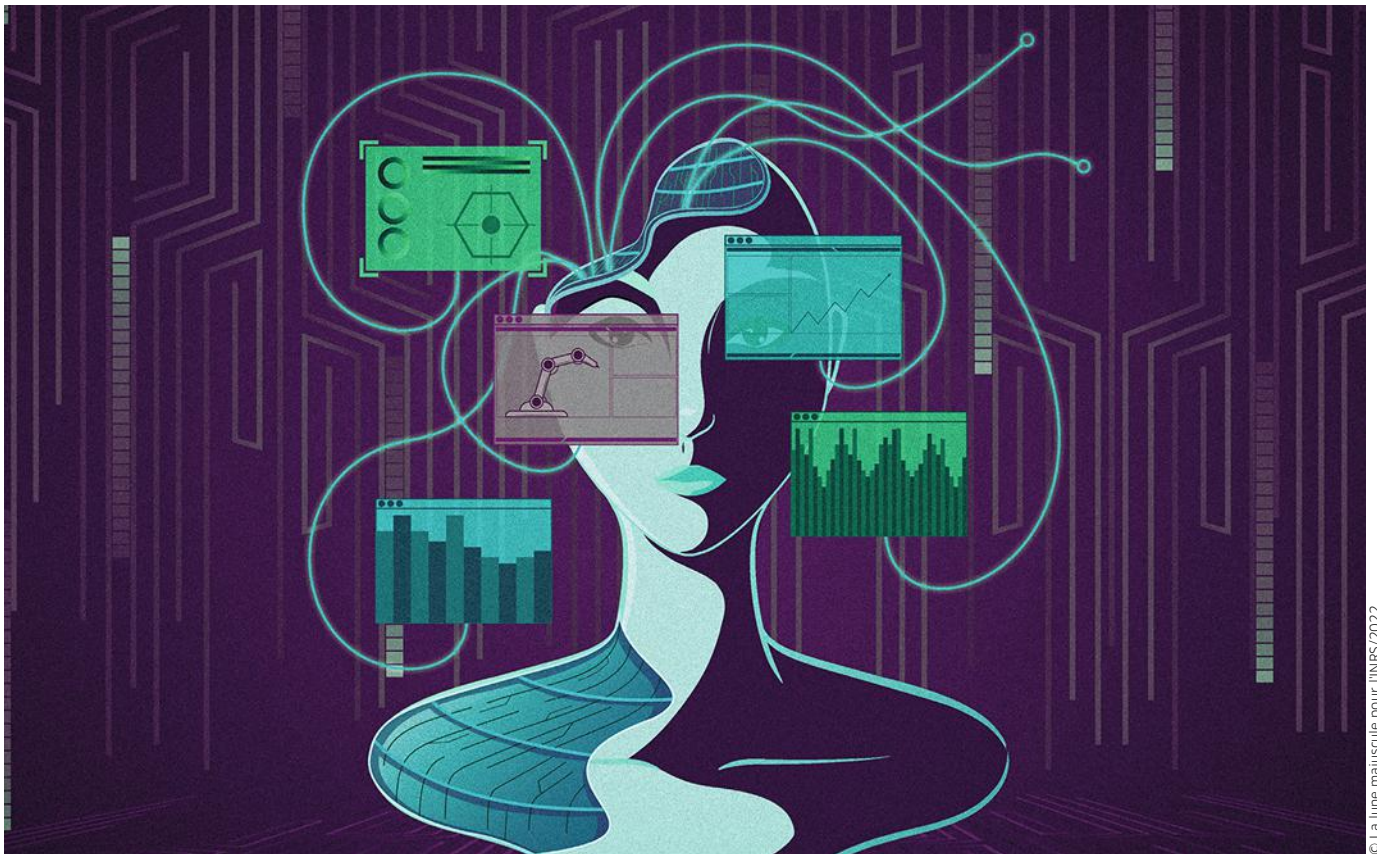
les agents conversationnels, les logiciels d'analyse d'image, les moteurs de recherche et les systèmes de reconnaissance de visages. D'autres systèmes d'IA sont ceux où l'IA est embarquée dans des appareils comme les robots avancés, les voitures autonomes, les drones ou les applications de l' "Internet des objets". En contraste avec les systèmes d'IA logiciels, de tels systèmes perçoivent leur environnement grâce à des capteurs, et agissent dessus ou s'y déplacent, exigeant donc des mesures de sécurité plus robustes. » La troisième relève d'une autre approche qui consiste à définir l'IA en extension, par la somme des diverses sous-disciplines qui la composent : apprentissage automatique, raisonnement automatique, traitement de la langue naturelle, vision artificielle, représentation des connaissances, etc. C'est le choix fait par l'Académie des technologies dans son rapport « *Renouveau de l'Intelligence artificielle et de l'apprentissage automatique* », également publié en 2018⁴ :

- perception, un domaine très vaste contenant la reconnaissance des images, des formes, et des sons (et de toutes les formes de signaux produits par des « senseurs »). La reconnaissance des images est le domaine le plus important (on parle de machine vision) à cause de ses applications et de par les progrès spectaculaires de ces dernières années, cités en introduction, et liés au *deep learning* ;
- traitement du langage naturel, sous sa forme écrite ou orale (ce qui implique de le coupler avec la perception de la parole) ;
- planning et navigation, qui pourraient être étendus aux domaines de la résolution de problèmes formalisés dans l'univers de la recherche opérationnelle ;
- représentation de connaissances, qu'il s'agisse de les manipuler, de les rechercher ou de les combiner. Ce domaine a fortement évolué avec le développement des mégadonnées (*Big Data*) ;
- raisonnement logique, liée à la forme symbolique de l'IA, rendue célèbre à partir des années 1980 par les « systèmes experts. »

On retrouve notamment, dans ces trois définitions complémentaires, les capacités données à des machines (matériel et logiciel) de réaliser des tâches demandant de l'intelligence quand elles sont exercées par des humains, des exemples d'applications de ces capacités (véhicule autonome, agent conversationnel, reconnaissance d'images, etc.), et des technologies au service de ces fonctions (représentation des connaissances, raisonnement, apprentissage, planification, etc.).

Usages de l'intelligence artificielle en santé et sécurité : quel périmètre ?

L'ambition de cet exercice n'est pas de traiter de manière exhaustive l'ensemble des liens entre



© La lune majuscule pour l'INRS/2022

IA et SST. Il s'agit d'explorer les usages des systèmes d'IA spécifiquement dédiés à la prévention des risques professionnels. Ce périmètre n'est pas simple à délimiter. La première difficulté est de définir quelles innovations technologiques mobilisent ou non de l'IA. Sur ce point, comme évoqué précédemment, le groupe de travail a retenu une définition assez englobante. La seconde difficulté est de s'accorder sur les usages à considérer, car certains développements peuvent être motivés par des objectifs autres que la préservation de la santé au travail (amélioration de la qualité, réduction des consommations de matières premières ou d'énergie) mais présenter des conséquences favorables pour cette dernière. Le parti pris du groupe a été d'essayer de se concentrer sur des développements de systèmes d'IA, motivés en premier lieu par des objectifs d'amélioration de la SST. Il peut s'agir d'usages à des fins d'études et recherche sur les risques professionnels (par exemple en épidémiologie), de systèmes visant à supprimer des expositions de travailleurs à des dangers (par exemple en recourant à des robots autonomes), ou bien de dispositifs de détection de situations à risques avant la survenue d'un dommage (par exemple, par le traitement en temps réel de données collectées par des objets connectés). Certains usages « indirects » ne sont donc pas traités ici. C'est par

exemple le cas des dispositifs visant la substitution de l'humain par la machine à des fins avant tout productivistes, comme l'automatisation de tâches répétitives dans les services : les conséquences possibles (positives ou négatives) en SST sont importantes, mais relèvent du sujet plus global de l'impact des technologies sur le travail, qui a déjà été traité dans d'autres travaux de prospective conduits par l'INRS⁵. Les développements de dispositifs de réalité virtuelle à des fins de formation à la prévention ne sont pas abordés non plus, dans la mesure où il s'agit essentiellement d'appliquer au domaine de la SST des technologies qui ont été initialement développées à d'autres fins ; ce sujet a également été évoqué dans le cadre d'un précédent exercice⁶.

Tout au long de cette réflexion, les membres du groupe ont identifié des circonstances dans lesquelles l'introduction de l'IA est susceptible de modifier de façon significative les situations de travail et l'approche globale de la prévention des risques professionnels. Sans remettre en cause *a priori* l'utilisation de ces technologies, ces conséquences possibles imposent une vigilance et une réflexion particulières de la part des partenaires sociaux et, plus généralement, des organismes chargés de la prévention des risques professionnels. Certains des messages clés issus de l'exercice relèvent ces points de vigilance.



Méthodologie suivie

Un exercice en cinq phases

Cet exercice a été mené par un groupe de travail pluridisciplinaire composé d'experts⁷ et accompagné sur le plan méthodologique par le cabinet Futuribles. Il a été conduit en cinq phases.

→ Première phase : définition collective du sujet et de son périmètre

Du fait de la technicité du sujet traité, la définition des concepts a constitué une étape importante. Une note de cadrage a servi de base à la première réunion du groupe de travail visant à clarifier les concepts et délimiter le sujet. L'horizon temporel de l'exercice de prospective a également été défini durant cette phase.

→ Deuxième phase : analyse du sujet

L'analyse des facteurs clés a permis d'établir une liste de douze variables, réparties au sein de trois composantes : évolution de l'offre en IA, acceptabi-



© Gael Kerbaol/INRS/2022

Lors de la restitution de cet exercice de prospective, une mise en scène impliquant une fausse machine intelligente a été présentée.

lité des usages possibles de l'IA, travail et prévention. La rédaction de chacune des fiches variables a été confiée à un membre du groupe puis discutée en réunion.

→ Troisième phase : élaboration des scénarios

Sur la base des hypothèses d'évolution des différentes variables, quatre scénarios ont été établis. Ils ont été construits de manière contrastée, afin d'explorer un champ des possibles suffisamment large.

→ Quatrième phase : étude de cas d'usage

Trois ateliers ont ensuite été organisés, afin de poursuivre la réflexion sur les usages possibles de l'IA et sur ses conséquences potentielles en prévention aussi bien pour les acteurs de la SST que pour les conditions de travail des salariés.

Ces ateliers ont été conduits sur la base des différents contextes envisagés dans les scénarios, en essayant de considérer toutes les conséquences, tant favorables que défavorables, dans trois champs d'usages possibles : épidémiologie/ accidentologie, surveillance des lieux de travail et des travailleurs et robotique avancée.

→ Cinquième phase : élaboration des messages clés

Une dernière série d'échanges avec le groupe de travail et des experts extérieurs au groupe a permis de stabiliser une liste de vingt-deux messages clés, principaux enseignements issus de cette démarche, et qui sont présentés en fin d'article (Cf. Focus).

Scénarios prospectifs : l'IA au travail en 2035

Les quatre scénarios ont été construits par le groupe sur la base de l'instruction de douze variables clés. Ils donnent à voir des évolutions possibles et contrastées du contexte de développement des usages de l'IA en milieu de travail à l'horizon 2035. Ils constituent avant tout un outil de projection et d'aide à la réflexion.

Scénario 1 : Les géants du numérique imposent leurs solutions et leur vision

Dans ce scénario, l'exubérance technologique se poursuit dans un cadre de compétition pour la maîtrise de l'IA, mais les acteurs principaux en sont les géants du numérique, essentiellement les AMAMA⁸ en Occident et les BATX⁹ 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 incontournables. 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.

Scénario 2 : Les États garantissent un cadre pour l'intégration de l'IA

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 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 standards 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'expertise 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.

Scénario 3 : Développement démocratique

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, no code*). De plus, les recherches en IA lancées depuis les années 2010 aboutissent 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 instaurant le principe 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.

Scénario 4 : Déclin de l'IA

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, etc.) et pour les travailleurs (diminution de la pénibilité, sécurité, etc.).

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 IA dans le monde du travail. À partir de 2030, ce rejet mène à un déclin de cette technologie et de ses usages professionnels.

Cas d'usage

La deuxième phase de cet exercice a consisté à sélectionner trois types d'applications possibles de l'IA en matière de SST, et à les confronter aux futurs possibles décrits dans les scénarios.

Ce travail a été conduit sous la forme d'ateliers rassemblant des membres du groupe de travail et des experts en SST n'ayant pas participé à la phase initiale. La catégorisation des usages a été inspirée par des réflexions analogues conduites par des équipes américaine [9] et québécoise [10].

Usages possibles des outils d'IA en épidémiologie et accidentologie

De par son potentiel de traitement intelligent de données massives, l'IA ouvre des possibilités prometteuses pour les acteurs en SST, notamment pour des usages :

- en épidémiologie, l'IA peut offrir de nouvelles possibilités de traitements sophistiqués et rapides des données collectées pour des populations d'intérêt en matière de vulnérabilité, d'exposition, etc. Mais aussi, sans doute, une meilleure exploitation croisée de bases de données distinctes (indicateurs de santé, trajectoires professionnelles, données de mesures individuelles ou d'ambiance, etc.) ;
- en accidentologie, les systèmes de traitement automatique du langage ouvrent également des possibilités de meilleure exploitation de données mal structurées, données textuelles telles que certificats de décès, mais demain pourquoi pas déclarations d'AT, données provenant des services de santé au travail relatives aux aptitudes médicales, bases de données de retour d'expérience de type « Epicea¹⁰ », données intra-entreprise : registre de dysfonctionnements de machines, données AT-MP stockées *via* des logiciels spécifiques, enquêtes QSE, etc. L'usage de ces outils pour optimiser l'exploitation des données contenues dans la littérature scientifique est également à considérer, il dépasse le champ de l'accidentologie.

→ Potentiel de développement

Les principaux facteurs susceptibles d'accélérer le développement de ces types d'usage sont :

- la disponibilité de grandes masses de données de qualité ;
- la volonté des acteurs de partager et d'exploiter leurs données et la mise en place d'un cadre permettant de le faire dans de bonnes conditions (protection des données, compatibilité des bases...) ;
- l'augmentation des capacités de stockage et de traitement de ces données ;



- l'amélioration des performances des outils d'IA pour atteindre un niveau de fiabilité important. Les bénéfices attendus sont des progrès dans l'évaluation et l'analyse des risques, qui devraient permettre des avancées en prévention.

→ **Atouts et points de vigilance pour la SST**

Le développement de ces outils pourrait permettre de détecter de nouveaux facteurs de risques (ou combinaisons de facteurs), que ce soit pour les accidents ou pour les maladies, ou apporter des progrès dans le suivi de santé des travailleurs tout au long de leur carrière, et notamment de populations dont le suivi pose des difficultés aux acteurs institutionnels, comme les travailleurs saisonniers ou les intérimaires.

Les principales interrogations portent sur les données nécessaires au fonctionnement de ces systèmes. La protection de ces données est un enjeu important : il s'agit de garantir la sécurisation des données personnelles, notamment de santé, d'être transparent sur les objectifs de ces traitements, de prévenir les détournements d'usage et les biais. De plus, la qualité des jeux de données, notamment ceux qui vont servir à entraîner les systèmes d'IA, est déterminante.

Ces systèmes peuvent à terme faire évoluer les approches de la santé au travail. Ils peuvent notamment inciter à une individualisation accrue du suivi de travailleurs en intégrant des facteurs de risque individuels (génétique, hygiène de vie, etc.).

→ **Quelles implications pour les acteurs de la SST ?**

Les acteurs de la veille sanitaire et de la prévention devront développer leurs compétences pour pouvoir conduire des études sur ces systèmes, et former, conseiller et assister les acteurs concernés, notamment les administrations, les préventeurs et les partenaires sociaux.

Il leur faudra aussi arriver à gérer quelques écueils possibles, comme le risque de se focaliser sur les outils eux-mêmes, au détriment d'autres missions et approches ou de limiter leur attention aux domaines ou aux populations pour lesquels ils disposent de données au détriment d'autres, rendues « invisibles » faute de données exploitables (cela pourrait être le cas pour des facteurs organisationnels, par exemple).

Technologies de sécurisation des environnements de travail recourant à l'IA

Ces solutions visent à sécuriser les environnements de travail en recourant à des dispositifs permettant d'analyser en temps réel des informations provenant d'objets connectés. Ces dispositifs fonctionnent grâce à des capteurs capables de mesurer différents types de valeurs, et à des actionneurs. On distinguera ici deux types d'usage de ces systèmes de surveillance :

- les systèmes visant à surveiller l'environnement de travail et capables d'actionner une alerte avant la survenue d'un phénomène dangereux : émission d'un produit toxique, proximité d'un équipement en mouvement... Ces systèmes sont définis ici comme des solutions de détection ;
- les systèmes visant à surveiller le travailleur lui-même. Il pourra s'agir d'un équipement de protection individuelle (EPI) connecté, effectuant des mesures régulières de données biométriques, ou encore d'un équipement de travail muni de capteurs biométriques.

→ **Potentiel de développement**

Actuellement, les avancées sont liées à l'évolution de la puissance de calcul qui permet une augmentation de la performance et de la précision. Par ailleurs, la consommation d'énergie des puces embarquées est de plus en plus faible. Les développements technologiques à venir pourraient encore augmenter cette capacité de calcul et ainsi augmenter la pertinence, perfectionner la miniaturisation et réduire la consommation d'énergie (sur le véhicule).

Tout cela pourrait finalement réduire les prix de solutions fondées sur l'IA, qui limitent actuellement fortement leur déploiement. La 5G, en réduisant les temps de latence, pourrait aussi permettre d'assurer un niveau de sécurité élevé.

→ **Atouts et points de vigilance pour la SST**

Ces technologies permettent de sécuriser les environnements de travail, y compris pour des salariés peu formés ou peu sensibilisés aux règles de sécurité du site (comme les intérimaires, les nouveaux embauchés ou les intervenants extérieurs). Par ailleurs, les informations qu'elles peuvent apporter sur les situations et comportements à risque peuvent aider au ciblage des actions de prévention en fonction de l'occurrence des presque-accidents. Cependant, la facilité d'usage et d'implémentation de ces solutions risque d'encourager les acteurs à ne prendre en considération que les risques identifiés par le système d'IA, sans évaluer régulièrement les risques plus organisationnels. La méconnaissance des possibles défaillances de ces outils pourrait entraîner une baisse de la vigilance des salariés comme des employeurs. Ces outils pourraient aussi favoriser le développement d'approches « court-termistes » et plus coercitives que préventives, avec pour conséquences possibles de reporter la responsabilité sur le salarié en cas d'accident, ou bien de faciliter une surveillance intrusive des salariés, à la fois sur leur état de santé par la collecte de données (rythme cardiaque, suivi médical sans accord formalisé), mais également sur leurs comportements et rythmes de travail. Une telle surveillance peut induire une intensification du travail chez le

salarié qui se sait surveillé et conduire à des risques psychosociaux ou d'accidents accrus.

→ Quelles implications pour les acteurs de la SST ?

Le regroupement et la mutualisation de données de différentes entreprises d'un même secteur pour constituer une base sous l'égide d'organismes de recherche pourraient permettre de faire avancer à grands pas la recherche dans le domaine de la prévention en SST.

Parallèlement, face aux conséquences d'usages potentiellement délétères de ces technologies émergentes, un effort d'information et de formation en direction des préventeurs et des entreprises est souhaitable pour faire connaître les points de vigilance à garder à l'esprit, ainsi que les bonnes pratiques à mettre en œuvre lors de leur implémentation.

Enfin, les préventeurs auront pour responsabilité d'alerter et de rappeler aux entreprises que ces outils ne s'apparentent pas à des systèmes de sécurité et qu'ils ne les exonèrent pas de la démarche d'évaluation des risques et de prévention primaire à mettre en œuvre dans leur entreprise.

Robotique avancée recourant à l'IA

Certaines technologies de robotique avancée (embarquant de l'IA) offrent des solutions potentiellement bénéfiques à la SST. Deux types de cas d'usage ont été discutés : la téléopération, qui permet d'éloigner l'opérateur de situations délétères ou dangereuses, et la collaboration homme – robot, qui permet de faire exécuter les tâches physiquement pénibles de l'opérateur par un robot apte à la collaboration.

- La téléopération : Les compacteurs ou les brise-roches téléguidés du BTP permettent d'éloigner l'opérateur de ces machines dangereuses. Mais pour les tâches plus complexes ou à réaliser dans un espace dangereux, le téléguidage n'est pas toujours possible. Dans ces situations, la machine doit embarquer une partie de l'expertise et des capacités de l'opérateur humain (voire des capacités supérieures à celles de ce dernier). Il s'agit notamment de la vue stéréoscopique, du déplacement dans un environnement fait pour des humains, ou encore de la préhension d'objets... Dans ces cas, l'opérateur distant ne pilotera que des « ordres macro » de type, « Rends-toi à tel emplacement », « Ouvre cette porte », « Effectue la mission XDK74 », etc. On parle alors de téléopération, celle-ci nécessite des fonctions de robotique avancée faisant appel à l'IA.
- La collaboration : La robotique collaborative définit les conditions grâce auxquelles on peut faire travailler sans risque corporel un opérateur à proximité immédiate d'un robot. On distingue trois types de collaboration : la collaboration

directe où l'opérateur et le robot travaillent simultanément sur une même pièce, la collaboration indirecte où l'opérateur et le robot travaillent alternativement sur la même pièce et le partage d'espace de travail où opérateur et robot travaillent indépendamment dans un espace commun. Pour être sûres, ces collaborations imposent que le robot soit doté de facultés de perception des actions de son collaborateur humain très pointues. Ces facultés peuvent faire appel à de l'IA.

→ Potentiel de développement

Ces deux types d'usage de robotique avancée sont destinés à se déployer si les progrès suivants sont réalisés :

- démocratisation de la technologie (coûts d'achat, d'intégration, d'exploitation et de maintenance) ;
- extension des domaines de fonctionnement opérationnel ;
- progression de l'acceptabilité et de la confiance par les collaborateurs humains ;
- progression de la qualité de l'intégration de ces technologies dans les organisations de travail (transparence de l'objectif visé par cette intégration [augmentation de la productivité ou réduction de la pénibilité], préparation et concertation des équipes, cadence au poste de travail, accompagnement du changement, formation).

Pour ces deux domaines d'application et à moyen terme, les potentiels de développement se concentrent autour de deux enjeux :

- augmentation de la flexibilité de ces équipements (variabilité des produits et des tâches) ;
- augmentation de la capacité de ces équipements à réaliser des tâches plus complexes.

→ Atouts et points de vigilance pour la SST

Le principal atout de ces technologies réside dans la réduction, voire la suppression des risques auxquels les opérateurs sont soumis, grâce à leur remplacement par le robot sur les sites confinés ou pollués, ou pour la réalisation de tâches répétitives, ou encore le port de charges lourdes.

Elles pourraient permettre le maintien ou le retour à l'emploi d'opérateurs déchargés de la partie physique du poste, ainsi que la requalification des opérateurs, dans les domaines liés à la maintenance et la supervision de ces nouveaux outils.

Cependant, ces technologies peuvent également faire émerger certains risques, en favorisant une augmentation délétère des cadences, du fait du maintien des objectifs de production auxquels s'ajouteraient les activités de maintenance et de supervision. Le remplacement du salarié par le robot peut aussi entraîner une perte du geste professionnel, conduisant à une progressive déqualification de la main-d'œuvre.



POUR EN SAVOIR +

- Pour visionner le replay de la restitution de l'exercice organisée le 18 novembre 2022 : <https://www.inrs.fr/footer/actes-evenements/prospective-IA.html>.
- Pour consulter le rapport publié par l'INRS : <https://intelligenceartificielle2022.inrs.fr>.
- L'ensemble des travaux de prospective est accessible sur : <https://www.inrs.fr/inrs/prospective-quel-travail-demain>

Par ailleurs, des comportements inattendus de la machine, provoqués par une situation imprévue ou une cyberattaque, pourraient conduire à des réflexes de rattrapage ou à des interventions en situation dégradée et être la cause d'accidents du travail. La gravité de ces situations pourrait être accentuée par la perte de savoir-faire induite par l'automatisation de nombreuses tâches.

→ **Quelles implications pour les acteurs de la SST ?**

Les acteurs de la prévention pâtissent d'un manque de compétences en matière d'IA et de robotique intelligente, ainsi que de terrains d'expérimentation, qui risque de leur poser des difficultés pour accompagner l'émergence de ces technologies.

Des efforts de formation – des préventeurs, des concepteurs mais également des salariés – sont nécessaires pour permettre de débattre de manière critique de ces sujets, condition nécessaire à l'instauration de la confiance et de l'acceptabilité. ●

1. Voir : Article L. 4121-2 du Code du travail. Accessible sur : https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000033019913/.
2. Composition du groupe de travail : Anani Olympio (CNP Assurances) ; Bertrand Braunschweig (« confiance.ai ») ; François de Jovenel, Jean-François Soupizet (Futuribles) ; Jorge Munoz (université de Brest) ; Martin Bieri, Régis Chatellier (Cnil/Linc ; Nazim Fatès (Inria/Loria) ; Nicolas Bouby (Bouygues construction) ; Sylvain Halluin (Cramif) ; Timothée Silvestre (CEA) ; Vincent Mandinaud (Anact) ; Jennifer Clerté, Michaël Sarrey, Michel Héry, Marc Malenfer (INRS).
3. L'intension (anglicisme) est un terme utilisé notamment en logique, et qui désigne « l'ensemble des caractères qui constituent un concept » (par opposition à extension).
4. Voir : https://academie-technologies-prod.s3.amazonaws.com/2018/04/06/13/49/30/183/Rapport_IA_DEF.pdf.
5. Voir : « Modes et méthodes de production en France en 2040 : quelles conséquences pour la santé et la sécurité au travail ? ». Accessible sur : <https://www.inrs.fr/media.html?refINRS=VEP%203>.
6. Voir : « Quelle formation à la santé et sécurité au travail en 2030 ? ». Accessible sur : <https://www.inrs.fr/media.html?refINRS=PV%2015>.
7. Les auteurs remercient l'ensemble des contributeurs et notamment les membres du groupe de travail (Cf. composition du groupe en note 2).
8. Ce sigle désigne les firmes : Alphabet, Meta, Amazon, Microsoft, Apple.
9. Ce sigle désigne les firmes : Baidu, Alibaba, Tencent, Xiaomi.
10. La base de données « Epicea » (Études de prévention par l'informatisation des comptes rendus d'accidents) est gérée par l'INRS et rassemble des récits d'accidents graves, mortels ou particulièrement significatifs pour la prévention. Pour plus d'informations, voir : <https://www.inrs.fr/publications/bdd/epicea.html>.
11. Consulter : <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A52021PC0206>.

BIBLIOGRAPHIE

[1] BRYNJOLFSSON E., ROCK D., SYVERSON C. – *Artificial intelligence and the modern productivity paradox: a clash of expectations and statistics*. In: AGRAWAL A., GANS J., GOLDFARB A. (dir.) – *The economics of artificial intelligence: an agenda*. University of Chicago Press, 2017. Chapitre accessible sur : <https://www.nber.org/system/files/chapters/c14007/c14007.pdf>.

[2] TOORAJIPOUR R., SOHRABPOUR V., NAZARPOUR A. ET AL. – Artificial intelligence in supply chain management: A systematic literature review. *Journal of business research*, 2021, 122, pp. 502-517. Accessible sur : <https://doi.org/10.1016/j.jbusres.2020.09.009>.

[3] COSIMATO S., VONA R. – Digital innovation for the sustainability of reshoring strategies: a literature review. *Sustainability*, 2021, 13 (14), p. 7601. Accessible sur : <https://doi.org/10.3390/su13147601>.

[4] ROMERO D., MATTSSON S., FAST-BERGLUND Å. ET AL. – *Digitalizing occupational health, safety and productivity for the operator 4.0*. In: MOON I., LEE G., PARK J. ET AL. (dir.) – *Advances in production management systems. Smart manufacturing for Industry 4.0*. Springer, 2018. Chapitre accessible sur : https://doi.org/10.1007/978-3-319-99707-0_59.

[5] KOUTROUMPINAS P., ZHANG Y., WALLIS S. ET AL. – An artificial intelligence empowered cyber physical ecosystem for energy efficiency and occupation health and safety. *Energies*, 2021, 14, p. 4214. Accessible sur : <https://doi.org/10.3390/en14144214>.

[6] BRUNET S., HÉRY M., LEVERT C. ET AL. – Quels risques professionnels et quelle prévention en 2040 ? *Hygiène & sécurité du travail*, 2016, 245, pp. 105-111. Accessible sur : <https://www.inrs.fr/publications/hst/veille-et-prospective.html>.

[7] HÉRY M., MALENFER M., DEVEL S. ET AL. – Evolution of working conditions under the impact of ICTs. *Journal of safety research*, 2021, 77, pp. 268-276. Accessible sur : <https://doi.org/10.1016/j.jsr.2021.03.009>.

[8] ATAIN-KOUADIO J.J., AUBLET-CUVELIER A., BARBET-DETRAYE R. ET AL. – Quelle place pour les robots d'assistance physique en 2030 ? *Hygiène & sécurité du travail*, 2014, 235, pp. 64-68. Accessible sur : <https://www.inrs.fr/media.html?refINRS=CC%205>.

[9] PISHGAR M., ISSA S.F., SIETSEMA M. ET AL. – REDECA: A novel framework to review artificial intelligence and its applications in occupational safety and health. *International journal of environmental research and public health*, 2021, 18, p. 6705.

[10] COMEAU M. – *Intelligence artificielle en santé et SST*. IRSST, 2021. Rapport accessible sur : <https://www.irsst.qc.ca/media/documents/PublIRSST/QR-1130-fr.pdf?v=2022-09-21>.

FOCUS SUR... Principaux enseignements issus de l'exercice

Les enseignements issus de cet exercice de prospective sont rassemblés sous forme de vingt-deux messages clés, organisés ici en quatre thématiques.

Un marché en plein développement

- ❶ Les différentes avancées en cours en matière d'intelligence artificielle (IA), permises par des investissements massifs de la part d'acteurs privés et publics, préfigurent le développement d'un marché conséquent dans les prochaines années. La sécurisation des environnements de travail est un des domaines d'usages professionnels de ces innovations.
- ❷ Plus l'automatisation sera intelligente, plus la machine (ou l'algorithme) remplira les tâches effectuées jusqu'alors par des travailleurs. Cette automatisation peut permettre de soustraire certains travailleurs à des risques. Elle conduira aussi à une modification des tâches d'autres travailleurs vers des fonctions d'entraînement, d'accompagnement et de contrôle.
- ❸ Il existe un enjeu pour les acteurs à favoriser le développement de systèmes d'IA qui soient compatibles avec les valeurs essentielles de l'approche européenne et française de la santé et sécurité au travail (SST) : approche collective, protection des données, dialogue social. La position hégémonique actuelle des géants du numérique américains et chinois pose donc question.
- ❹ Face aux problématiques éthiques que soulèvent ces nouvelles technologies, de nombreux organismes ont formulé des principes à respecter, afin que le développement de l'IA puisse se faire au bénéfice de tous les acteurs de la société. Il est nécessaire d'identifier et de promouvoir auprès des entreprises les recommandations pertinentes pour un usage éthique des technologies d'IA en SST.
- ❺ Le scénario d'un nouvel « hiver de l'IA », du fait de blocages (« murs ») technologiques, de crises énergétiques, de rejets sociétaux ou de cybercriminalité, etc., n'est pas exclu. Il est donc important de ne pas faire reposer toutes les avancées en SST sur ces solutions technologiques. La recherche de solutions de prévention efficaces ne recourant pas à ces systèmes ne doit pas être abandonnée.
- ❻ L'usage des systèmes d'IA en SST peut se heurter au « mur de l'explicabilité » (notamment pour les dispositifs faisant appel à l'apprentissage profond) et donc de la compréhension des messages et décisions générés par l'IA. Le développement des usages de l'IA à des fins de prévention des risques professionnels supposera à la fois de favoriser la bonne compréhension de ces outils (vertus et limites) par les salariés et employeurs, et de favoriser l'émergence de solu-

tions transparentes (par exemple, des systèmes d'IA hybrides : puissance de l'IA et transparence des systèmes de raisonnement logique) facilitant les débats entre parties prenantes.

Des promesses en santé et sécurité au travail

- ❼ Les avancées de l'IA sont potentiellement porteuses d'usages divers en SST. Des avancées sont à attendre dans le domaine des traitements de grandes masses de données au service de l'accidentologie et de l'épidémiologie, dans la toxicologie *in silico*, dans celui de la sécurisation des environnements de travail, dans le développement de technologies de robotique avancée.
- ❽ Les progrès de l'IA mettant en œuvre des techniques d'apprentissage ouvrent des perspectives intéressantes, par exemple en épidémiologie et en accidentologie, sous réserve de disposer de masses de données fiables et de ne pas occulter certaines dimensions de la SST pour lesquelles on ne dispose pas forcément de données exploitables (avec une dimension organisationnelle notamment).
- ❾ L'IA ouvre aussi des possibilités de supervision d'un environnement de travail, un chantier ou un site industriel par exemple. Au-delà de la détection et de l'alerte, on doit attendre de ces dispositifs qu'ils fournissent des informations utiles à l'élaboration de mesures de prévention durables (organisationnelles) ; cela implique une exploitation par des personnes en capacité de les analyser.
- ❿ Certaines technologies de robotique avancée (embarquant de l'IA) offrent des solutions potentiellement bénéfiques à la SST. C'est le cas notamment des dispositifs de téléopération et de robotique collaborative, qui peuvent atténuer voire supprimer des expositions à des facteurs de risque. L'implémentation de ces dispositifs doit cependant faire l'objet d'une évaluation systématique, permettant de garantir qu'ils ne génèrent pas de nouveaux risques (intensification du travail, perte de sens...).

Limites et points de vigilance sur les usages de l'IA en SST

- ⓫ La logique de rentabilisation des investissements dans ces technologies, parfois coûteuses, peut amener à positionner ces systèmes au centre de l'organisation du travail, au risque de placer le travail humain au second plan.
- ⓬ De façon générale, un usage inadapté, détourné, ou l'absence d'une réflexion préalable sur l'organisation de l'intégration de ces nouvelles technologies



pourraient conduire à des effets délétères en matière de SST. La facilité apparente d'usage et d'implémentation de ces solutions « intelligentes » risque d'entraîner un effet de contentement des acteurs, les incitant à ne prendre en considération que les risques identifiés par le système d'IA, sans évaluer régulièrement les risques plus organisationnels ne faisant pas l'objet d'un monitoring technologique.

13 L'usage de l'IA en SST peut conduire à développer des outils de surveillance des travailleurs et d'alerte, lorsque les conditions d'un travail en sécurité ne sont pas remplies (consignes non respectées, état de santé du travailleur hors norme, etc.). Cette surveillance permanente peut générer des risques psychosociaux (RPS) et également conduire à une individualisation de la SST et à une responsabilisation exclusive du travailleur, au détriment de la mise en place par l'employeur de mesures de prévention collectives.

14 Une attention devra être portée aux risques éventuels associés à des usages, dans des dispositifs de SST, d'algorithmes d'IA qui n'ont pas été développés spécifiquement à cette fin (bibliothèques *open source*, produits standards vendus sur étagère).

15 Les technologies d'apprentissage profond (*deep learning*) reposent sur l'entraînement d'un modèle à partir d'un jeu de données. L'utilisation de l'IA dans les entreprises va donc impliquer la collecte et le stockage de nombreuses données. Dès lors qu'un usage en prévention est envisagé, se pose la question de la constitution, de la qualification et de l'étiquetage du jeu de données utilisées lors des phases d'apprentissage dans les systèmes supervisés. Une attention particulière doit être apportée aux jeux de données, afin qu'ils correspondent aux domaines d'applicabilité qui peuvent varier selon les activités et situations de travail et qu'ils ne soient pas biaisés (Cf. aussi Point n° 19).

16 Les accidents du travail surviennent fréquemment lors de situations atypiques par rapport au déroulement classique d'un processus de production : situations dégradées, pannes, opérations de maintenance... Ces situations sont souvent imprévues, donc non anticipées dans les procédures, ce qui les rend particulièrement dangereuses. Elles constituent ainsi une limite possible à l'entraînement des systèmes d'IA, les jeux de données nécessaires n'étant pas en mesure d'intégrer de manière exhaustive la palette des aléas qui peuvent survenir dans de nombreux contextes de travail (chantiers, grands sites industriels, travaux sur la voie publique...).

Pistes d'actions

17 Du fait des opportunités offertes par ces nouvelles technologies, ainsi que des risques potentiels qu'elles sous-tendent, la formation des acteurs de la prévention (employeurs, représentants du personnel, préventeurs) constitue un enjeu primordial de l'inté-

gration à venir de l'IA aux équipements de travail et aux solutions de prévention. Ces formations devront permettre une bonne compréhension du mode de fonctionnement de ces outils, des enjeux éthiques, du cadre réglementaire qui les régit, des possibilités de les piloter, des risques qu'ils peuvent représenter, mais aussi l'acquisition de méthodes permettant la définition des besoins, la rédaction de cahiers des charges et l'intégration des dispositifs dans l'entreprise. L'acculturation des acteurs du dialogue social, aussi bien au niveau des branches professionnelles que des entreprises (notamment dans les TPE-PME), est indispensable, afin qu'ils soient en mesure d'appréhender et de discuter en amont les transformations des modes et méthodes de travail que ces nouveaux systèmes induisent.

18 Le développement et la mise sur le marché de dispositifs utilisant des techniques d'IA présentés comme des outils de prévention doivent être assurés par des personnes disposant de solides compétences en matière de SST. Au-delà des formations à prévoir dans le cadre de l'entreprise, des modules doivent également être mis en œuvre dans les cursus d'écoles de management et d'ingénieurs, afin de sensibiliser les futurs commanditaires et développeurs de systèmes d'IA aux opportunités et risques qu'amènent avec elles ces nouvelles technologies en matière de SST.

19 Il convient de promouvoir auprès des entreprises les démarches reposant sur l'expérimentation et l'évaluation, qui permettent de mesurer en conditions réelles les conséquences des nouveaux systèmes sur l'organisation de l'entreprise et sur le travail des opérateurs, et de conserver une possibilité de revenir en arrière.

20 Les normes et réglementations encadrant l'IA se développent (*AI act 11*). Il est donc essentiel que les principes de SST soient portés au sein des instances d'élaboration. C'est notamment vrai au niveau des réglementations européennes, mais également dans les comités de normalisation.

21 Une réflexion collective (de type conférence de consensus) devra être menée sur la question des données utilisées dans des dispositifs d'IA touchant à la SST. Il s'agira notamment de définir des règles pour la constitution des jeux de données, l'encadrement de leur utilisation en fonction des domaines d'application. Au-delà des utilisateurs, il s'agira d'impliquer les partenaires sociaux et des personnalités qualifiées (experts, philosophes spécialistes de l'éthique, juristes...).

22 De façon générale, les avancées de l'IA offrent des perspectives de progrès pour la prévention des risques professionnels. Elles entraînent aussi, comme toute modification, certains risques. Le développement et la diffusion d'outils méthodologiques, permettant de guider les acteurs face à ces innovations, constituent un enjeu fort pour les organismes de prévention. ●