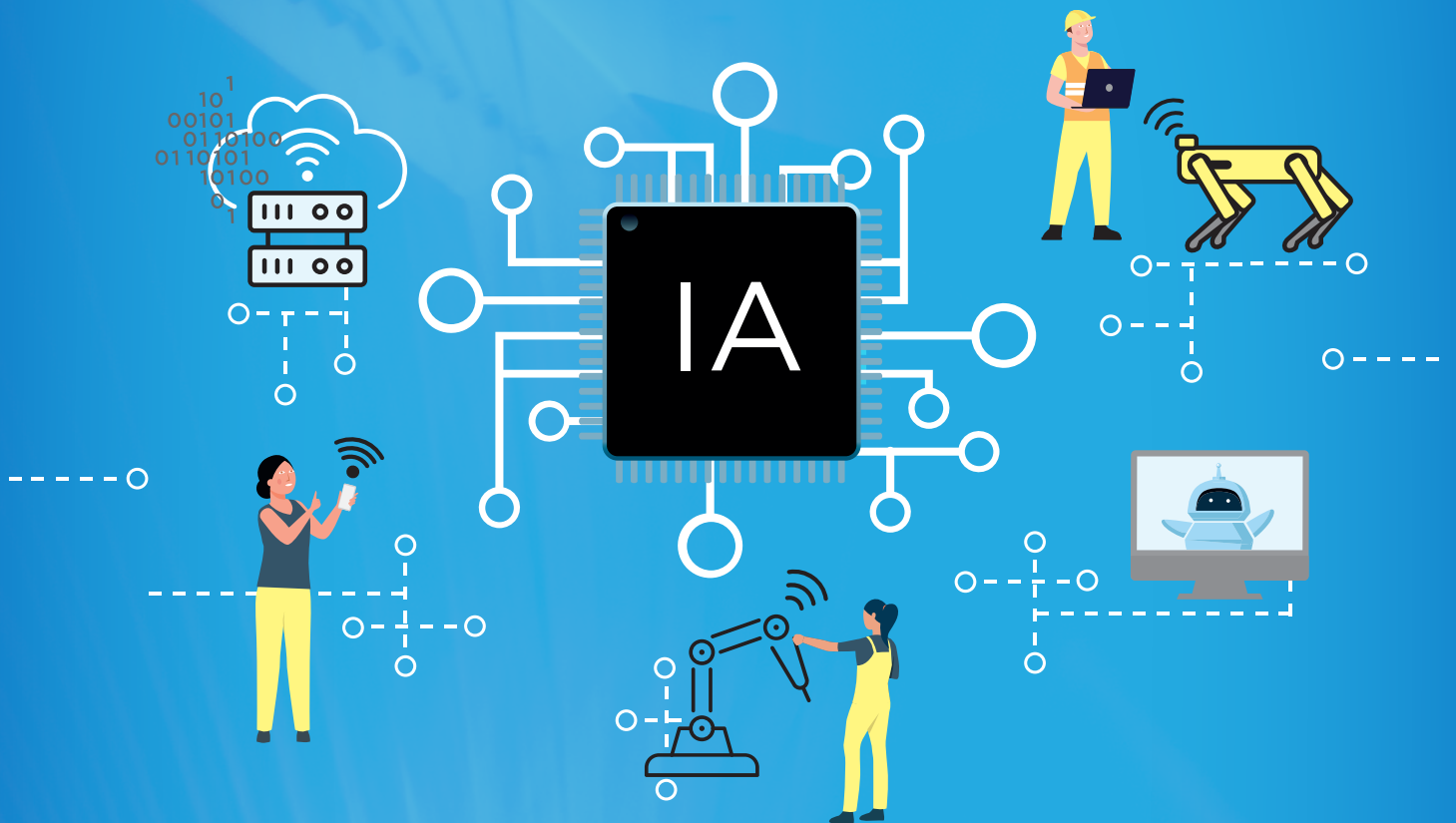




**l'Assurance
Maladie**
RISQUES PROFESSIONNELS

inrs
Institut National de Recherche et de Sécurité



L'intelligence artificielle au service de la santé et sécurité au travail

Enjeux et perspectives
à l'horizon 2035

Groupe projet

Anani Olympio (CNP Assurances) ; Bertrand Braunschweig (Bilab, 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, Marc Malenfer, Michaël Sarrey, Michel Héry (INRS).

Participants aux ateliers

Guillaume Boulanger (Santé publique France) ; Sabri Bayouhd (Blaxtair) ; Cegarra Julien [INU Champollion (Laboratoire SCoTE)] ; Joséphine Hurstel (CNNum) ; Sebastien Lissare (Cognitive Engine) ; Claire Tissot, Frédéric Clerc, Guy Hedelin, Lisa Chedik, Adriana Savescu, Lien Wioland, Louis Laurent, Patrice Marchal, Adel Sghaier, David Tihay (INRS).

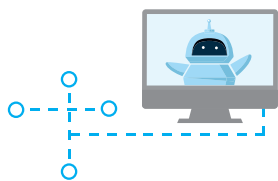
Autres contributeurs

Loïck Briot (Mines Nancy) ; Yann Ferguson (Icam) ; Ariane Leroyer (Université de Lille) ; Agnès Aublet-Cuvelier, Benjamin Paty, Jean Theurel, Jean-Christophe Blaise, Cordelia Genzel, Nadiège Félicie, Virginie Govaere (INRS).

Les auteurs remercient également Céline Raynal de l'entreprise Safety Data et Florian Ruen de l'entreprise Eyes'R.

➡ Ce document est une synthèse. Un rapport complet est disponible sur le site de l'INRS : www.inrs.fr/prospective





Introduction

Parmi les 9 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 technique »¹. 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 des risques rigoureuse nécessite bien souvent 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 ces dernières années. Que ce soit à des fins d'amélioration de la productivité, dans l'optique de diminuer la vulnérabilité des chaînes d'approvisionnement ou pour des raisons géopolitiques de relocalisation d'activités dans des secteurs stratégiques, 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.

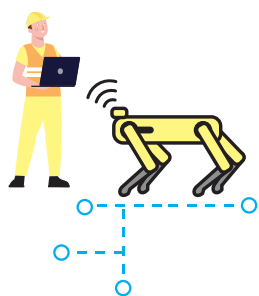
Les acteurs de la santé et sécurité au travail (S&ST) s'intéressent naturellement aux conséquences que peut avoir l'introduction de systèmes d'IA en milieu professionnel. Il s'agit généralement pour eux d'essayer d'évaluer les risques que peuvent faire peser ces dispositifs sur la santé physique et mentale des travailleurs et d'émettre des préconisations de prévention. L'INRS prend logiquement sa part dans ce travail qui relève pleinement de ses missions.

Le parti pris de l'exercice de prospective, dont les résultats sont présentés dans ce document, est de se focaliser sur les usages possibles des dispositifs d'IA à des fins de protection de la santé et de la sécurité au travail à un horizon d'une douzaine d'années. En pratique, on verra que certaines conclusions vont au-delà de ce cadre initial mais il a été jugé pertinent de les conserver. Ces technologies sont d'ores et déjà utilisées sur certains lieux de travail, à un niveau encore marginal, mais qui soulève déjà des questions.

La démarche suivie reprend les éléments suivants :

- les fondamentaux de la pratique de la prospective à l'INRS, à savoir un travail collaboratif et pluridisciplinaire, impliquant de nombreux experts internes et externes à l'institut ;
- l'application d'un cadre méthodologique (voir annexe) permettant de donner à voir différents futurs possibles, tous cohérents, 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 siégeant au conseil d'administration de l'INRS.

1. Article L4121-2 du Code du travail : https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000033019913/



Intelligence artificielle : éléments de définition

De quoi parle-t-on ?

Il existe de très nombreuses définitions de l'intelligence artificielle, aucune ne réglant totalement la question car la discipline a des contours flous. Pour les besoins de cet exercice, le groupe de travail a retenu trois définitions qui sont compatibles entre elles.

Deux sont des définitions par intension. Celle proposée par la Commission d'enrichissement de la langue française et publiée au *Journal Officiel* du 9 décembre 2018 :

« 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 2018² :

« – perception, un domaine très vaste contenant la reconnaissance des images, des formes, et des sons (et de toutes 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'Intelligence artificielle, rendue célèbre dans les années quatre-vingt par les « systèmes experts ».

2. https://academie-technologies-prod.s3.amazonaws.com/2018/04/06/13/49/30/183/Rapport_IA_DEF.pdf

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 en santé et sécurité, quel périmètre ?

Comme indiqué en introduction, l'ambition de cet exercice n'est pas de traiter de manière exhaustive l'ensemble des liens entre IA et S&ST. 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 définir.

La première difficulté est de définir quelles innovations technologiques mobilisent ou non de l'IA. Sur ce point, comme évoqué ci-dessus, le groupe de travail a retenu une définition assez englobante.

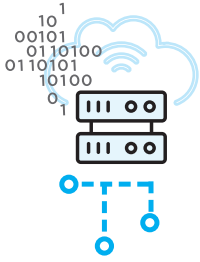
La seconde est de s'accorder sur les usages à considérer, car certains développements peuvent être motivés par des objectifs autres que la préservation de la santé au travail (amélioration de la qualité, réduction des consommations de matières premières ou d'énergie) mais avoir des conséquences favorables 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 santé et de la sécurité au travail. Il peut s'agir d'usages amont, à des fins d'étude et recherche sur les risques professionnels, par exemple en accidentologie, ou bien de systèmes visant à supprimer des expositions de travailleurs à des dangers, par exemple en recourant à des robots autonomes, ou 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 principalement 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 S&ST 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'INRS³. 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 S&ST des technologies qui ont été initialement développées à d'autres fins ; ce sujet a par ailleurs également été évoqué dans le cadre d'un précédent exercice de prospective⁴.

Tout au long des travaux, 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.

3. « Modes et méthodes de production en France en 2040 : quelles conséquences pour la santé et la sécurité au travail ? », <https://www.inrs.fr/media.html?refINRS=VEP%203>

4. « Quelle formation à la santé et sécurité au travail en 2030 ? », <https://www.inrs.fr/media.html?refINRS=PV%2015>



Intelligence artificielle : messages clés

Les principaux enseignements issus de cet exercice de prospective sont rassemblés ici sous forme de vingt-deux messages clés.

Un marché en plein développement

- 1** Les différentes avancées en cours en matière d'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.
- 2** 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 entraînera aussi une modification des tâches d'autres travailleurs vers des fonctions d'entraînement, d'accompagnement et de contrôle.
- 3** Il y a un enjeu pour les acteurs à favoriser le développement des systèmes d'IA qui soient compatibles avec les valeurs essentielles de l'approche européenne et française de la S&ST (approche collective, protection des données, dialogue social). La position hégémonique actuelle des géants du numérique américains et chinois pose donc question.
- 4** Face aux problématiques éthiques que posent 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 S&ST.
- 5** 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 S&ST sur ces solutions technologiques. La recherche de solutions de prévention efficaces ne recourant pas à ces systèmes ne doit donc pas être abandonnée.
- 6** L'usage des systèmes d'IA en S&ST 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 solutions 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

- 7** Les avancées de l'IA sont potentiellement porteuses d'usages divers en S&ST. 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.
- 8** 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 S&ST pour lesquelles on ne dispose pas forcément de données exploitables (dimension organisationnelle notamment).
- 9** 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.
- 10** Certaines technologies de robotique avancée (embarquant de l'IA) offrent des solutions potentiellement bénéfiques à la S&ST. C'est le cas notamment des dispositifs de téléopération et de robotique collaborative qui peuvent atténuer, voire supprimer des expositions à des facteurs de risques. 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 S&ST

- 11** 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.
- 12** De façon générale, un usage inadapté, détourné ou l'absence d'une réflexion préalable sur l'organisation de l'intégration de ces nouvelles technologies, pourrait conduire à des effets délétères en matière de S&ST. 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 S&ST 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 S&ST 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 S&ST d'algorithmes d'IA qui n'ont pas été développés spécifiquement à cette fin (librairies open source, produits standards vendus sur étagères).
- 15** Les technologies d'apprentissage profond (*deep learning*) reposent sur l'entraînement d'un modèle sur 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 (voir aussi message clé 21).
- 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'intégration 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'elles 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 santé et sécurité au travail. Au-delà des formations à prévoir dans le cadre de l'entreprise, des modules de formation 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 S&ST.
- 19** Il convient de promouvoir auprès des entreprises les démarches reposant sur l'expérimentation, 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). Il est donc essentiel que les principes de la S&ST 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 S&ST. 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, 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.



L'IA au travail en 2035 : scénarios prospectifs et cas d'usage

Les scénarios

Les quatre scénarios, dont les résumés sont présentés ici, donnent à voir des évolutions possibles et contrastées du contexte de développement des usages de l'intelligence artificielle en milieu de travail à l'horizon 2035. Ils constituent avant tout un outil de projection et d'aide à la réflexion.

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'intelligence artificielle, mais les acteurs principaux en sont les géants du numérique, essentiellement les AMAMA⁵ à l'Ouest et les BATX⁶ 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.

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 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.

5. Alphabet, Meta, Amazon, Microsoft, Apple

6. Baidu, Alibaba, Tencent, Xiaomi

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, 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.

4 – Déclin de l'IA

En début de période, les usages des systèmes d'intelligence artificielle se développent dans tous les domaines professionnels. Poussé 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 (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 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.



Cas d'usage

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

1. 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 S&ST. On pense notamment ici à certains domaines d'usage :

- En épidémiologie : l'IA peut offrir de nouvelles possibilités de traitements sophistiqués et rapides des données collectées pour des populations d'intérêt en termes de vulnérabilité, d'exposition, etc. Mais aussi sans doute d'une meilleure exploitation croisée de bases de données différentes (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 [exemple : certificats de décès, mais demain pourquoi pas déclarations d'accident du travail (AT), données provenant des services de prévention et de santé au travail relatives aux aptitudes médicales, bases de données de retour d'expérience type Épicea, données intra-entreprise : registre de dysfonctionnements de machines, données AT-MP (accident du travail-maladie professionnelle) stockées via des logiciels spécifiques, enquêtes QSE (qualité, sécurité, environnement), etc.].

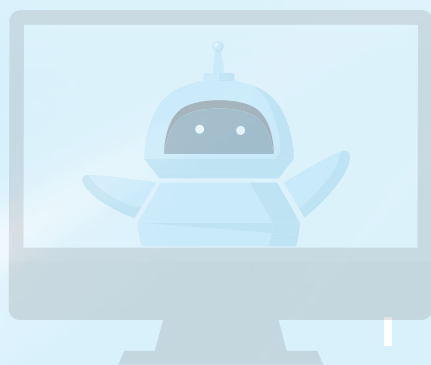
L'usage de ces outils pour optimiser l'exploitation des données contenues dans la littérature scientifique est également à considérer, il dépasse le champ de l'accidentologie.

■ Potentiel de développement

Les principaux facteurs susceptibles d'accélérer le développement de ces types d'usages 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

Le développement de ces outils ouvre des perspectives intéressantes en matière de prévention des risques et de suivi médical des travailleurs. Ils pourraient notamment permettre de détecter de nouveaux facteurs (ou combinaisons de facteurs) de risques que ce soit pour les accidents ou pour les maladies. Il est également possible d'imaginer 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 des données de santé, d'être transparent sur les objectifs de ces traitements, de prévenir les détournements d'usage. De plus, la qualité des jeux de données, notamment ceux qui vont servir à entraîner les systèmes d'IA est déterminante. La nature des données d'entraînement devra être représentative des situations traitées et inversement les usages devront être limités aux domaines d'applicabilité prévus, correspondant aux données d'entraînement utilisées.

Ces systèmes peuvent à terme faire évoluer les approches de la santé au travail. Ils peuvent notamment inciter à une individualisation accrue du suivi des travailleurs en intégrant des facteurs de risque individuels (génétique, hygiène de vie) au côté des facteurs de risques professionnels et par conséquent conduire à des mesures de prévention de plus en plus individualisées, ce qui pose des questions d'éthique. Il y a donc un enjeu important autour des compétences, des objectifs et du contrôle des acteurs qui développent et utilisent ces outils.

Implications pour les acteurs de la S&ST

Les acteurs de la veille sanitaire et de la prévention sont confrontés à la question de leur montée en compétences sur la compréhension de ces solutions. Il s'agira pour eux de pouvoir non seulement conduire des études sur ces systèmes, mais également de pouvoir développer une capacité de conseil et d'assistance à l'intention des parties prenantes : administrations, partenaires sociaux...

Une implication des organismes de la santé et sécurité au travail dans les instances de normalisation et le dialogue avec les régulateurs seront également des enjeux importants pour s'assurer que les solutions développées constituent des avancées en prévention et ne conduisent pas à une dérive de l'approche de la santé au travail.

Il leur faudra aussi arriver à gérer quelques écueils possibles. Il y a d'une part le risque de se focaliser sur les outils eux-mêmes, au détriment d'autres missions et approches. Ces dispositifs étant dépendants des données disponibles, ils peuvent par conséquent orienter l'attention des acteurs vers des domaines ou des populations pour lesquels ils disposent de données au détriment d'autres, rendus « invisibles » faute de données exploitables disponibles (cela pourrait être le cas pour des facteurs organisationnels par exemple). D'autre part, si les développements de ces dispositifs nécessitent des compétences spécifiques et des capacités de stockage, de protection, de traitement des données importantes, ils ne pourront être gérés que par des acteurs disposant de moyens importants (publics ou privés). Cela pourrait contribuer à une centralisation du pilotage de la santé et sécurité au travail et possiblement au creusement d'un écart entre les informations de pilotage centralisées issues de ces outils et les contextes de travail réels dans les entreprises.

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

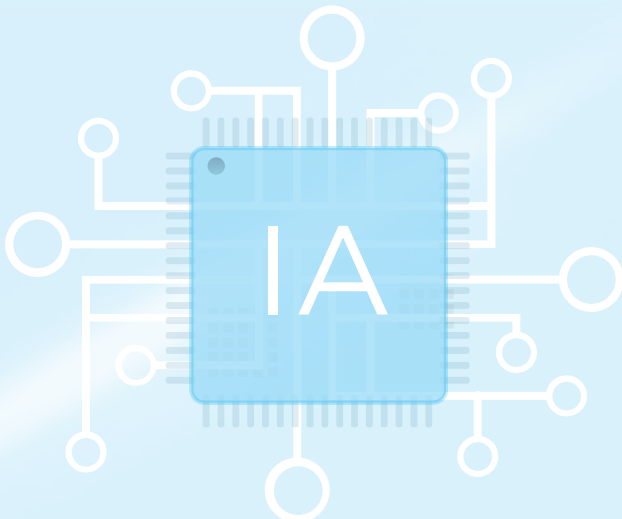
L'intelligence artificielle permet d'extraire du sens à partir de volumes considérables de données. C'est le cas notamment pour l'analyse de flux vidéo ou audio en temps réel, de paroles ou encore de données biométriques. Combinés à des dispositifs de captation connectés, les algorithmes d'IA peuvent réaliser une surveillance instantanée ou consolidée des situations observées. Pour la S&ST, de nombreuses initiatives promettent d'alerter de dangers imminents ou de prévenir de troubles chroniques en informant les opérateurs de mauvaises pratiques. Ces applications se basent sur deux typologies : les objets de surveillance de scène (exemple : caméra intelligente de détection de situations dangereuses dans un entrepôt) et les équipements à porter (*Wearable technology*) qui mesurent et transmettent des données biométriques du travailleur (exemple : mesure de la fréquence cardiaque pour la prévention des maladies cardiovasculaires) ou des positions articulaires (pour la prévention des lombalgies).

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. Ils 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 de mesures biométriques.

■ Potentiel de développement

Actuellement, les avancées sont liées à l'évolution en termes de puissance de calcul qui permet une augmentation de la performance et de la précision. La consommation d'énergie des puces embarquées est de plus en plus faible. Les développements dans le domaine quantique pourraient encore augmenter cette capacité de calcul et ainsi augmenter la pertinence, réduire la compacité, de même que la consommation d'énergie (sur le véhicule). Tout cela pourra en définitive permettre de réduire les prix de ces solutions qui limitent actuellement fortement leur déploiement. La 5G permettra aussi d'assurer un niveau de sécurité élevé en réduisant les temps de latence.



Atouts et points de vigilance

La facilité d'usage de ces technologies permet le développement de nouveaux outils performants dans le domaine de la prévention des risques professionnels, qui peuvent aider à réduire de façon significative la sinistralité. Elles apportent la possibilité de sécuriser les environnements, y compris pour des salariés peu formés ou sensibilisés aux règles de sécurité du site (comme les intérimaires, les nouveaux embauchés ou les intervenants extérieurs). La disponibilité de ces outils constituant une promesse d'amélioration de la sécurité dans les environnements de travail, elle pourrait conduire le législateur et le juge à renforcer leurs exigences en matière d'obligation de moyens et de résultats de l'employeur. Enfin, du fait des informations qu'ils peuvent apporter sur les situations et comportements à risque, ils peuvent aider au ciblage des actions de prévention en fonction de l'occurrence des presque-accidents et à prioriser les efforts de formation sur certains publics : intérimaires, nouveaux embauchés, chargés de la maintenance.

Cependant, un usage inadapté, détourné ou l'absence d'une réflexion préalable sur l'organisation de l'intégration de ces nouvelles technologies, pourrait conduire à des effets délétères en matière de santé et de S&ST. 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 la machine, 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. Cette déresponsabilisation à la fois collective et individuelle en matière de prévention pourrait être accompagnée de l'apparition de comportements déviants (comme le fait de chercher à éviter la surveillance opérée par ces technologies) possiblement générateurs d'accidents. Ces outils pourraient aussi favoriser le développement d'approches plus coercitives que préventives, avec pour conséquence de reporter la responsabilité sur le salarié en cas d'accident.

Détournées de leur finalité première de prévention, ces technologies sont aussi susceptibles 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.

Implications pour les acteurs de la S&ST

La collecte et l'exploitation de masses importantes de données sur les lieux de travail offrent une opportunité d'analyse inédite pour objectiver les expositions et révéler certains facteurs de risques.

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 grand pas la recherche dans le domaine de la prévention en santé et sécurité au travail. 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 et 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 à mettre en œuvre dans leur entreprise.

3. Robotique avancée recourant à l'IA

Certaines technologies de robotique avancée (embarquant de l'IA) offrent des solutions potentiellement bénéfiques à la S&ST. Pour cet atelier, deux types de cas d'usage ont été identifiés : 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 ;
 - 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'usages 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 sur 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 :

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

Atouts et points de vigilance

Le principal atout de ces technologies réside dans la réduction, voire la suppression des risques auxquels les opérateurs sont soumis, grâce au remplacement du salarié 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 aussi favoriser le maintien ou le retour à l'emploi d'opérateurs déchargés de la partie physique du poste. Par ailleurs, elles peuvent constituer une opportunité de requalification des opérateurs, qui une fois déchargés des tâches répétitives désormais réalisées par le robot, peuvent monter en compétence, notamment dans les domaines liés à la maintenance et la supervision de ces nouveaux outils.

Cependant, ces technologies peuvent également faire émerger certains risques.

Une organisation du travail, qui n'aurait pas impliqué une expérimentation en amont et une réflexion collective sur l'intégration de ces nouveaux outils, pourrait conduire à 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, une dévalorisation des opérateurs, une réduction de leurs marges de manœuvre et une potentielle perte de sens au travail. Les risques psychosociaux pouvant en découler sont donc nombreux.

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 situations dégradées et être la cause d'accidents du travail. La perte de confiance qui pourrait en résulter pourrait mener à un abandon de ces technologies en dépit des atouts qu'elles représentent.

Implications pour les acteurs de la S&ST

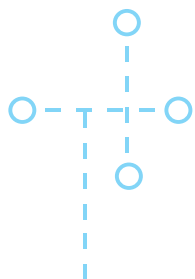
Les acteurs de la prévention, de par leur pratique pluridisciplinaire, leur expertise des risques mécaniques et organisationnels et des dispositifs juridiques, sont à même d'accompagner l'émergence de ces technologies et d'intervenir en amont, lors des phases de développements. Cependant, ces acteurs pâtissent pour le moment d'un manque de compétences en matière d'intelligence artificielle et de robotique intelligente et de terrains d'expérimentation.

À l'avenir, une évaluation de la pertinence du recours à ces outils – réalisée autour d'un dialogue social au sein des entreprises – pourrait être favorisée par des méthodes adaptées. Celles-ci pourraient émerger de travaux de veille et de recherche concernant les risques spécifiques liés à l'IA et de collaborations entre acteurs. Ces recherches pourraient également contribuer à la validation de la robustesse des systèmes, ainsi qu'à faire évoluer la réglementation. Des produits d'information et de sensibilisation pourraient également en découler.

Des efforts de formation – des préventeurs, des concepteurs mais également des salariés – sont nécessaires pour permettre aux acteurs concernés de pouvoir débattre de ces sujets et instaurer confiance et acceptabilité.

Annexe

La méthodologie en cinq phases



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

1. 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 au moment d'engager cet exercice. 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.

2. Analyse du sujet

L'analyse des facteurs clés a permis d'établir une liste de 12 variables réparties au sein de trois composantes : évolution de l'offre en IA, acceptabilité 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 réunions.

3. 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.

4. Étude de cas d'usage

Trois ateliers ont ensuite été organisés afin de poursuivre la réflexion sur les usages possibles de l'IA et de ses conséquences potentielles en prévention aussi bien pour les acteurs de la santé et la sécurité au travail 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 ; robotique avancée.

5. 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.



Avis au lecteur

La prospective n'est pas une prédiction de l'avenir. Elle n'est pas non plus une prévision qui serait le prolongement des tendances passées.

La prospective prend en compte les tendances et les discontinuités pour décrire des futurs possibles et proposer une aide à la prise de décision.

Les travaux présentés dans ce document sont le fruit d'une production collective. Ils ne reflètent pas nécessairement les opinions et souhaits des participants y ayant contribué et n'engagent ni les organismes dont ils sont issus, ni l'INRS.

Dans le cadre de son activité de prospective, l'INRS a animé une réflexion collective visant à explorer comment des systèmes mobilisant de l'intelligence artificielle pourraient être utilisés afin d'améliorer la santé et la sécurité au travail à l'horizon 2035. Quelles sont les opportunités et menaces pour la prévention des risques professionnels ? Dans quels domaines des avancées sont-elles possibles ? À quelles conditions ? Comment les acteurs peuvent-ils se préparer ?



Institut national de recherche et de sécurité
pour la prévention des accidents du travail
et des maladies professionnelles
65, boulevard Richard-Lenoir 75011 Paris
Tél. 01 40 44 30 00 • info@inrs.fr

Édition INRS PV 20

1^{re} édition | novembre 2022 | 200 ex. | ISBN 978-2-7389-2802-3

L'INRS est financé par la Sécurité sociale
Assurance maladie - Risques professionnels

www.inrs.fr   