

Rapport

L'intelligence Artificielle au Travail

Accompagner
& sécuriser
les initiatives
collaborateurs

01

07

25



Auteur : Yann Ferguson, Inria

Équipe projet : Isabelle Hilali, Jeanne Godard pour datacraft ; Edouard Havis, Jean-Michel Lefèvre, Laurence Mari, Julia Savalli pour Inria

SOMMAIRE

LES ENSEIGNEMENTS CLÉS DE CETTE ÉTUDE	5
REMERCIEMENTS	7
1. INTRODUCTION : UNE INTÉGRATION DE L'IA PILOTÉE PAR LES USAGES ?	9
2. L'INTELLIGENCE ARTIFICIELLE AU TRAVAIL : DU CAS D'USAGE AU « BRICOLAGE »	13
2.1. L'IA est-elle enfin devenue employable ?	13
2.1.1. « Dans 20 ans... »	13
2.1.2. L'IA d'apprentissage : la fin du paradoxe de Polanyi ?	15
2.1.3. Des effets pressentis significatifs sur l'emploi et le travail	16
2.2. Le plafond de verre du déploiement des cas d'usage : l'IA ne passe pas la période d'essai	17
2.3. L'intelligence artificielle, une initiative des collaborateurs : à chacun son usage	20
2.3.1. L'IA, initiative discrète des collaborateurs	20
2.3.2. Le paradigme du bricolage	22
2.3.3. De surprenants assistants	24
2.3.4. La valeur recherchée : s'augmenter plutôt que déléguer	30
2.4. Des initiatives qui inquiètent et menacent les bricoleurs et les organisations	33
2.4.1. Des usages profanes qui mènent à des mésusages	33
2.4.2. Des pratiques clandestines qui intègrent des données sensibles engageant la responsabilité des organisations	36
2.4.3. Des pratiques honteuses sur fond de peur pour l'avenir du travail	38
2.4.4. Une clandestinité au détriment du collectif de travail	42
2.5 Du côté des organisations syndicales	44
3. ENCADRER LES BRICOLEURS : LES RÉPONSES DES EMPLOYEURS AUX USAGES INFORMELS	54
3.1. Méthodologie de notre enquête : à la rencontre d'organisations pionnières	54
3.2. Enseignements généraux : des initiatives qui accélèrent l'intégration et la gouvernance de l'IA	56
3.3. Une définition élargie du Shadow AI	58
3.3.1. L'usage individuel informel et clandestin : « ils ont tout intérêt à rester discrets, voire à se cacher »	59
3.3.2. L'usage informel légitimé : « On a aussi choisi de faire confiance aux employés »	62
3.3.3. L'usage informel sécurisé : « des modules développés sans réel accompagnement »	66
3.3.4. L'usage socialisé : « présenter ses meilleurs prompts »	68

3.3.5. L'usage « tacite » via IA embarquée : « un logiciel officiel non-maîtrisé »	69
3.3.6. L'usage masqué au client : « le client achète surtout du temps humain »	70
3.3.7. Le shadow learning : « des espoirs qui ne cadrent pas avec la ligne de l'organisation »	71
3.4. Trouver l'équilibre entre opportunités et menaces pour les organisations	72
3.4.1. Une étude de marché interne : « qu'est-ce qu'on pourrait avec ça ? »	74
3.4.2. Des mésusages et des risques : « un vrai défi d'entreprise »	75
3.4.3. Une intégration anarchique	77
3.4.4. Une accélération de la stratégie IA : « apporter vision et sens »	79
3.4.5. Une montée en maturité des organisations provoquée par les initiatives des collaborateurs	82
3.5. Quatre postures face aux pratiques spontanées de collaborateurs	84
3.5.1. Ouverture et pilotage des pratiques informelles	84
3.5.2. De l'ignorance à l'innovation encadrée	85
3.6. Convertir une créativité diffuse en intelligence collective durable et en stratégie robuste	86
4. RECOMMANDATIONS ET OUTILS	88
4.1. Sortir du « shadow AI » ? Des pratiques spontanées à des usages institutionnalisés.	88
4.1.1. Piloter : formaliser l'invisible	89
4.1.2. Partager : socialiser les pratiques et les savoirs	90
4.1.3. Institutionnaliser et diffuser : Sécurité	91
4.1.4. Un enchaînement cumulatif	92
4.2. Libérer les usages par des conflits de qualité	93
4.2.1. Cartographier les « user-stories »	93
4.2.2. La valeur recherchée dans le recours à l'IA	93
4.2.3. La valeur obtenue par le recours à l'IA	94
4.2.4. Les gestes professionnels pour optimiser le recours à l'IA	94
4.2.5. La sérendipité : les bénéfices inattendus de l'IA	95
4.2.6. Mise en débat collective des référentiels	96
4.3. Elaborer un cadre de confiance pour un usage responsable de l'IA générative	97
4.3.1. Distinguer clairement les usages autorisés et interdits	98
4.3.2. Définir la responsabilité des collaborateurs	99
4.3.3. Prendre en compte les risques spécifiques à l'IA générative	100
4.3.4. Accompagnement des collaborateurs et évolution continue de la charte	101
4.3.5. Modalité de pilotage et gouvernance de la charte IA générative	102
4.4. Une formation holistique à l'intelligence artificielle	103
CONTRIBUTEURS	106

LES ENSEIGNEMENTS CLÉS DE L'ÉTUDE

Shadow AI : des bricolages individuels instructifs mais risqués

Depuis l'essor de l'IA générative fin 2022, de nombreux salariés s'approprient ces outils de manière informelle, en dehors des cadres officiels. Ce phénomène, appelé **“Shadow AI”**, révèle un décalage persistant entre les ambitions stratégiques des entreprises et les usages réels sur le terrain. Alors que seuls 20 % des projets d'IA pilotés “par le haut” sont industrialisés, les pratiques informelles se développent car elles répondent à des besoins concrets et semblent favoriser l'autonomie, l'efficacité et la créativité des collaborateurs. L'étude décrit des formes multiples de Shadow AI, ces pratiques informelles pouvant être secrètes, clandestines, mais également connues de la direction tout en demeurant personnelles ou cachées aux clients. Les collaborateurs utilisent l'IA pour des tâches variées (rédaction, traduction, structuration d'idées...), dans une logique de « bricolage » mêlant expérimentation et autonomie. Ces usages, bien que générant un sentiment d'efficacité individuelle, exposent à des risques collectifs : sécurité des données, non-conformité, biais, atomisation des collectifs, apprentissage et inégalités.

Une opportunité stratégique à encadrer

Notre étude auprès de 14 organisations pionnières montrent comment celles-ci entreprennent de désindividualiser les pratiques pour les convertir en opportunités stratégiques. Elle montre en effet que le Shadow AI, loin d'être un simple risque, peut devenir un accélérateur de stratégie IA s'il est reconnu, partagé et encadré.

Il permet de :

- Identifier des usages métiers pertinents ;
- Engager un dialogue interfonctionnel (DSI, RH, métiers...) ;
- Transformer des pratiques diffuses en intelligence collective.

Quatre postures organisationnelles sont identifiées, selon le degré d'ouverture à des solutions tiers et de pilotage par les organisations : indifférence, permissivité, accompagnement structuré ou verrouillage. Plutôt qu'une interdiction ou un laissez-faire, se balisent des trajectoires d'innovation qui recherchent l'équilibre entre reconnaissance des initiatives individuelles et capitalisation collective, entre sécurisation des données et des meilleures pratiques et gouvernance ouverte et négociée.

Une trajectoire en trois temps : Piloter- Partager- Sécuriser

Les organisations pionnières sortent du Shadow AI par une démarche progressive :

- **Piloter** : rendre les usages visibles, poser un cadre minimal ;
- **Partager** : organiser des ateliers, socialiser les pratiques ;
- **Sécuriser** : mettre en place des environnements sécurisés, élaborer des chartes et formateurs.

Toutes les organisations pionnières ne suivent pas strictement les mêmes étapes ni le même rythme. Elles ont toutefois en commun de vouloir responsabiliser les utilisateurs, puis de désindividualiser les pratiques et enfin de les inscrire dans un cadre de confiance.

Des recommandations et des ressources concrètes

Notre étude aboutit à des recommandations et propose des ressources pour les activer.

1. **Négocier collectivement la sortie du Shadow AI**: Onze actions sont détaillées autour de trois objectifs: piloter, partager, sécuriser.
2. **Des ateliers pour co-définir la qualité du travail avec l'IA**: Mettre en discussion les transformations du travail liées à l'IA, en faisant émerger des critères partagés de qualité, au lieu de se limiter à des normes techniques ou des règles descendantes.
3. **Elaborer un cadre de confiance co-construit**: Construire une régulation des usages de l'IA générative qui combine clarté, adaptabilité et légitimité sociale des usages.
4. **Définir une formation "holistique" à l'IA**: Engager une montée en compétences au-delà des seuls savoir-faire techniques, en formant des professionnels capables d'un usage critique, réfléchi et responsable de l'IA.

Ces ressources appellent à dépasser les réponses technocentrées pour construire une gouvernance partagée de l'IA, ancrée dans les usages réels et portée par l'intelligence collective et le dialogue social technologique.

REMERCIEMENTS

Nous souhaitons exprimer notre profonde gratitude à l'ensemble des acteurs qui ont contribué à la richesse de cette étude sur les usages émergents de l'intelligence artificielle générative dans les environnements professionnels.

Nos remerciements vont tout d'abord aux entreprises et institutions pionnières qui ont accepté de partager leurs pratiques, leurs questionnements et leurs expériences : Airbus, l'Assurance Maladie, Believe, CHU de Montpellier, Crédit Agricole, Ekimetrics, L'Oréal, MAIF, Malakoff Humanis, Métropole de Montpellier, Ministère des Armées, Région Île-de-France, Skyrock et Veepee.

Nous tenons également à saluer l'engagement de Nicolas Blanc, secrétaire national à la transition économique à la CFE-CGC, pour ses apports éclairés sur les enjeux sociaux et professionnels de l'automatisation.

Nos remerciements sincères à Odile Chagny, économiste à l'Institut de Recherches Économiques et Sociales (IRES) et Erell Thevenon-Poullennec, déléguée générale de l'Institut pour l'innovation économique et sociale (ZIES), pour leurs contributions à l'amélioration de nos livrables dans la perspective notamment du dialogue social.

Un grand merci à Samuel Durand, auteur et conférencier, pour ses travaux inspirants sur le futur du travail.

Nous remercions chaleureusement Apolline Guillot, rédactrice en chef de Philonomist, pour son regard philosophique sur les usages de l'IA au travail et sa capacité à articuler enjeux techniques et questions de sens.

Inria remercie le Ministère du Travail et de l'Emploi pour le partenariat engagé dans le cadre du LaborIA qui a initié la constitution d'une équipe de recherche dédiée aux enjeux de l'intégration de l'intelligence artificielle dans les environnements de travail.

Enfin, nous adressons nos plus vifs remerciements au Village by CA pour leur accueil chaleureux et leur soutien logistique, qui ont grandement facilité la tenue de nos ateliers et temps d'échange.

À toutes et tous, votre collaboration a été essentielle pour nourrir une réflexion collective, exigeante et ouverte sur l'appropriation responsable de l'intelligence artificielle dans le monde du travail.

Les propos, analyses et recommandations formulés dans ce rapport n'engagent en rien les organisations citées ni les personnes remerciées. Ils relèvent de la seule responsabilité des auteurs.

Inria

Inria est l'institut national de recherche en sciences et technologies du numérique et a la responsabilité depuis janvier 2024 de l'Agence de programmes dans le numérique pour renforcer les dynamiques collectives de l'enseignement supérieur et de la recherche. La recherche de rang mondial, l'innovation technologique et le risque entrepreneurial constituent son ADN. Au sein de 220 équipes-projets, pour la plupart communes avec les grandes universités de recherche, plus de 3 800 scientifiques y explorent des voies nouvelles, souvent dans l'interdisciplinarité et en collaboration avec des partenaires industriels pour répondre à des défis ambitieux. Institut technologique, Inria soutient la diversité des voies de l'innovation : de l'édition open source de logiciels à la création de startups technologiques (Deeptech). Inria est labellisé Institut Carnot confirmant sa volonté de rapprochement entre la recherche et le monde industriel.

datacraft

datacraft est un modèle unique de Club, inspiré du compagnonnage, pour permettre aux data scientists/data ingénieurs de se former entre pairs et permettre aux entreprises d'accélérer leurs projets IA. Le premier Club a ouvert à Paris en 2020. Il compte aujourd'hui 50 grandes entreprises membres (Accor, Airbus, BNP Paribas, Bouygues... mais aussi AP-HP, CNES, Marine nationale...), des startups et plus de 600 chercheurs et freelances experts en IA.

Etre membre du Club datacraft c'est :

- permettre aux collaborateurs experts data de participer aux ateliers hebdomadaires organisés par datacraft et rejoindre une communauté d'experts data/IA
- trouver des talents data/IA : proposer des missions aux chercheurs et freelance en résidence ; trouver des partenaires pour des projets.
- accéder à toutes les autres activités du Club : soirées CDO, rencontres data gouvernance, ateliers pour les différents métiers : RH, finance, juridique...et accéder à tous les services du Club : formations innovantes d'acculturation, formations tech sur mesure, escape game, hackathon, Kaggle, journées inspirantes autour de l'IA pour les comex, les équipes...

INTRODUCTION : UNE INTÉGRATION DE L'IA PILOTÉE PAR LES USAGES ?



Dès le départ, j'ai senti que l'arrivée de l'IA allait révolutionner notre façon de travailler — y compris mon propre métier. Je commençais à prendre la mesure de la puissance de cette technologie. Je n'en parlais pas à mon manager, mais certains jours, je pouvais facilement gagner deux heures. Personne ne s'en rendait compte ! Mes collègues avaient vaguement testé les premières versions de ChatGPT et les avaient vite mises de côté. Pendant nos pauses, j'essayais d'orienter la conversation vers l'IA, mais ça ne suscitait pas vraiment leur intérêt. Je me souvenais de ce que ma mère m'avait raconté : comment les ordinateurs avaient remplacé les sténographes. Certains avaient perdu leur emploi, d'autres avaient appris à se servir des ordinateurs et à en faire bien plus ! Mais l'arrivée de l'IA me paraissait être un bouleversement encore plus profond. Je me demandais si je devais parler de mon utilisation à mes supérieurs. Est-ce qu'ils allaient m'imposer plus de travail ? Ou forcer tout le monde à s'y mettre ? Mes collègues n'avaient pas l'air motivés, et je craignais que si leur productivité ne suivait pas, l'entreprise décide de s'en séparer. Durant ces premiers mois, je ne savais pas du tout comment gérer la situation...



Dans la docu-série « A.I. at Work » écrite et produite par Samuel Durand¹, l'héroïne raconte sa découverte de l'IA générative et de son utilisation professionnelle qu'elle cache à ses responsables et à ses collègues : rédactions d'e-mails, relecture de contrats, création de présentations... Grâce à l'IA, elle gagne parfois « facilement deux heures dans une journée », sans que « personne ne s'en rende compte », mais elle redoute que cette efficacité perçue n'entraîne des effets indésirables : une surcharge de travail, une injonction collective à adopter l'outil, ou pire, des licenciements parmi ses collègues moins performants. L'inconfort est également déontologique : « Ça me donnait l'impression de tricher, mais personne ne semblait s'en rendre compte ». La narratrice décide pourtant de révéler à sa direction l'étendue de son usage de l'IA, ainsi que le potentiel qu'elle y voit pour l'ensemble de l'entreprise. Plutôt que de la sanctionner, sa direction lui confie une mission : « Et ce vendredi pluvieux de mai, je suis sortie du bureau avec la plus grosse responsabilité qu'on m'ait jamais confiée : le déploiement de l'IA pour tous les collaborateurs de l'entreprise – j'avais un an pour le faire. » Forte de cette nouvelle légitimité, elle élabore un plan d'action, accompagne ses collègues dans leur montée en compétences et pose les bases d'une gouvernance éthique.

L'arrivée de l'IA générative initiée par la sortie mondiale de ChatGPT, un grand modèle de langage (LLM) conçu par OPEN AI a accéléré l'introduction de l'IA dans les environnements de travail. Jusque-là, l'avènement de l'IA dans les organisations, annoncé dans les années 2010 comme majeur et rapide, tardait à se réaliser. De nombreux travaux de recherche et d'observatoires internationaux se rejoignent pour constater un faible taux de déploiement (20% environ) des « preuves de concept » (Proof of Concept, POC), développées par les organisations. Du côté de la science, la difficulté à rendre les résultats interprétables, explicables, fiables et robustes limite encore la confiance dans les systèmes d'IA (SIA), a fortiori pour les tâches critiques. Dans les organisations, la faiblesse des infrastructures et l'insuffisance de la quantité et de la qualité des données sont pointées du doigt. Mais de plus en plus de chercheurs et d'experts s'accordent pour considérer que ce sont les causes organisationnelles et humaines qui sont au cœur des difficultés actuelles. Or, tandis que les cas d'usage expérimentés par les organisations peinent à passer la période d'essai, le « bricolage » des collaborateurs avec l'IA générative ne cesse d'augmenter en volume et en fréquence. Le dernier baromètre Ifop/Talan² montre ainsi que 43% des utilisateurs d'IA génératives les utilisent dans un cadre professionnel, 29% perçoivent un gain de productivité et d'efficacité supérieur à 40% grâce aux IA génératives. Personnellement comme professionnellement, ils font des recherches (34%), corrigent leurs fautes d'orthographe (23%), augmentent leur créativité (13%) et disent que l'IA augmente leur confiance en eux-mêmes (10%). Malgré une adoption en hausse, plusieurs défis subsistent. Si 52 % des utilisateurs professionnels se sentent encouragés à utiliser les IA génératives par leurs employeurs, seuls 15 % ont bénéficié d'une formation, tandis que 73 % des Français estiment ne pas disposer des connaissances nécessaires pour les utiliser efficacement. Plus préoccupant encore, seuls 9 % des salariés déclarent que leur entreprise met ces outils à leur disposition, et près de la moitié (49 %) affirment que leur organisation n'a pas l'intention de le faire. En outre, cette progression remarquable dissimule toutefois une accentuation de la fracture générationnelle : 85 % des 18-24 ans utilisent désormais les IA génératives (contre 68 % en 2024), 63 % des 25-34 ans (47 % en 2024), mais seuls 31 % des personnes de 35 ans et plus y ont recours. **L'ensemble témoigne d'un effet de ciseau croissant entre l'augmentation des usages par les collaborateurs et la faible institutionnalisation de ces usages qui se déploient alors sans gouvernance claire des employeurs.** Ainsi, ChatGPT est devenu en 2023 la première application « shadow IT » de web. Cette notion ancienne désigne l'ensemble des technologies (logiciels, applications, services cloud, etc.) utilisées au sein d'une organisation sans l'approbation, la supervision ou la connaissance du service informatique.


¹ <https://www.wipdocumentary.com/aiatwork>

² Baromètre 2025 IFOP pour Talan - Les Français et les IA génératives. <https://www.talan.com/france/fr/Ifop-Talan-barometre-2025-francais-ia-generatives>

Cela inclut des outils que les employés adoptent de manière autonome pour répondre à leurs besoins professionnels en contournant les procédures officielles. Bien que souvent motivé par la recherche d'efficacité ou de flexibilité, le shadow IT pose d'importants risques en matière de sécurité, de conformité, de gouvernance des données et de continuité des systèmes. Son pendant dans le domaine de l'intelligence artificielle est appelé « Shadow AI ».

Le Shadow AI représente une opportunité en ce qu'il témoigne de l'initiative des employés à intégrer spontanément l'IA dans leur travail pour gagner en efficacité et en créativité. Plus globalement, il permet également de repenser les méthodologies antérieures d'intégration de l'IA dans les organisations et les métiers en explorant de façon ascendante des cas d'usage innovants avant même leur reconnaissance officielle par l'organisation. Cependant, cette adoption non encadrée expose l'entreprise à des risques majeurs : fuites de données sensibles, non-conformité réglementaire, et dépendances techniques non maîtrisées. Sans stratégie claire, le Shadow AI peut fragiliser la sécurité, la cohérence et la souveraineté numérique de l'organisation. Elle peut également conduire à des mésusages, une perte de qualité du travail, à une atomisation du collectif de travail ou encore à une remise en cause des engagements éthiques des organisations en matière d'IA. Ces aspects sont développés dans la première partie de ce rapport.

La deuxième partie de ce rapport repose sur une étude menée par Inria et datacraft auprès de 14 « entreprises pionnières » qui ont accepté de témoigner sur l'appréhension de ces usages spontanés de l'IA générative par leurs collaborateurs. Le choix de l'échantillon était d'abord déterminé par l'adoption d'une posture volontariste de ces organisations vis-à-vis du « shadow AI » : avoir identifié ces pratiques et entrepris des actions concrètes : interdiction, sécurisation, sensibilisation, régulation, formation, animation, etc. Ces auditions ont suivi une grille d'entretien semi-directive autour de l'axe principal de l'étude : **Comment accompagner les usages spontanés de l'IA par les collaborateurs en bénéficiant de la fenêtre d'opportunité tout en réduisant les risques qu'ils engendrent ?** Car les bricolages individuels, souvent invisibles des directions, font entrer l'IA dans les métiers avant même qu'une stratégie soit posée. Notre enquête, fondée sur quatorze entretiens et une journée de co-construction, montre que les organisations pionnières suivent trois paliers : Piloter (rendre visibles les pratiques et poser les garde-fous), Partager (socialiser les savoir-faire et construire un référentiel commun) et Sécuriser (déployer des outils sécurisés et massifier la formation) .



La troisième partie de ce rapport exploite les deux précédentes pour exprimer des recommandations et proposer des ressources aux organisations qui souhaitent mieux accompagner le « Shadow AI » et en faire un levier stratégique de l'intégration sociale de l'IA.

- Élaborer un chemin responsable et équitable de sortie du « shadow AI ».
- Engager des conflits de qualité du travail avec l'IA.
- Formaliser une gouvernance des pratiques spontanées.
- Former à toutes les dimensions de l'IA.

Pour mener ce travail, notre équipe s'est régulièrement appuyée sur des SIA dont certains n'ont pas été validés par leurs directions respectives. Les outils et usages seront spécifiés en note de bas de page au début de chaque section. Quelques remarques sont faites sur l'intérêt de ces applications pour cette recherche.

L'INTELLIGENCE ARTIFICIELLE AU TRAVAIL : DU CAS D'USAGE AU « BRICOLAGE »³

Alors que l'intelligence artificielle a longtemps été pensée comme une technologie à intégrer par le haut, à travers des cas d'usage formalisés et des stratégies d'optimisation, son déploiement effectif dans les organisations suit parallèlement un chemin alternatif. Loin des promesses industrielles ou des POC par les directions, c'est par les usages spontanés, souvent invisibles et non encadrés, que l'IA générative s'installe dans les pratiques professionnelles. Cette section explore ce renversement : comment, en miroir des limites des approches descendantes, les collaborateurs s'approprient l'IA sur le mode du bricolage — avec leurs propres moyens, pour leurs propres besoins — transformant progressivement leurs gestes professionnels et les pratiques collectives. Mais cette appropriation spontanée, si elle témoigne d'une ingéniosité réelle, se heurte aussi à ses propres limites : elle fragilise la cohérence collective, expose les organisations à des risques juridiques ou techniques, et laisse les travailleurs souvent seuls face à des dilemmes éthiques, sans cadre partagé ni reconnaissance institutionnelle.

2.1. L'IA est-elle enfin devenue employable ?

2.1.1. “Dans 20ans...”

Le projet d'automatisation par l'IA entend dépasser la simple imitation d'actions régies par des règles fixes : en lui conférant une « rationalité procédurale », elle doit proposer des solutions satisfaisantes dans des contextes où la rationalité est intrinsèquement limitée, ambition qu'Herbert Simon annonçait déjà en 1958 sans qu'elle se soit concrétisée à ce jour : « d'ici vingt ans, les machines seront capables d'effectuer n'importe quel travail réalisable aujourd'hui par un humain ».

³Dans cette section, nous avons régulièrement utilisé l'application « NotebookLM » qui permet aux utilisateurs d'interagir avec leur document, quelque-soit leur langue d'origine. Tous les documents utilisés étant publics, cela ne posait pas de problème de sécurité. Nous avons également utilisé plusieurs modèles pour reproduire des éléments présents sur des documents, rendant une capture d'écran modifiable. Nous avons ponctuellement utilisé le Chat de Mistral et ChatGPT 40 pour nous aider à harmoniser le style d'écriture entre des morceaux de textes extraits de différents documents. Si le résultat était satisfaisant, cela a déplacé notre travail d'auteur à éditeur et la satisfaction a davantage résulté du sentiment d'avancer plus vite que de faire. Nous avons découvert qu'une supervision efficace n'était possible que si les textes générés étaient courts. L'IA générative produit beaucoup et trop vite pour relire avec vigilance. Enfin, nous avons centré les usages d'écriture aux “investissements de forme” : introductions et conclusions brèves, transitions plutôt que sur la connaissance, dont nous voulions garder un total contrôle. A noter : lorsque nos prompts intégraient une référence aux sciences humaines et sociales, les textes de sorties étaient rédigés en écriture inclusive.

Pour expliquer la difficile automatisation avancée du travail, de nombreux chercheurs et experts recourent au paradoxe de Polanyi développé par le philosophe dans « The Tacit Dimension »⁴ : « Nous en savons plus que nous pouvons en dire... ». Une grande partie de notre savoir est en effet tacitement incorporée dans nos gestes, nos perceptions et notre expérience pratique, sans que nous soyons capables de le formaliser ou de l'exprimer clairement. Ce paradoxe souligne une limite structurelle de l'automatisation, notamment pour les activités qui mobilisent des compétences sensori-motrices fines (telle que la conduite d'un véhicule), du jugement professionnel, c'est-à-dire des décisions prises dans l'incertitude, fondées sur des signaux faibles ou des contextes ambigus ou des interactions humaines complexes, comme dans le soin, l'éducation ou la négociation. Puisque ces savoirs ne sont ni codifiés, ni entièrement verbalisables, il est difficile de les traduire en instructions précises ou en données exploitables pour des systèmes automatisés.

Ces limites rappellent que le courant symbolique dominant, qui fait à écho à une vision taylo-fordienne où toute activité serait décrivable, décomposable, standardisable et optimisable, a longtemps sous-estimé la dimension corporelle et adaptative de l'intelligence. C'est pourquoi il est devenu commun d'imputer les échecs antérieurs de l'automatisation du travail par l'intelligence artificielle à l'hégémonie du courant « symbolique » ou « logique » sur la discipline, suite notamment à l'ouvrage de Minsky et Papert, « Perceptrons », publié en 1969, qui a condamné durablement le courant connexionniste de l'IA. L'approche logique concentrerait l'IA sur les raisonnements sciemment mobilisés par les professionnels durant leur activité, sans considérer les savoirs et savoir-faire tacites qui leur permettent de s'adapter aux situations réellement vécues échappant aux procédures et aux processus.

En 2006, Nilsson avait d'ailleurs proposé de remplacer le célèbre test de Turing, épreuve emblématique mais controversée de mesure de l'intelligence des machines, par un « test de l'employabilité », jugé plus significatif pour comparer les intelligences artificielles et humaines⁵. Dans ce test, les programmes d'IA devraient être capables d'accomplir les tâches ordinairement effectuées par les humains, et les progrès vers l'IA de niveau humain pourraient être mesurés par la fraction des emplois humains qui peuvent être réalisés de manière acceptable par des machines. Pour atteindre cet objectif, il faudrait selon Nilsson développer des systèmes à usage général et éducatifs capables d'apprendre et d'être enseignés pour effectuer diverses tâches humaines, plutôt que de construire de nombreux systèmes spécialisés.

⁴ Polanyi, M. (2009), *The Tacite Dimension*, University of Chicago Press.

⁵ Nilsson, N. (2006), *Human-Level Artificial Intelligence? Be Serious!*, *AI Magazine* Volume 26 Number 4.

Une approche possible pour construire ces « systèmes habiles » serait de commencer par des systèmes ayant des capacités d'apprentissage minimales mais étendues, similaires à celles d'un enfant, et de les éduquer progressivement. Il suggère enfin que les systèmes ainsi conçus passent des examens professionnels testant les compétences requises pour divers emplois. Nilsson préfigure l'avènement de l'IA dite « moderne » comme approche libératrice de l'employabilité de l'IA.

2.1.2. L'IA d'apprentissage : la fin du paradoxe de Polanyi ?

L'IA moderne qui repose sur des techniques d'apprentissage plutôt que l'intégration de règles dans un programme va, pour beaucoup d'experts, conduire l'IA à dépasser ces limites historiques. L'apprentissage de la conduite par la Google Car, par exemple, démontre selon eux la maturité nouvelle prise par l'IA, la conduite étant depuis longtemps l'illustration majeure du paradoxe de Polanyi : « Les voitures sans conducteur ne sont pas une anomalie : elles font partie d'une évolution fascinante et plus générale », analysent Erik Brynjolfsson et Andrew McAfee dans leur célèbre ouvrage « Le deuxième âge de la machine »⁶. En 2011, une machine conçue par IBM, Watson, terrasse deux champions de jeu TV américains à Jeopardy, en langage naturel et dans les conditions du direct. En 2012, lors du concours annuel ImageNet de vision par ordinateur, un programme de reconnaissance d'images améliore significativement les performances des machines, faisant passer leur taux d'erreurs de 27% à 17%. En 2016 enfin, AlphaGo, un programme de jeu de go, terrasse le 3e joueur mondial, surprenant les experts par sa créativité. Ces machines ont en commun de réhabiliter une approche ancienne mais marginalisée, le « deep learning ». Ce courant « neuronal » repose sur des méthodes d'apprentissage à l'intérieur d'ensemble de données, à partir desquelles elles extraient des règles par corrélation, puis appliquent ces règles pour formuler des jugements sur des situations nouvelles. Grâce au big data qui lui fournit sa base d'apprentissage et l'activité de centaines de milliers de « travailleurs du clic » (Casilli, 2019⁷), le deep learning obtient des avancées significatives et rapides sur des tâches impossibles à formaliser par les approches traditionnelles : le langage, la traduction, la génération de textes, la reconnaissance des images, des émotions, des sons, etc. « Longtemps les progrès concernant les ordinateurs, les robots et les autres équipements numériques sont restés très lents ; puis en quelques années, ils se sont faits très rapides.

⁶ Brynjolfsson, E., McAfee, A. (2014), Le deuxième âge de la machine. Travail et prospérité à l'heure de la révolution technologique, Odile Jacobs, Paris.

⁷ Casilli, A. (2019), En attendant les robots, Seuil, Paris.

La technologie numérique a commencé à accélérer, réalisant des tâches qu'elle n'avait jamais maîtrisées, et montrant des compétences qu'elle n'était pas censée acquérir aussi vite » (Brynjolfsson, McAfee : 30). L'IA empirique serait désormais en capacité d'automatiser l'empirie des pratiques et des savoir-faire professionnels. En s'attaquant aux limites que posait le paradoxe de Polanyi, l'IA contemporaine semble désormais en mesure d'automatiser des compétences inaccessibles à la formalisation a priori. Pour de nombreux experts, cette mutation technologique ouvre la voie à des transformations profondes de l'emploi et du travail.

2.1.3. Des effets pressentis significatifs sur l'emploi et le travail

Le « deuxième âge de la machine » annonce une automatisation croissante des tâches cognitives : certaines machines pourraient décider mieux que l'humain, devenant de véritables substituts plutôt que de simples compléments. Cette perspective nourrit, depuis 2013, des projections spectaculaires sur l'avenir de l'emploi. L'étude emblématique de Frey & Osborne estime que 47 % des postes américains sont automatisables à moyen terme ; pour le démontrer, elle relie 70 professions à trois domaines où la technologie reste fragile : perception/manipulation, créativité et intelligence sociale – puis généralise l'algorithme à 632 métiers⁸. Un exercice similaire mené par le cabinet Roland Berger (2014) porte à 42 % la part d'emplois français menacés ; les plus vulnérables sont ceux dominés par des routines.

Par contraste avec ces travaux dits « centrés sur l'emploi », le Conseil d'orientation pour l'emploi privilégie une approche « centrée sur la tâche ». Seules les tâches routinières, clairement programmables, sont pleinement substituables ; les autres, plus complexes, révèlent une complémentarité entre l'IA et le travailleur. L'exposition d'un métier s'évalue donc à la faible présence de quatre compétences : flexibilité, adaptation, résolution de problèmes et interactions sociales. Ainsi, 10 % des emplois français présentent une forte probabilité d'automatisation (plus de 70% de tâches potentiellement automatisables), tandis que 50 % subiront surtout une transformation profonde (au moins 50%).

Au-delà de ces chiffres divergents, trois tendances se dégagent :

- L'IA accroîtra la productivité en automatisant les routines ;
- Les salariés libérés se recentreront sur des tâches à plus forte valeur et plus épanouissantes ;
- La créativité, la résolution de problèmes, l'adaptation, la dextérité et les relations humaines deviendront décisives.

⁸ Benedikt, C., Osborne, M. (2013). « The Future of Employment : How Susceptible are Jobs to Computerisation ? », Technological Forecasting and Social Change, vol. 11. 4, issue C.

Si les scénarios d'automatisation massive ont d'abord cristallisé les craintes autour de la disparition des emplois, une lecture plus fine des transformations du travail montre que l'impact de l'IA dépend avant tout de la nature des tâches et des compétences mobilisées. Loin d'un effacement uniforme de l'humain, c'est une recombinaison des activités professionnelles qui s'opèrerait, fondée sur des dynamiques de complémentarité entre intelligences humaine et artificielle. Cette transition conduit à reconsidérer ce qui, dans le travail, résiste à l'automatisation : le jugement, l'interaction, l'adaptation etc. Ces travaux théoriques alimentent les représentations d'un avenir où les SIA se déploient massivement et rapidement dans les organisations. Or, en une dizaine d'années, si les expérimentations se sont effectivement multipliées, en particulier dans les grandes entreprises et les administrations, la plupart d'entre elles ne sont pas mises en production et déployées.

2.2. Le plafond de verre du déploiement des cas d'usage : l'IA ne passe pas la période d'essai

L'apprentissage automatique nourri de données suscite un foisonnement d'expérimentations : optimisation logistique, maintenance prédictive, aide à la décision, relation client, détection de fraude ou pilotage budgétaire... Autant d'usages capables de remodeler la performance des organisations privées comme des administrations.

Toutefois, l'industrialisation de ces cas d'usage reste marginale : 80 % environ de proofs of concept (PoC) demeurent au stade de prototype. Le Partenariat Mondial pour l'IA⁹ (2021), dans son rapport de 2021, s'interrogeait en sens : « pourquoi les PoC réussis ne sont presque jamais déployés ? ». A partir d'une étude sur plus de 100 cas d'usage dans une quinzaine de pays, trois raisons majeures étaient invoquées :

- L'acceptation organisationnelle pointait sur des SIA décorrélés des activités des travailleurs, des SIA utiles mais non utilisables, ou, à l'inverse, utilisables mais inutiles.
- L'acceptation sociale décrivait des valeurs professionnelles de front, en particulier la responsabilité, qui motivaient une résistance éthique à des applications par ailleurs jugées utiles, utilisables et performantes.

⁹ Le Partenariat mondial sur l'intelligence artificielle (PMIA) est une initiative lancée en 2020 pour promouvoir une IA responsable, éthique et centrée sur l'humain, en partenariat avec l'OCDE et impliquant 29 pays. Il anime quatre groupes d'experts, dont le « Future of Work Working Group » et « Innovation & Commercialisation ». <https://www.oecd.org/en/about/programmes/global-partnership-on-artificial-intelligence.html>

- L'acceptation pratique montrait comment la suppression ou la transformation de gestes professionnelles emblématiques, et la création de nouvelles tâches, souvent orientées vers le fonctionnement de l'IA et peu valorisées, irritent les travailleurs.

Dans la continuité, le LaborIA¹⁰, dans son étude longitudinale de plusieurs cas d'usage pointe sur « un conflit de rationalité » entre une logique gestionnaire de rationalisation du travail à l'œuvre dans le cas d'usage, et le travail réel. Autrement dit, les SIA se rapportent encore trop largement au travail prescrit, ce qui génère d'importants décalages entre les recommandations, notifications, alertes, etc. de l'IA et les situations réellement vécues au travail. La rationalité gestionnaire promeut l'IA pour optimiser performance et productivité, positionnant décideurs et intégrateurs comme moteurs du changement. En face, la rationalité du travail ordinaire exprime les attentes des salariés : reconnaissance, savoir-faire, relations sociales, autonomie et responsabilité. Faute de dialogue, ces logiques s'affrontent, l'IA imposée d'en haut transformant le travail sans concertation. Les projets de SIA pilotés par la direction illustrent cette tension : alors que le management vante l'efficacité, les opérateurs soulignent risques, pertes de sens et redéfinition des compétences, remettant en cause l'idée que l'IA enrichit automatiquement la valeur des salariés.

Le rapport Rand¹¹ met en lumière les causes structurelles et organisationnelles de cet écart entre les promesses technologiques et leur mise en œuvre effective dans les entreprises. Il identifie cinq raisons principales expliquant le faible déploiement de l'IA dans l'industrie :

- Mauvaise communication sur les objectifs à atteindre, entraînant l'optimisation de métriques inappropriées ou un mauvais ancrage dans les processus métier.
- Manque de données de qualité, avec une sous-estimation des ressources nécessaires à leur collecte.
- Priorité donnée à la technologie de pointe plutôt qu'à la résolution de problèmes concrets pour les utilisateurs.
- Infrastructure insuffisante, freinant la gestion des données et le déploiement efficace des modèles.
- Ciblage de problèmes trop complexes, mal adaptés aux capacités actuelles de l'IA.

¹⁰ En réponse à l'une des recommandations du Rapport Villani et en complément de l'implication française dans le PMIA, le ministère chargé du travail et de l'emploi et Inria ont fondé ensemble en 2021 le LaborIA, un laboratoire visant à construire et consolider une vision terrain pour mieux cerner l'intelligence artificielle et ses effets sur le travail, la population active, l'emploi, les compétences et le dialogue social. <https://www.laboria.ai/>

¹¹ [James Ryseff, Brandon F. De Bruhl, Sydne J. Newberry, 2024, "The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed", https://www.rand.org/pubs/research_reports/RRA2680-1.html](https://www.rand.org/pubs/research_reports/RRA2680-1.html)

Les causes de cette difficulté de déploiement ne se localisent pas uniquement dans les entreprises. Dans le champ académique, le rapport pointe le peu d'intérêt de la recherche pour l'intégration réelle de l'IA dans les organisations. Cela s'expliquerait par un désalignement des incitations : un projet est peu valorisé s'il ne génère pas de prestige ou de publications, même s'il présente des bénéfices concrets. Ainsi, la logique de carrière prime souvent sur l'impact réel des travaux. Ces constats soulignent que les blocages ne tiennent pas uniquement à la maturité des technologies, mais relèvent d'enjeux humains, organisationnels et culturels profonds. Sans clarification des finalités, sans alignement entre les besoins métier, les capacités techniques et les logiques d'acteurs, l'IA reste cantonnée à l'état de prototype, bien loin de transformer durablement les pratiques professionnelles et le fonctionnement des organisations.

Dans le même sens, le rapport de BCG « Where's the Value in AI? »¹² met en lumière les facteurs de succès des entreprises qui tirent une réelle valeur de l'intelligence artificielle. Contrairement à celles restées au stade de l'expérimentation, les organisations performantes adoptent une approche structurée basée sur la règle du 70-20-10 : 70 % des ressources sont investies dans les personnes et les processus (formation, gestion du changement, refonte des flux de travail, gouvernance), 20 % dans la technologie et les données (infrastructure, plateformes, gestion), et seulement 10 % dans les algorithmes. L'essentiel de la valeur provient donc des dynamiques humaines et organisationnelles, plus que des choix techniques. BCG souligne que les échecs tiennent souvent à une focalisation excessive sur les solutions technologiques au détriment des conditions d'appropriation, d'alignement des rôles et d'évolution des compétences. À l'inverse, les entreprises qui réussissent construisent en priorité les fondations organisationnelles d'une adoption à grande échelle, condition sine qua non pour générer une valeur significative et durable.

C'est dans ce contexte de faible déploiement des cas d'usage de l'IA, initiatives d'employeurs d'intégrer des SIA dans l'activité des collaborateurs, que se développent, suite à l'avènement de l'IA générative grand public, les pratiques spontanées, initiatives d'employés. À l'inverse des premières, ces pratiques passent rapidement à l'échelle. Opportunité d'expérimentation et de démystification, elles soulèvent simultanément de nombreux enjeux de sécurité, de qualité du travail et de qualité de vie au travail.

¹² Nicolas de Bellefonds, Tauseef Charanya, Marc Roman Franke, Jessica Apotheker, Patrick Forth, Michael Grebe, Amanda Luther, Romain de Laubier, Vladimir Lukic, Mary Martin, Clemens Nopp, and Joe Sassine, (2024), "Where's the value in AI?", [https://www.bcg.com/publications/2024/wheres-value-in-ai?](https://www.bcg.com/publications/2024/wheres-value-in-ai?utm_source=search&utm_medium=cpc&utm_campaign=digital-transformation&utm_description=paid&utm_topic=ai&utm_geo=global&utm_content=dsa_wheres-value-in-ai&gclid=Cj0KCQjwotDBBhCQARISAG5pinOeaJ8lViSjOAzD_UO39_SFljeyha100E5dOV6UI7gX7W-TsNBwIXAaAoeDEALw_wcB)

2.3. L'intelligence artificielle, une initiative des collaborateurs : à chacun son usage

2.3.1. L'IA, initiative discrète des collaborateurs

Le lancement mondial de ChatGPT d'OpenAI, version gratuite d'un LLM a fait basculer l'IA au travail dans une nouvelle dimension que l'on peut qualifier à l'aune de deux dynamiques spécifiques.

Un passage à l'échelle

Tandis que l'IA concernait un très faible nombre d'actifs français avant l'arrivée de l'IA générative, ils seraient aujourd'hui autour de 20% (Beauchene et al, 2024¹³ ; Ifop/Learnthings, 2024¹⁴). Une enquête Ifop Talan (2023) révélait que 71% des français ont entendu parlé de l'IA générative et que 44% des Français qui utilisent les IA génératives les utilisent à la fois dans un cadre personnel et professionnel. . En 2025, le taux de notoriété est mesuré à 86%, avec un pic à 91% chez les moins de 35 ans. Un autre sondage réalisé auprès de quatre grandes écoles estime que 92% des étudiants l'utilisent régulièrement et que 65% font de la présence de l'IA générative un critère majeur du choix de leur future entreprise. En effet, 51% constatent qu'ils auraient du mal à s'en passer¹⁵ (Ifop Talan 2024): « La fracture générationnelle se confirme avec 85% des 18-24 ans qui utilisent désormais les IA génératives (contre 68% en 2024), 63% des 25-34 ans (contre 47% en 2024), mais seulement 31% des 35 ans et plus ».

Un changement de trajectoire

D'initiatives exclusives d'employeurs, l'IA est largement devenue une initiative d'employés. En 2023, une étude de Salesforce¹⁶ révèle que si seulement 18% des employés français utilisent l'IA générative au travail, plus de la moitié le font en dehors d'un cadre défini par leur employeur (58%). S'ils se disent plus productifs (68%) et plus impliqués (63%), 71% ont déjà présenté des travaux effectués par l'IA générative comme les leurs et 27% envisageraient d'exagérer leurs compétences en matière d'IA générative pour obtenir une opportunité d'emploi. Enfin, 87% des employés français déclarent que leur employeur n'a pas de politique clairement définie quant à l'utilisation de l'IA générative en entreprise et 73 % déplorent l'existence de freins à l'utilisation de l'IA générative dans leur entreprise.

¹³ <https://www.bcg.com/press/26june2024-ia-generative-au-travail-amie-ou-ennemie-des-salaries>

¹⁴ <https://www.learnthings.fr/sondage-ifop-intelligence-artificielle-statistique/>

¹⁵ Baromètre 2024 IFOP pour Talan "les français et les IA génératives", 2024, <https://www.talan.com/global/fr/barometre-2024-ifop-pour-talan-les-francais-et-les-ia-generatives>

¹⁶ <https://www.salesforce.com/fr/company/news-press/press-releases/2023/11/231127/>

En 2023, le baromètre Ifop Talan souligne que 68% des Français qui utilisent les IA génératives en entreprise le cachent à leur supérieur hiérarchique. Un an plus tard, un sondage Ifop/Learnthings, évalué à 55% les utilisateurs de l'IA au travail n'en auraient pas informé leur hiérarchie. Cette même enquête précise que seul 10% des répondants auraient suivi une formation dans le cadre professionnel et que la moitié estiment cette formation non pertinente. Les deux tiers des salariés (63%) indiquent ne pas avoir été formés et, surtout, ne pas vouloir l'être quand 27% y aspirent au contraire. Le baromètre 2025 d'Ifop Talan indique un taux de formation de 15%. Il montre également une diminution sensible des pratiques clandestines en trois ans :

Avez-vous informé votre supérieur hiérarchique que vous utilisez les IA génératives ?

	MAI 2023 (%)	MARS 2024 (%)	MARS 2025 (%)
OUI	25	36	39
NON	52	36	37
(Vous n'avez pas de supérieur hiérarchique)	23	28	24
TOTAL	100	100	100

Source : Ifop pour Talan, 2025

En un an, ChatGPT serait ainsi devenu la principale application en Shadow IT. Selon un rapport de Productiv (2023)¹⁷, qui a analysé plus de 100 milliards de points de données d'usage couvrant près de 100 millions de licences SaaS entre 2021 et 2023, ChatGPT a en effet bondi directement à la première place des applications non autorisées (« shadow IT »), détrônant LinkedIn Business Solutions, numéro 1 les deux années précédentes. En 2024, Productiv note toutefois que l'usage global des licences ChatGPT officiellement approuvées reste faible : malgré l'intérêt manifeste des employés démontré par le recours au Shadow IT, l'IT peine à obtenir l'adoption formelle des licences.¹⁸

¹⁷ <https://productiv.com/blog/2023-state-of-saas-series-while-companies-make-progress-cutting-costs-previous-investments-and-growth-of-shadow-apps-like-chatgpt-challenge-efforts-to-manage-saas-spend/>

¹⁸ <https://productiv.com/state-of-saas/2024-saas-trends-usage/>

Le Shadow IT, ou informatique fantôme, désigne l'ensemble des systèmes et outils numériques créés et utilisés dans une organisation sans validation de la direction des systèmes d'information. Perçu par ses partisans comme un vivier d'innovation – les solutions élaborées pouvant servir de prototypes à des applications officielles – il est jugé par ses détracteurs comme porteur de risques majeurs : non-respect des bonnes pratiques, failles de sécurité, absence de documentation et fragilité globale. Si le Shadow IT catalyse l'expérimentation d'idées neuves, il peut aussi freiner l'introduction de technologies plus sûres.

Dans la continuité de cette définition, le Shadow AI rend compte de l'utilisation non autorisée d'outils ou d'applications d'intelligence artificielle par les employés ou les utilisateurs finaux sans l'approbation ou la supervision formelle du département informatique (IT). Plus précisément, il s'agit le plus souvent de l'utilisation non approuvée ou ponctuelle d'IA générative au sein d'une organisation qui échappe à la gouvernance informatique. Cela inclut l'accès par les employés à des outils d'IA générative comme ChatGPT pour effectuer des tâches telles que la rédaction de textes, la création d'images ou l'écriture de code. Le Shadow AI se distingue du Shadow IT en se concentrant spécifiquement sur les outils, les plateformes et les cas d'utilisation liés à l'IA, tandis que le Shadow IT concerne tout déploiement non autorisé de logiciels, de matériel ou de technologies de l'information. Pour les départements informatiques, le Shadow AI représente un enjeu critique en matière de gouvernance, car il nécessite de décider quelle utilisation de l'IA autoriser ou restreindre afin de soutenir les initiatives tout en assurant la sécurité de l'organisation.

2.3.2. Le paradigme du bricolage

Le paradigme du bricolage, art de faire avec ce que l'on a, est une logique d'action et de connaissance fondée sur l'exploitation inventive des ressources disponibles, qu'elles soient matérielles, conceptuelles ou sociales. Dans *La Pensée sauvage* (1962), Lévi-Strauss y voit l'« exploitation créative des ressources disponibles » : le bricoleur puise dans « l'univers clos » des matériaux sous la main ; il recycle les résidus d'expériences antérieures et « s'arrange avec les moyens du bord ».

« Le bricoleur est apte à exécuter un grand nombre de tâches diversifiées ; mais, à la différence de l'ingénieur, il ne subordonne pas chacune d'elles à l'obtention de matières premières et d'outils, conçus et procurés à la mesure de son projet : son univers instrumental est clos, et la règle de son jeu est de toujours s'arranger avec les “moyens du bord” ». ¹⁹

¹⁹ Claude Lévi-Strauss (1962), *La Pensée sauvage*, Paris, Ed. Plon, 1960, p. 27.

Contrairement à la science idéale, qui construirait ses outils de toutes pièces, le bricolage avance par ajustements, improvisations et alliances contingentes. D'où l'idée centrale : la création naît moins d'un plan parfaitement tracé que d'une série d'opportunités saisies.

Le bricolage peut être rapporté à trois notions clés :

- Le rhizome : empruntée à Deleuze et Guattari²⁰, l'image du rhizome aide à visualiser un réseau sans centre ni hiérarchie fixe. Les projets scientifiques ou techniques évoluent en tissant des connexions latérales entre disciplines, instruments et acteurs ; une idée surgit ici, une contrainte apparaît là, et l'ensemble se reconfigure en permanence.
- Le braconnage : Michel de Certeau décrit la façon dont les usagers s'approprient les dispositifs socio-techniques pour en faire autre chose que ce qui était prévu²¹. Le « faire-avec » décrit par Michel de Certeau à travers le braconnage est « caché » et « se dissémine dans les régions définies et occupées par les systèmes de la “production” ». Les pratiques bricolantes et tactiques, le « faire-avec » que Michel de Certeau cherche est donc « caché », il « se dissémine dans les régions définies et occupées par les systèmes de la “production” (télévisée, urbanistique, commerciale, etc.) » (xxxvii). En tant qu'elle module les pratiques de consommation, la tactique est « rusée, [...] dispersée, mais elle s'insinue partout, silencieuse, quasi invisible puisqu'elle ne se signale pas avec des produits propres mais en manières d'employer les produits imposés par un ordre économique dominant » (ibid. : xxxvii).
- La tactique : toujours chez de Certeau, la tactique consiste à saisir les occasions dans un terrain que l'on ne contrôle pas complètement. La tactique est un « calcul qui ne peut pas compter sur un propre, ni donc sur une frontière qui distingue l'autre comme une totalité visible » (p.14). La tactique fait partie intégrante des « arts de faire » du quotidien.

S'appuyer sur le paradigme du bricolage, c'est admettre, à la suite de Lévi-Strauss, que l'innovation progresse souvent « par morceaux ». On avance pas à pas, on combine ce qui est disponible, et l'on transforme les contraintes en occasions de créer. Cette vision n'oppose pas amateurisme et rigueur ; elle rappelle simplement que, même dans la science la plus formalisée, une part essentielle du travail reste faite d'ingéniosité pratique et de ré-agencements successifs. Aujourd'hui, cette logique se retrouve dans les usages informels d'intelligence artificielle par les collaborateurs.

²⁰ Deleuze G., Guattari F., 1980, *Capitalisme et schizophrénie 2. Mille plateaux*, Paris, Éd. de Minuit.

²¹ De Certeau M. (1990 [1980]). *L'invention du quotidien t.1. Arts de faire*, Paris, Gallimard.

2.3.3. De surprenants assistants

Loin des discours abstraits sur la productivité ou des craintes massives de substitution, les témoignages de terrain révèlent une autre réalité : celle d'une appropriation pragmatique et inventive de l'IA générative par les professionnels eux-mêmes. En 2024, la journaliste Apolline Guillot (Philonomist) a recueilli une série de récits éclairants, où l'IA agit comme un assistant discret du travail réel, tantôt traducteur culturel, mentor méthodologique, copilote graphique ou moteur de réflexivité. À ces pratiques souvent informelles s'ajoute le cas du Dr David Morquin, médecin hospitalier, qui expérimente l'usage de l'IA pour personnaliser les courriers médicaux, illustrant un usage à la fois empathique et structuré de ces outils dans le champ du soin. Ces récits donnent à voir des usages incarnés, bricolés, profondément situés, qui transforment en profondeur les gestes, les rythmes et les arbitrages professionnels.



Les débuts : J'ai commencé à utiliser ChatGPT quand c'est devenu grand public. Je ne suis pas à l'affût de ce genre d'outils, j'utilise juste des dictionnaires de synonymes et des correcteurs, vu que le français n'est pas ma langue maternelle. Mais avec ChatGPT, très rapidement, c'était le grand amour.

L'usage : Comme je suis commerciale, j'écris énormément de mails de prospection, de relance, à des interlocuteurs variés. Et même si tu écris bien, sans faute, en France, il y a différents types de langages pour différents types d'interlocuteurs. On peut perdre des lead, en relançant mal une personne ! Par exemple, quand j'ai commencé, on m'a déjà fait remarquer sèchement que mes relances étaient « cavalières ». Dans ma culture, c'est comme ça qu'on demande les choses, alors qu'en France c'est vu comme insistant. ChatGPT m'a aidé à apprivoiser cette différence culturelle entre l'Italie et la France. Concrètement, j'écris le mail de manière directe, en proposant par exemple un rendez-vous commercial, puis je demande à ChatGPT de transformer mon mail de manière « corporate ».

Mon utilisation de l'IA va au-delà de la simple consigne-exécution. Je suis un peu comme un prof qui corrige une copie d'un élève : je donne une consigne, je repasse, je corrige. Ça ne prend pas forcément beaucoup de temps : j'ai découvert un truc qu'il suffit de cliquer sur reformuler, la deuxième proposition est meilleure — je ne sais pas pourquoi, d'ailleurs.

²² Trois témoignages ainsi que notre analyse sont accessibles en ligne : <https://www.philonomist.com/fr/article/comment-jai-dresse-mon-ia>

Le grand avantage de ce genre d'outils, c'est qu'il me suggère des modalités de communication qu'en tant qu'étrangère, je n'aurais pas eues. Des codes implicites qui m'auraient demandé beaucoup plus de temps à acquérir par moi-même!

Un des prompts les plus utilisés en 2023, c'est « dis-le de manière voilée ». Pour un étranger, dire quelque chose sans avoir l'air de le dire, c'est compliqué ; l'IA m'apprend un peu ça — et moi je la corrige quand elle en fait des tonnes. C'est un travail d'enrichissement mutuel, de correction réciproque.

Limites ? Et si ChatGPT part dans la mauvaise direction, que ça bloque, c'est généralement moi qui n'ai pas été assez claire ou la tâche a été trop complexe, pas assez décomposée. C'est le signe que je dois reformuler ma pensée. En fait, j'ai compris que ce qu'on donne en entrée est important : si ce que je lui fournis ne sont pas structurées au départ, impossible d'avoir quelque chose de bien.



Les débuts : J'ai été diplômée d'expertise comptable en 1991, et le sujet de mon mémoire était la conception d'un système expert d'évaluation d'entreprise... Bref, j'étais acculturée au machine learning, j'étais loin d'en avoir peur.

L'usage : Pour la collecte d'information, très chronophage, ChatGPT me sert de moteur de recherche augmenté. Par exemple, pour l'évaluation d'entreprise : c'est énormément d'analyse de données, de compréhension des modèles d'affaires des entreprises, de leur marché, de la concurrence, etc. Une autre utilisation que j'en fais, c'est pour comparer des données d'entreprises de plusieurs pays, classées en fonction de codes différents, propres à chaque pays : l'IA fait office de « traducteur » des normes d'un pays à l'autre, d'un système comptable à l'autre. Enfin, je l'utilise aussi pour m'apprendre des choses. Exemple : j'avais un graphique à produire sur Excel, je ne savais pas comment le faire, et ChatGPT m'a donné en 5 minutes un mode d'emploi.

ChatGPT, c'est un peu le professeur Nimbus, il sait tout et me dépanne. En plus, quand je lui demande de m'aider, je m'adresse à lui en langage naturel, alors que sur Google je mettrais des mots-clés. Je ne suis pas naïve : je sais bien que lorsqu'il ne connaît pas la réponse il est capable d'inventer. Mais dans le métier de comptable, par définition, on est sceptiques par déformation professionnelle, donc je ne me fie pas immédiatement à ce qui est donné.

Je n'ai pas de crainte sur mon remplacement par la machine, parce que dans la comptabilité il y a toute une partie qui relève du « jugement professionnel ». Imaginons que je sois en train de revoir un rapport qui a été fait par des collaborateurs qui ont déroulé une certaine mécanique : pour comprendre les choix de l'entreprise, je dois me mettre à la place du chef d'entreprise. Cette sensibilité, je ne vois pas comment ChatGPT peut l'avoir.

Limites ? Pour moi, demain l'IA sera capable de « prémâcher » notre travail. On devra davantage s'interroger sur la qualité de la donnée et la qualité des chiffres qui sortent, ne pas se contenter de ça : ça veut dire développer une capacité de réflexion. Un bon auditeur, c'est quelqu'un qui a su traiter la donnée avant de réfléchir après à sa signification. Notre valeur ajoutée en tant que comptables, est d'alerter les clients en regardant la cohérence des chiffres, les anomalies, retourner à l'information source.

Mais est-ce que les jeunes comptables, demain, qui n'auront manipulé que des chiffres pré-traités par de l'iA, auront ce savoir-faire ? Nous, les anciens, on est toujours capable d'aller vérifier dans l'information brute. Est-ce que les jeunes, qui n'auront jamais eu à traiter de l'information brute, pourront faire la même chose? En fait, on le voit déjà : à mesure que les outils de production de chiffre s'améliorent, le niveau global de réflexion baisse.

Dans nos métiers, un outil qui vous assure une fiabilité arithmétique amène à sous-développer son bon sens économique.



Les débuts : Avant Covid, je regardais ça de loin. Il y avait Disco Diffusion, qui permettait de générer des images à base de prompts, mais c'était lent et le rendu atroce. Et puis, il y a trois ou quatre ans, je me suis retrouvé par hasard, sur les forums de Midjourney. Et là, c'est le choc : les IA parlent mieux que moi, codent mieux que moi, font moins de fautes d'orthographe que moi, chantent juste, dessinent mieux que moi.

Ça fait un coup. J'ai été vexé, et mes étudiantes aussi. On se rend compte que certaines réalisations qu'on pensait humaines, artistiques, réalisées avec talent...peuvent s'automatiser comme le travail d'une caissière ou de manutentionnaire. J'ai même parfois des réactions très violentes ! Un étudiant m'a avoué qu'il était en colère : « On m'avait dit que dans les professions intellectuelles, on était irremplaçables! ». Je lui réponds : « Ben voilà, raté. ».

J'utilise l'IA comme une sorte d'assistant infatigable, qui prendrait beaucoup de drogues...et aurait une certaine créativité ! Car pour moi ça ne fait aucun : Midjourney est créatif. Les IA, comme nous, utilisent le passé pour créer le futur, il n'y a pas de différence fondamentale.

Limites ? Un bon graphiste est en dialogue avec l'IA, fait des aller-retours sans cesse. On dessine, on envoie l'image, la machine nous propose quelque chose, on dessine par dessus, puis la machine habille, complète, etc. Mais c'est un dialogue un peu particulier... avec un extraterrestre. La « chose » ne réfléchit pas comme nous; elle ont été nourries de milliards de données que nous n'avons pas vues, mais elle garde la fraîcheur d'un enfant de deux ans.

C'est ça qui la rend passionnante : elle propose des choses avec une naïveté incroyable, car elle ne sait pas que ce n'est pas possible. Nous, humains, nous nous fermons des pistes parce que nous suivons notre intuition, nos habitudes, nos goûts esthétiques. Nous disons d'emblée : « Ça ne marchera pas ». Mais tout ce qui est possible a été impossible à un certain moment, non?

Je pousse l'IA dans ses retranchements, je la fais halluciner, puis ensuite je fais le tri. Je vois d'ailleurs que les professionnels vont souvent chercher des versions anciennes de ChatGPT, moins réalistes, plus loufoques, pour avoir plus de marge de manœuvre. Si on ne l'utilise que pour faire des images ultra-réalistes, ça n'a pas grand intérêt.



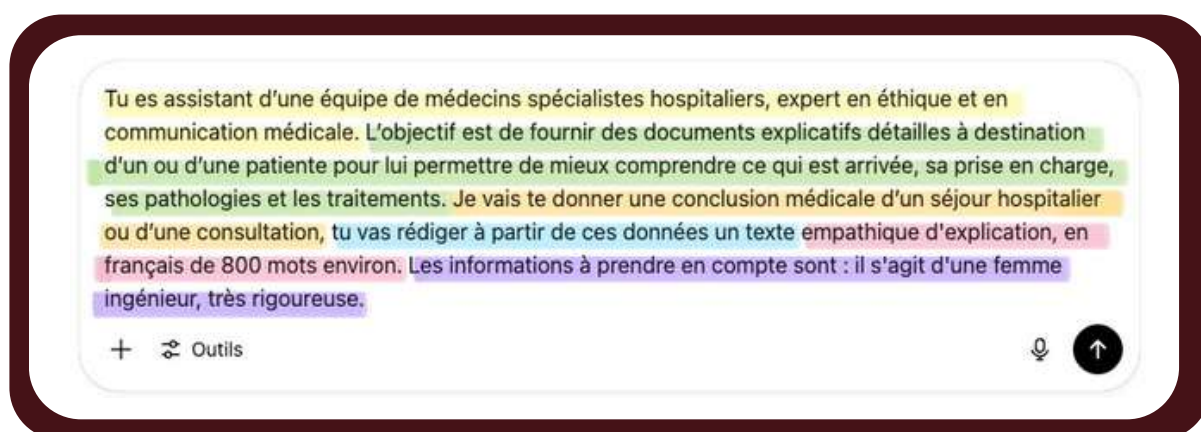
« Je travaille depuis trois ans dans une banque d'affaires internationale internationale. Et il y a eu un « avant » et un « après » ChatGPT. Évidemment, il faut faire le travail de relecture, mais pour traiter une grande quantité d'informations ou une question un peu vague, maîtriser rapidement un sujet, c'est extraordinaire. Et très vite, je me suis mis à challenger : je pose une première question, il répond de manière vague, puis je lui demande de continuer à améliorer sa réponse jusqu'à ce qu'il ait abouti à ce que je voulais précisément.

En fait, l'IA va « senioriser » mon métier. Aujourd'hui, on a besoin de stagiaires, d'analystes, d'associés pour pouvoir faire des slides, pour pouvoir chercher de l'information pour des clients, etc. Si demain on a une IA ultra performante, entraînée sur nos propres données, qui sera capable de faire des présentations, de rédiger des mails automatiques, de gérer la logistique... je pense que tous les grades un peu juniors (et je m'inclus dedans!) vont être amenés à disparaître progressivement. Après, c'est aussi propre à notre métier... C'est du 80/20 : on passe la plupart de notre temps à faire des slides, et la majorité de ce qui se trouve dessus n'est pas réellement intéressant.

En vérité, les gens paient une banque d'affaires pour son réseau, pour son rôle d'intermédiaire. C'est ça, le nerf de la guerre : trouver les bonnes personnes, savoir les lire. Quelqu'un qui met des millions d'euros sur la table obéit à des motifs ultra rationnels, certes... mais là où il y a des grosses sommes il y a aussi de l'affects, des émotions qu'il faut être capable d'anticiper, de gérer.

Après, rien qu'aujourd'hui ça a modifié considérablement mon quotidien. Il faut savoir qu'en banque d'affaires, on doit devenir expert rapidement sur tous les sujets qui concernent les entreprises qu'on conseille. C'est très chronophage, ce sont des métiers où traditionnellement on finit très tard. Aujourd'hui, c'est moins le cas : grâce à l'IA, je fais du sport tous les jours entre midi et deux. »

A ces expériences « ordinaires » s'ajoute celle d'un médecin hospitalier en maladie infectieuse et spécialiste des systèmes de gestion de l'information dans le milieu hospitalier au CHU de Montpellier, David Morquin, qui a élaboré empiriquement un prompt dédié à la rédaction de compte rendu de séjour hospitalier pour ses patients.



■ Rôle ■ Objectif ■ Action ■ Tonalité, résultats ■ Contraintes

Proposition de ChatGPT :

“

1. Mandibulectomie interruptive avec reconstruction par lambeau de fibula :

Vous avez subi une intervention chirurgicale complexe sur la mâchoire, nommée mandibulectomie interruptive. Cette opération a été nécessaire pour traiter une ostéoradionécrose de la mandibule, une complication rare mais grave qui se caractérise par la mort du tissu osseux de la mâchoire, souvent survenue après une radiothérapie. »

”

Remplaçant les contraintes par « il s'agit d'un adolescent fan de rugby », il obtient un courrier factuellement identique mais avec une autre tonalité :

“

« Salut Champion,

« J'espère que tu te sens un peu mieux et que tu es prêt à récupérer ta forme pour revenir sur le terrain de rugby dès que possible. On sait que tu as traversé une période difficile, et c'est pourquoi on veut t'expliquer tout ce qui s'est passé pendant ton séjour à l'hôpital, comme si c'était un « replay » d'un match important. »

”

Tous ces usages se situent hors radar (à l'exception de ceux du Dr. Morquin qui suit un protocole expérimental dans un cadre sécurisé) : licences non déclarées, prompts privés, itérations. L'innovation s'effectue « par le bas », dans les interstices où la gouvernance IT n'a pas (encore) pris pied, reproduisant le phénomène classique de Shadow IT mais à l'échelle de l'IA générative. À défaut de cadre officiel, chaque professionnel « bricole » ainsi sa propre chaîne sociotechnique : Elisabetta combine mails bruts + reformulation, Maud entrelace recherche, traduction de normes comptables et micro-formation, Étienne dialogue en boucle son “extraterrestre”, Pierre itère ses requêtes jusqu'au livrable. Tous décrivent un échange plus qu'une simple exécution ; l'agent conversationnel reflète la pensée, oblige à clarifier la demande, fractionner la tâche.

L'outil agit comme « miroir réflexif » : il met en évidence les zones floues du raisonnement et renvoie l'utilisateur à son propre cadrage du problème. L'IA compresse la phase routinière (collecte, rédaction, mise en forme) ; le temps ainsi libéré est réinvesti en activités personnelles (sport pour Pierre) ou en créativité (expérimentations d'Étienne). Une réarticulation vie pro/vie perso est exprimée, mais aussi une intensification de l'injonction à produire toujours plus vite. Enfin, Étienne exploite l'IA pour sortir des sentiers battus ; inversement, Elisabetta vise la conformité culturelle de ses démarches. La même technologie nourrit deux dynamiques opposées : uniformisation des échanges, « s'emparer des codes », et élargissement du champ des possibles, « sortir des codes ». **L'IA devient un outil à composer plutôt qu'un service packagé, illustrant la logique de « bricolage» où savoir « prompter » et vérifier devient une compétence-clef, marqueur d'un capital technique informel.** Le cas du Dr Morquin illustre qu'un usage maîtrisé et encadré de l'IA peut conjuguer précision clinique et humanité, en adaptant le ton des documents médicaux à la sensibilité des patients. Il montre qu'un prompt bien conçu, dans un cadre sécurisé, peut transformer un compte rendu standardisé en message empathique, sans perte d'exactitude. Mais ces pratiques soulèvent aussi des inquiétudes profondes. Étienne exprime un trouble identitaire face à une IA jugée plus performante que lui sur des compétences qu'il croyait humaines et protégées. Maud s'interroge sur la perte de savoir-faire critique chez les jeunes professionnels qui n'auraient jamais manipulé des données brutes. Pierre anticipe, avec lucidité, la disparition des fonctions juniors dans son secteur. Ces récits témoignent d'une ambivalence forte : pour ces bricoleurs, l'IA est autant un levier d'efficacité qu'un facteur de déstabilisation cognitive, affective et professionnelle.

2.3.4. La valeur recherchée : s'augmenter plutôt que déléguer

Plusieurs études convergent pour identifier deux logiques distinctes dans les usages de l'IA générative : une logique de délégation et une logique d'augmentation. Un sondage GetApp (2023²³) révèle que les gains les plus attendus concernent la productivité (36 %), l'automatisation des tâches répétitives (29 %) ou leur organisation (23 %), selon un modèle où l'IA prend en charge des actions standardisées pour libérer du temps humain. Mais des bénéfices plus qualitatifs émergent autour de la coopération humain-machine sur des tâches cognitives non routinières : créativité (32 %), analyse (29 %), résolution de problèmes (27 %), innovation (25 %) et inspiration (21 %).

²³ Getapp (2023), « IA générative: comment les entreprises devraient-elles aborder cette révolution ? ». <https://www.getapp.fr/blog/3801/ia-generative-guide-pour-pme>

Le sondage du Capgemini Research Institute²⁴ confirme cette dualité : si la rédaction d'e-mails (52 %), la révision de textes (20 %) ou la traduction (16 %) relèvent d'une automatisation ciblée, des usages plus réflexifs apparaissent, comme le « brainstorming » (28 %), la veille technologique (28 %), ou la conception de supports pédagogiques (20 %). Ces résultats montrent que l'IA n'est pas seulement un outil d'exécution, mais aussi un **partenaire intellectuel capable de stimuler la pensée critique et l'apprentissage, en catalysant de nouvelles formes de coopération cognitive.**

En février 2025, une analyse de plusieurs millions de conversations sur Claude.ai menée par les équipes de recherche d'Anthropic montre que l'utilisation se concentre principalement dans les professions informatiques et le développement de logiciels (37,2 %), suivi des professions artistiques, de conception, de divertissement, de sport et de médias (10,3 %)²⁵. L'analyse des compétences représentées dans les conversations montre que les compétences cognitives comme la compréhension en lecture, l'écriture et la pensée critique sont très présentes, tandis que les compétences physiques (comme l'installation, la maintenance d'équipement) et les compétences managériales (comme la négociation) montrent une présence minimale. L'analyse des conversations sur Claude.ai a classifié les modes d'interaction en cinq modèles de collaboration, regroupés en comportements d'automatisation, par exemple, l'exécution d'une demande avec une implication humaine minimale ou d'augmentation, par exemple l'apprentissage ou l'itération sur une sortie.

- Directive (comportement d'automatisation) : L'humain délègue complètement l'exécution d'une tâche à l'IA avec une interaction minimale.
- Boucle de rétroaction (comportement d'automatisation) : L'achèvement de la tâche est guidé par le retour d'information environnemental fourni par l'humain. L'utilisateur relaie à plusieurs reprises l'erreur qu'il a reçue au modèle. Cet usage est largement lié au codage et au débogage.
- Itération de tâche (comportement d'augmentation) : Il s'agit d'un processus de raffinement collaboratif « Collaborative refinement process » où l'humain et l'IA s'engagent dans un dialogue itératif pour accomplir une tâche, l'humain affinant les sorties de l'IA.

²⁴ Capgemini Research Institute. (2023). Why consumers love generative AI. Report.

<https://prod.ucwe.capgemini.com/be-en/wp-content/uploads/sites/14/2023/06/Final-Web-Version-Report-Creative-Gen-AI.pdf>

²⁵ Cette étude ne précise pas si ces conversations relèvent d'initiatives personnelles ou d'un cadre institutionnel. Nous l'évoquons surtout pour mettre en avant le paradigme du bricolage que les pratiques décrites élaborent. En outre, dans la section 2 de ce rapport, nous allons proposer une définition plus étendue du « shadow AI ».

Handa K. et al. (2025), "Which Economic Tasks are Performed with AI? Evidence from Millions of Claude Conversations", <https://arxiv.org/abs/2503.04761>.

- Apprentissage (comportement d'augmentation) : L'humain cherche l'acquisition et la compréhension de connaissances plutôt que l'achèvement direct d'une tâche.
- Validation (comportement d'augmentation) : L'humain utilise l'IA pour vérifier ou valider son propre travail, usage surtout constaté dans la traduction.

L'article montre que les interactions suggérant une augmentation étaient légèrement plus fréquentes (57%) que celles suggérant une automatisation (43%). La plupart des professions étudiées présentaient un mélange d'automatisation et d'augmentation à travers les tâches.

Enfin, une étude France Travail publiée en janvier 2025 montre une forte utilisation de l'IA par les demandeurs d'emploi²⁶ : « les demandeurs d'emploi se sont emparés massivement de l'IA générative dans le cadre de leurs recherches d'emploi et de la préparation de leur candidature ». 77 % déclarent avoir déjà utilisé l'IA pour optimiser leurs recherches d'emploi, indépendamment du niveau d'éducation, 62 % ont ressenti un impact positif de l'IA dans la recherche d'emploi et 60% disent gagner en efficacité sur des critères de gain de temps et d'amélioration de leurs candidatures. Le niveau d'éducation compte peu dans l'utilisation de l'IA, les personnes ayant un niveau inférieur au bac comme les Bac + 2 l'utilisant de façon équivalente. L'utilisation par les femmes (79 %) est légèrement supérieure à celle des hommes (74 %). Les moins de 25 ans l'intègrent plus fortement (83 %) dans leur stratégie de recherche par rapport aux plus de 50 ans (69 %). 40 % utilisent l'IA pour rédiger des C.V. plus percutants et des lettres de motivation personnalisées. Un tiers l'exploite pour accéder à davantage d'offres d'emploi via les recommandations automatiques. Les robots conversationnels interactifs sont populaires (31 %) et servent à obtenir des conseils sur les entretiens ou à résoudre des questions administratives. Une partie plus modeste, surtout chez les personnes peu ou pas diplômées, utilise l'IA pour simuler des entretiens d'embauche.

Ainsi, la convergence de ces travaux éclaire un constat sans équivoque : l'intelligence artificielle générative s'inscrit désormais au cœur d'une double dynamique qui reconfigure la création de valeur et l'organisation du travail. D'une part, son potentiel d'automatisation délégitime, rationaliser les chaînes de tâches répétitives et libérer du capital-temps ; d'autre part, son rôle d'auxiliaire cognitif pour le spectre des capacités humaines, catalysant la créativité individuelle, la résolution de problèmes complexes et l'apprentissage continu, y compris chez les demandeurs d'emploi. Entre ces deux pôles, les utilisateurs bricolent, composent, assemblent des hybridations inédites où le gain de productivité se conjugue avec la recherche d'élévation des compétences.

²⁶ <https://www.francetravail.org/accueil/actualites/infographies/2025/intelligence-artificielle-et-recherche-d%27emploi.html?type=article>

2.4. Des initiatives qui inquiètent et menacent organisations et “bricoleurs”

2.4.1. Des usages profanes qui mènent à des mesusages

L'usage croissant de l'IA générative dans la recherche d'informations soulève des enjeux importants de fiabilité et de compréhension de la nature même de ces outils. Dans une affaire juridique rapportée en 2023, un avocat, Maître Schwartz, a cité dans son mémoire de défense la jurisprudence « Varghese vs. China Southern Airlines Co. Ltd. », laquelle s'est révélée, selon le juge, partiellement fictive. Au total, « six des cas mentionnés se révèlent être des décisions judiciaires imaginaires ». L'avocat reconnaît son erreur en déclarant : « Je n'étais pas au courant de la possibilité que le contenu de ChatGPT soit erroné ». Cet incident illustre un phénomène plus large : une partie significative des utilisateurs recourent à l'IA générative pour rechercher de l'information, sans nécessairement comprendre ses limites épistémiques.

Le baromètre Ifop Talan 2025 révèle que 34 % des répondants utilisent principalement l'IA générative à des fins de recherche, une hausse de 10 % par rapport à 2024. Comme le souligne l'institut, « cela conforte l'idée que les moteurs de recherche traditionnels sont de plus en plus concurrencés par certaines IA génératives, notamment ChatGPT ». Or, cette assimilation est problématique : contrairement à un moteur de recherche qui localise et fournit des contenus existants, l'IA générative produit des contenus originaux, dont l'exactitude factuelle ne peut être, par design, garantie...

Pour mieux évaluer la capacité des modèles à produire des réponses factuelles, OpenAI a publié en novembre 2024 le benchmark SimpleQA, composé de 4 326 questions factuelles couvrant des domaines variés (sciences, histoire, géographie, etc.)²⁷. Les réponses sont évaluées selon trois catégories : correctes (la réponse correspond à la réponse de référence), incorrectes (elle la contredit), ou non tentées (la réponse n'est ni donnée ni contredite). Les résultats obtenus illustrent les limites actuelles des modèles : le modèle o1-preview obtient 47 % de réponses correctes, tandis que ChatGPT 4.0 mini n'en atteint que 8,6 %. Ces résultats confirment les limites actuelles de l'IA générative en termes de factualité que le problème des hallucinations demeure central :

²⁷ Wei J. et al., (2024), “Measuring short-form factuality in large language models”.



Les modèles d'IA les plus avancés produisent parfois des réponses fausses ou non étayées par des preuves, un problème appelé "hallucinations". Ces hallucinations constituent l'un des principaux obstacles à l'adoption plus large des formes générales d'intelligence artificielle telles que les grands modèles de langage.



Le terme d'hallucination, aujourd'hui largement diffusé, a été introduit dès 2018 par les chercheurs de Google DeepMind pour désigner des réponses « hautement pathologiques » générées par des systèmes de traduction neuronale. En 2021, Meta constate également que son modèle BlenderBot 2 produit des « déclarations confiantes qui ne sont pas vraies ». Médicalement, l'hallucination est une perception sans objet ; de même, l'IA générative opère dans un univers asémantique²⁸, où les mots ne renvoient pas à une réalité, le signifiant n'est pas relié au signifié, le signe à la chose qu'il désigne. Les hallucinations ne sont donc pas des erreurs corrigibles, mais relèvent du fonctionnement même de l'IA générative, que l'on peut seulement réduire. Paradoxalement, ce qui permet à l'IA de s'exprimer librement constitue aussi son principal point faible. Or, l'enquête Ifop Talan 2024 indique que 44 % des Français reprennent les résultats générés par l'IA sans les modifier, et que 52 % des étudiants déclarent que l'IA influence leurs choix.

Dans ce contexte, le développement de l'esprit critique apparaît comme un enjeu essentiel. La chercheuse Elena Pasquinelli et ses collègues le définissent comme : « l'ensemble des capacités et des critères qui permettent d'évaluer la qualité épistémique des informations disponibles et de doser de façon conséquente notre confiance en ces informations, en vue de prendre une décision, de se forger une opinion, d'accepter ou de rejeter une affirmation à bon escient »²⁹. Leurs travaux montrent que ce savoir-faire cognitif est rarement enseigné de manière explicite à l'école. Bien qu'il soit jugé prioritaire par les enseignants, ceux-ci considèrent souvent que la méthode scientifique suffit à développer une pensée critique généralisable, postulat que ne partagent pas les auteurs, a fortiori quand de plus en plus d'individus s'informent principalement sur Internet et les réseaux sociaux via des algorithmes de recommandation.³⁰

²⁸ Grinbaum, A., 2023, Parole de machines, Paris, humenSciences.

²⁹ « Définir et éduquer à l'esprit critique », 2020.

³⁰ Sur les algorithmes de recommandation, voir Fourquet J.-L. et Lê Nguyễn (2024), La dictature des algorithmes, Paris, Tallandier.

Dans le monde professionnel, cette compétence est également ambivalente. En 2020, le Future of Jobs Report du Forum Économique Mondial classait la pensée critique parmi les compétences émergentes les plus recherchées à l'horizon 2025. Pourtant, dans l'édition 2025 de ce même rapport, elle a disparu du classement³¹. Les compétences cognitives mises en avant sont désormais la pensée analytique, la créativité, le multilinguisme, les fondamentaux (lecture, écriture, mathématiques) et la pensée systémique.

Au-delà des rapports prospectifs, les recherches en sciences sociales montrent que la pensée critique, bien que valorisée en principe, se heurte aux pratiques organisationnelles dominées par un positivisme de l'action qui privilégie l'exécution de solutions validées à leur remise en question. Le sociologue Vincent de Gaulejac en résume ainsi les implications :

“ Cette posture sanctionne l'esprit critique. La critique est perçue comme négative, inutile, voire nuisible. Celui qui l'exprime est considéré comme inefficace et improductif. Dans un monde qui prône l'efficacité dans tous les registres, la contestation de la hiérarchie et a fortiori celle du système ne sont plus recevables³² ”

En somme, ces pratiques illustrent un mésusage de l'IA générative fondé sur une compréhension erronée de sa nature et de ses limites épistémiques. Or, ce type de mésusage, intégration implicite de contenus erronés ou infondés dans les raisonnements ou décisions, échappe souvent aux dispositifs classiques de gouvernance, car il ne laisse pas nécessairement de trace technique explicite. Dans un contexte de Shadow AI, où les usages sont informels, discrets ou non déclarés, ce type de dérive cognitive ne peut être détecté ni encadré.

³¹ WEF (2025), “The Future of Jobs Report 2025”, <https://www.weforum.org/publications/the-future-of-jobs-report-2025/>

³² “Positivisme et "gestion des émotions", ces deux freins à l'esprit critique en entreprise”|<https://www.atlantico.fr/decryptage/3509437/positivisme-et-gestion-des-emotions--ces-deux-freins-a-l-esprit-critique-en-entreprise-vincent-de-gaulejac>

2.4.2. Des pratiques clandestines qui intègrent des données sensibles engageant la responsabilité des organisations

L'utilisation d'outils d'IA générative en dehors de tout cadre conduit de plus en plus souvent à la diffusion involontaire de données sensibles. Ces pratiques, à la fois massives, invisibles et difficilement traçables, exposent directement les organisations à des risques juridiques, sécuritaires et réputationnels. Elles soulèvent des enjeux critiques en matière de responsabilité, de gouvernance et de conformité, dans un contexte où les cadres réglementaires se renforcent et où l'exigence de vigilance devient systémique.

Fin 2024, Harmonic Research a examiné la manière dont les données sensibles sont partagées de manière involontaire avec les outils d'IA générative³³. La méthode utilisée par l'équipe de recherche de Harmonic pour ce rapport a consisté à analyser des dizaines de milliers d'invites (prompts). Ces invites provenaient de milliers d'utilisateurs qui les ont soumises à divers outils d'IA générative (GenAI), incluant ChatGPT, Copilot, Gemini, Claudet et Perplexity. Cette analyse a été menée au cours du quatrième trimestre 2024 (Q4 2024) dans le but d'observer les tendances de fuite de données (data leakage).

- **8,5 % des invites soumises aux outils GenAI contiennent des données sensibles, soit un prompt sur 12.**
- 45,77 % des données sensibles étaient des données client, telles que les informations de facturation, les rapports client et les données d'authentification client.
- Les données d'employés représentaient 26,83 % des invites sensibles, incluant les données de paie, les informations personnellement identifiables (PII) et les dossiers d'emploi.
- Les données juridiques et financières représentaient 14,88 %, couvrant des informations sur les pipelines de vente, les portefeuilles d'investissement, les fusions et acquisitions, la documentation relative aux brevets.
- Les politiques de sécurité et les rapports constituaient 6,88 %, notamment les résultats de tests d'intrusion, les configurations réseau et les rapports d'incidents.
- Le code sensible, comme les clés d'accès et le code source propriétaire, représentait 5,64 %, posant des risques concurrentiels et exposant des vulnérabilités.
- 63,8 % des utilisateurs de ChatGPT utilisaient la version gratuite, et 53,5 % des invites sensibles ont été saisies via ce niveau gratuit. Le rapport note que l'utilisation des niveaux gratuits est importante, car de nombreux outils gratuits indiquent explicitement qu'ils entraînent leurs modèles sur les données client, ce qui signifie que les informations sensibles saisies

³³ Harmonic, "From Payrolls to Patents: The Spectrum of Data Leaked into GenAI Copy", <https://www.harmonic.security/resources/from-payrolls-to-patents-the-spectrum-of-data-leaked-into-genai>

pourraient être utilisées pour améliorer les modèles. L'utilisation de ces plateformes gratuites avec des données sensibles soulève des préoccupations supplémentaires pour les organisations souhaitant préserver leur propriété intellectuelle des modèles d'IA.

Pour illustrer les risques que de telles pratiques font courir aux organisations, L'avocate France Charruyer, directrice du cabinet d'avocats pluridisciplinaire en droit des affaires et des nouvelles technologies, évoque un « **paradoxe de la Reine Rouge** »³⁴, un concept utilisé pour décrire une situation où une adaptation ou un effort constant est nécessaire juste pour maintenir sa position ou son existence, et où il faut redoubler d'efforts pour réellement progresser. Dans le contexte de l'article et du phénomène du « Shadow Gen AI » (sic), ce paradoxe illustre la dynamique effrénée à laquelle sont confrontées les organisations : plus elles courent après la sécurisation de ce phénomène, plus elles doivent accélérer pour maintenir le contrôle. C'est une course sans fin pour ne pas se faire déborder par la rapidité de l'adoption de l'IA par les employés.

Ce paradoxe s'inscrit dans un cadre réglementaire de plus en plus contraignant où les organisations ont des obligations de moyens renforcées pour prévenir les risques pour la santé et la sécurité des travailleurs. Elles ont également une obligation générale d'assurer la sécurité du traitement des données personnelles conformément à l'article 32 du RGPD, nécessitant des mesures techniques et organisationnelles appropriées. S'ajoutent à cela la multiplication des obligations de sécurité sectorielles « by design » en cybersécurité et résilience (DORA, NIS 2, LPM, REC). Le Cyber Resilience Act (CRA) du 23 octobre 2024 impose des exigences de cybersécurité aux fabricants de produits numériques. Des règlements d'exécution de NIS 2 (17 octobre 2024) et DORA (29 novembre 2024) établissent des exigences techniques et méthodologiques. Le règlement d'exécution de la NIS 2 oblige notamment les entités à surveiller et journaliser les activités sur leurs réseaux pour détecter et réagir aux incidents.

Dans ce contexte, le « Shadow Gen AI » accroît brutalement les risques de non-conformité et peut engager la responsabilité de l'employeur : faute de contrôle, les données peuvent s'échapper, les droits de tiers être violés, des failles cyber ou des troubles pour la santé des salariés apparaître, tandis que l'empreinte environnementale et les dérives éthiques (biais, perte d'autonomie, opacité) s'amplifient. Cette vulnérabilité tient à une adoption décentralisée, rapide et souvent improvisée, qui contourne les processus d'approbation et multiplie les accès non autorisés aux informations sensibles ; elle expose donc dirigeants et entités à des poursuites administratives, civiles, voire pénales.

³⁴ Charruyer, F. (2025), « L'accélération technologique et le défi du Shadow Gen AI : l'effet Reine Rouge », Expertises, pp. 24-28. « L'expression "paradoxe de la Reine rouge" provient d'une citation du livre "De l'autre côté du miroir" de Lewis Carroll : « Ici, vous voyez, il faut courir le plus vite possible pour rester au même endroit. Si vous voulez aller ailleurs, il vous faut courir encore deux fois plus vite ! ».

Pour y répondre, France Charruyer prône une stratégie multidisciplinaire couplant protection des données, droit social, RH et gouvernance transparente, afin de reprendre conscience des implications de leurs usages. Les organisations doivent décider de leur posture (interdiction, autorisation encadrée, hybridation), détecter les signaux faibles du Shadow AI (pics d'usage, trafics suspects), instaurer une gouvernance solide (analyses d'impact, contrôles, comités éthiques, audits, montée en compétences, veille réglementaire), anticiper les menaces cyber propres à l'IA générative, clarifier le régime de responsabilité en cas de dommage et vérifier la couverture assurantielle adéquate.

Ces travaux mettent en lumière la fréquence alarmante des fuites involontaires de données sensibles via les outils d'IA générative, particulièrement sur les plateformes gratuites telles que ChatGPT. Face à ces risques croissants, exacerbés par l'adoption rapide et non contrôlée ; les organisations doivent impérativement renforcer leur gouvernance et leurs pratiques en matière de sécurité et de conformité réglementaire. Une approche pluridisciplinaire combinant veille proactive, contrôle renforcé, sensibilisation et formation continue, devient indispensable pour gérer efficacement ces défis et limiter les risques.

2.4.3. Des pratiques honteuses sur fond de peur pour l'avenir du travail

Les usages « shadow » de l'IA au travail se déploient dans un cadre moral péjoratif qui tend à alimenter les stratégies clandestines des collaborateurs. **Les pratiques s'inscrivent en effet dans une dynamique sociale marquée par la honte, la peur et l'anticipation négative.** Loin d'être neutre, le recours dissimulé à ces outils traduit des tensions morales et identitaires, révélatrices de la difficulté à assumer publiquement ces pratiques. À cela s'ajoute une inquiétude plus structurelle liée à l'impact à long terme de l'IA sur les conditions de travail, la charge perçue, et la pérennité des emplois. Ces facteurs alimentent la réticence, voire l'angoisse, à l'égard de cette technologie.

Une étude de Slack (2024)³⁵ montre ainsi que près de la moitié (48 %) des employés de bureau se sentiraient mal à l'aise à l'idée d'admettre à leur responsable qu'ils utilisent l'IA pour des tâches courantes au travail. Les principales raisons de ce malaise étant l'impression que l'utilisation de l'IA équivaut à tricher, la peur d'être perçu comme moins compétent, et la crainte d'être considéré comme paresseux.

³⁵ <https://slack.com/intl/fr-fr/blog/news/the-fall-2024-workforce-index-shows-ai-hype-is-cooling>.

Les employés français ressentent notamment le plus de gêne à l'idée de révéler qu'ils utilisent l'IA pour la rédaction de messages destinés aux managers (38 %) et aux collègues du même niveau (37 %), suivie par la génération d'idées pour des projets (32 %) et l'évaluation des performances de leurs subordonnés directs (31 %). « Notre étude révèle que même lorsque l'IA permet d'accomplir une tâche plus rapidement et efficacement, de nombreuses personnes préfèrent que leurs responsables n'en soient pas informés », explique Janzer, responsable du Slack Workforce Lab. « Les dirigeants doivent comprendre que cette technologie ne se limite pas au contexte professionnel consistant à se demander 'Puis-je accomplir cette tâche aussi vite et efficacement que possible ?' mais s'inscrit également dans un contexte social où la question devient : 'Qu'est-ce que les autres penseront s'ils apprennent que j'ai eu recours à cet outil ? ».

Outre cette charge morale péjorative anticipée sur leurs usages professionnels de l'IA, l'étude de Slack rend compte de l'inquiétude des employés concernant l'exploitation de l'IA générative par leurs employeurs. A la question « Si l'IA vous faisait gagner plusieurs heures par semaine au travail, comment souhaiteriez-vous idéalement utiliser ce temps libéré ? », les répondants français déclarent vouloir consacrer ce temps à des tâches administratives, puis à des activités non professionnelles, et enfin à la formation. A la même question, les employeurs français disent vouloir le reporter sur des tâches administratives, puis à la formation et aux autres projets principaux existants. « Les employés craignent que le temps économisé grâce à l'IA augmente en réalité leur charge de travail, leurs dirigeants s'attendant à ce qu'ils accomplissent plus de tâches, plus rapidement », explique Janzer. « Cela représente une opportunité pour les dirigeants de redéfinir ce qu'ils entendent par 'productivité', en incitant leurs employés à améliorer la qualité de leur travail, et non simplement la quantité ». Ces travaux rejoignent les constats de BCG sur les inquiétudes des français sur l'IA générative qui dépassent d'ailleurs largement celles exprimées par les autres pays sondés³⁶. 31% des employés français se disent préoccupés par l'usage de l'IA générative au travail contre 15% en moyenne, ce qui fait de la France le pays le plus pessimiste sur le sujet. Cela fait écho à l'étude Ifop Learning (op.cit.), qui indique que 68% des Français interrogés anticipent des conséquences négatives pour leur entreprise, 4 salariés sur 10 estiment que l'IA pourra, à terme, effectuer l'essentiel de leur travail et que 27% estiment que ce transfert s'opérera dans les 10 prochaines années.

³⁶ BCG (2024), « Consumers know more about GenAI than business leaders think »

La réaction immédiate du monde académique génère également une lecture moralisante et déontologique des pratiques informelles dans des considérations morales. L'interdiction de l'usage de ChatGPT à Sciences Po, communiquée dès janvier 2023, en assimilant, au nom de « l'honnêteté intellectuelle », toute utilisation non encadrée à une faute relevant du plagiat ou de la fraude contribue, à enfermer l'expérimentation individuelle dans une logique de transgression à cacher ou camoufler :

“ L'utilisation, sans mention explicite, de ChatGPT à Sciences Po, ou de tout autre outil ayant recours à l'IA est, à l'exception d'un usage pédagogique encadré par un enseignant, pour l'instant strictement interdite lors de la production de travaux écrits ou oraux par les étudiants et étudiantes sous peine de sanctions qui peuvent aller jusqu'à l'exclusion de l'établissement, voire de l'enseignement supérieur. L'outil ChatGPT interroge fortement les acteurs de l'éducation et de la recherche dans le monde entier sur le sujet de la fraude en général, et du plagiat en particulier. Sciences Po est soucieux de garantir la qualité et l'intégrité de ses formations et de ses diplômes ³⁷ ”

Dans le même sens, le sociologue Bilel Benbouzid décrit et analyse comment « les IAG ébranlent les fondements de l'auctorialité et de la déontologie universitaire »³⁸, les problèmes se cristallisant selon lui, tant pour les chercheurs que pour les étudiants, autour de l'écriture. Le raisonnement s'articule autour de trois repères normatifs associés à la pratique de l'écriture : l'intégrité académique, l'idéal éthique d'authenticité, et le fondement de la moralité, l'autonomie. L'un des principaux enjeux déontologiques est l'intégrité académique, valeur fondamentale qui engage enseignants, chercheurs et étudiants à adopter des pratiques honnêtes, rigoureuses et responsables. Cela implique la reconnaissance du travail d'autrui, la transparence et la loyauté intellectuelle. Pour Benbouzid, l'usage des IAG met à l'épreuve chacune de ces dimensions. Les revues et éditeurs scientifiques peinent à encadrer l'usage des IAG, s'accordant à interdire qu'ils soient crédités comme co-auteurs, mais restant hétérogènes sur les modalités de divulgation. Ce flou normatif s'étend aux travaux des étudiants, contribuant à l'ambiguïté autour des notions de plagiat ou de triche, souvent mobilisées pour qualifier les problèmes posés par les IAG.

³⁷ <https://newsroom.sciencespo.fr/sciences-po-interdit-utilisation-de-loutil-chatgpt>

³⁸ « Écrire à l'université à l'heure des IA génératives : trouble dans l'auctorialité », AOC, 7 mai 2025.

Cependant, l'usage d'une IAG n'est pas exactement un cas de plagiat, car il ne s'agit pas d'utiliser les idées ou textes d'un auteur humain sans attribution. Le problème de la citation avec les IAG est paradoxal : leurs énoncés n'ont pas de valeur référentielle claire pour appuyer un argument, mais il est attendu que les lecteurs soient informés de leur usage. En ce qui concerne la triche, l'usage d'une IAG est sans ambiguïté une tricherie dans un contexte où elle est interdite (par exemple, un examen sur table). Mais en dehors de ces situations balisées, la ligne devient plus floue. S'appuyer sur une IAG pour un devoir à la maison, comme on demanderait de l'aide à un proche, rend difficile l'établissement du caractère « injuste » de l'avantage. Cela touche aussi les chercheurs concernant l'équité dans la compétition scientifique, d'autant que les évaluations des articles tendent elles-mêmes à être automatisées par les IAG. L'usage des IAG, n'étant ni tout à fait triche ni véritable plagiat, soulève une question plus fondamentale : quelle est la nature de l'effort intellectuel légitime à fournir dans un contexte académique ? Définir un seuil d'effort légitime reviendrait à réguler sur un mode strictement répressif. Cependant, cette approche se heurte à une impossibilité pratique de contrôle et de détection fiable, car il est difficile de distinguer un usage « passif » d'un usage « actif » de l'IAG. Un texte peut être formellement original, échappant aux détecteurs, tout en étant « frauduleux car non-authentiquement de l'auteur ». Ainsi, l'application du principe d'intégrité académique se heurte à cette difficulté de contrôle, « tant ces outils redéfinissent la frontière entre l'originalité formelle d'un contenu et l'expression authentique de son auteur ». Du côté des étudiants, le sociologue fait de l'écriture un acte fondateur de l'autonomie, bien plus qu'un acte éducatif, une « obligation morale envers soi-même ». Cette obligation consiste à cultiver les conditions de possibilité de sa liberté en développant les capacités de réflexion, de jugement et d'auto-détermination. L'acte d'écrire n'est pas juste une compétence technique, mais un acte de formation éthique, un moment où l'individu exerce sa capacité à réfléchir sur ses engagements, à formuler des arguments et à construire un point de vue justifiable. L'auteur distingue l'autonomie comme liberté de choix de l'autonomie comme fin éducative. La première est la capacité de choisir les moyens pour atteindre un but. Vue sous cet angle, l'utilisation d'une IAG pourrait sembler augmenter l'autonomie de l'étudiant en facilitant la tâche. La seconde suppose une lente élaboration de soi à travers des pratiques intellectuelles rigoureuses, cultivée par l'effort de penser, structurer ses idées et argumenter, des gestes rendus possibles par l'écriture. Or, l'autonomie procédurale, axée sur l'efficacité, contredit l'autonomie comme finalité morale de l'éducation. Déléguer l'écriture à une IA revient à renoncer à exercer les facultés qui rendent l'autonomie possible (délibération, critique, formulation d'un point de vue propre).

³⁷ <https://newsroom.sciencespo.fr/sciences-po-interdit-utilisation-de-loutil-chatgpt>

³⁸ « Écrire à l'université à l'heure des IA génératives : trouble dans l'auctorialité », AOC, 7 mai 2025.

L'étudiant qui utilise l'IA pour écrire à sa place cultive moins sa capacité à être l'"auteur" de ses choix et jugements. Selon Benbouzid, une possible éthique face aux IAG passerait par la capacité et la liberté de choisir de ne pas les utiliser, même quand c'est techniquement possible. Bien que déployé exclusivement pour penser les enjeux académiques de l'écriture avec l'IAG, le raisonnement est transposable à un cadre de réflexion plus général, dans des sociétés individualistes où chacun est supposé être l'auteur de sa vie et où le travail est présenté comme une des activités cardinales par laquelle les individus entreprennent ce projet.

Pour de nombreux utilisateurs, le développement du Shadow AI s'inscrit donc dans un cadre moral ambivalent, dominé par la honte, la peur de la transgression et l'injonction paradoxale à dissimuler. La clandestinité découle d'un cadre normatif culpabilisant qui associe largement l'usage de l'IA à la tricherie, un manque d'intégrité ou à un déficit de compétence. La nécessité de préserver l'intégrité des institutions et des organisations conduit souvent à stigmatiser le manque d'intégrité des usagers et à reporter sur eux la responsabilité morale de leurs pratiques. Ce régime de la honte fragilise la possibilité d'une appropriation éthique et ouverte des outils IA, et donc renforce les stratégies de dissimulation, ce qui révèle l'urgence de redéfinir les repères de légitimité, d'effort et d'authenticité dans des environnements professionnels et académiques transformés par ces technologies. En même temps, ce cadrage déontologique des pratiques informelles soulève des enjeux de fond, en particulier, sur l'autonomie comme principe moral et comme moteur de l'émancipation, qui ne peuvent être abordés dans une configuration clandestine de l'IA. Au travail, l'autonomie est en effet un "principe de justice" majeur. Le sociologue François Dubet montre en ce sens que chaque travailleur est conduit à juger de la justice de son travail en fonction de la liberté, de l'autonomie et de la réalisation de soi qu'il lui permet. A l'opposé, le sentiment d'injustice résulte de la fatigue, de l'usure, de l'absence d'intérêt pour la tâche, du sentiment de mépris et d'impuissance sur sa propre activité. L'autonomie consacre la réalisation de soi comme juste et l'aliénation comme injuste³⁹. Dans cette perspective, le développement du Shadow AI interroge profondément les conditions de justice au travail.

2.4.4. Une clandestinité au détriment du collectif de travail

Si l'usage de ChatGPT semble offrir des gains de productivité et de confort individuel, il engendre également des effets secondaires organisationnels plus insidieux. Loin d'être neutre, son intégration dans les pratiques quotidiennes reconfigure en profondeur les dynamiques de travail, de collaboration et de développement des compétences.

³⁹ Dubet, Injustices. L'expérience des inégalités au travail, Seuil, 2006.

Jana Retkowsky, Ella Hafermalz, Marleen Huysman (2024) ont mené une étude qualitative basée sur des entretiens menés auprès de 50 « early adopters » de ChatGPT. Les 50 participants provenaient de 50 organisations différentes, offrant une perspective large, et étaient tous engagés dans le travail du savoir ("knowledge work"), couvrant des domaines variés comme la gestion de projet, la recherche, le marketing, le copywriting, le design UX, l'informatique/programmation, l'investissement, le recrutement, le service client et le conseil. Les entretiens suivaient un protocole exploratoire avec des questions larges sur les expériences d'utilisation actuelles et passées de ChatGPT, ainsi que leurs perspectives. L'objectif était d'obtenir une perspective d'initié sur l'utilisation réelle de ChatGPT par les employés. L'approche par affordances a été choisie pour comprendre ce que ChatGPT permet concrètement aux employés de faire en pratique, plutôt que de se limiter à ses fonctionnalités conçues.

Six usages dominants de ChatGPT ont été ainsi identifiés, montrant une intégration croissante dans les routines professionnelles des utilisateurs. Le plus fréquent est la recherche d'informations (37 participants), où l'outil est perçu comme un moteur de recherche plus accessible et personnalisé, bien que les hallucinations nécessitent des vérifications. Le brainstorming (36) est valorisé pour sa disponibilité constante et sa capacité à stimuler la créativité. Ensuite vient la structuration du contenu (35), permettant de résumer ou reformuler des textes complexes selon des cadres précis, réduisant ainsi la dépendance aux experts internes. L'écriture de brouillons (34) aide à surmonter la page blanche et à passer plus vite à l'étape de l'édition. L'assistance à la relecture (25) offre un gain de temps, en particulier pour la correction de langage ou de code, et renforce la confiance des non-natifs. Enfin, l'embellissement du texte (16) permet d'ajuster le ton, d'ajouter des éléments stylistiques ou de peaufiner des prompts pour d'autres outils d'IA. Ces usages résultent d'un glissement d'une expérimentation informelle à une utilisation professionnelle pleinement intégrée.

L'étude identifie trois effets secondaires majeurs liés à l'usage de ChatGPT au travail, dont l'impact est peu visible des managers. D'abord, les liens de connaissance sont fragilisés : les interactions avec les collègues, auparavant mobilisées pour relire, échanger ou créer, sont remplacées par des échanges avec l'IA, affaiblissant la circulation des savoirs tacites et l'apprentissage collectif. Ensuite, une surveillance limitée de la qualité des contenus générés émerge : les employés peinent à distinguer ce qui provient de l'IA, ce qui complique le contrôle managérial et expose à des erreurs ou hallucinations intégrées dans les livrables.

Enfin, les rôles professionnels se reconfigurent, l'IA prenant en charge des tâches de base, et les employés devenant éditeurs ou chefs d'orchestre plutôt qu'exécutants. Cette évolution rend plus complexe l'évaluation des contributions et la montée en compétences, notamment pour les profils débutants, privés d'opportunités formatrices. L'étude pointe également un risque de polarisation des compétences, entre ceux qui exploitent l'IA avec efficacité et ceux qui restent en retrait.

Ces transformations silencieuses appellent une vigilance accrue des managers. Sans accompagnement adapté, l'usage massif de l'IA générative pourrait fragiliser les collectifs de travail, creuser les écarts de compétences et affaiblir les processus d'apprentissage organisationnel.

2.5. Du côté des organisations syndicales

Tout au long de cette étude, et en cohérence avec les principes de fonctionnement du LaborIA dont le Comex s'élargit, deux fois par an, aux partenaires sociaux, notre équipe a échangé avec les organisations syndicales. Elles développent des positions claires, en insistant sur les enjeux de régulation, de formation, de dialogue social et de gouvernance. Ces prises de position permettent de dégager plusieurs enseignements clés sur les risques, les opportunités et les conditions d'un usage encadré et responsable de l'IA par les salariés, dans un contexte où les pratiques échappent encore largement aux dispositifs formels des entreprises. La CFE-CGC, la CFTC, la CFDT et l'UGICT-CGT nous ont transmis chacune un texte, sous forme de prise de position, exprimant les enjeux identifiés pour la transformation du travail et des organisations.

- **Risques identifiés** : Les solutions d'IA non contrôlées peuvent aspirer de grandes quantités de données sensibles de l'entreprise, risquant ainsi leur exposition ou leur utilisation inappropriée. Des vulnérabilités peuvent également être introduites dans les systèmes de l'entreprise, car ces outils ne sont pas soumis aux mêmes contrôles de sécurité que les solutions approuvées par les services IT. Des risques éthiques et juridiques (vie privée, biais algorithmiques, responsabilité) sont également mentionnés.
- **Ne pas interdire, mais encadrer et accompagner** : L'idée générale n'est toutefois pas de prohiber l'usage de l'IA, même non officiel, mais de le structurer.
- **Nécessité de transparence et de gouvernance claire** : Il est crucial de favoriser la transparence dans l'usage des technologies d'IA et de mettre en place une gouvernance claire pour les outils utilisés.

- **Importance de la formation et de la sensibilisation** : Les salariés doivent être formés pour utiliser ces outils de manière responsable, éclairée et sécurisée. Il est essentiel de les sensibiliser aux enjeux juridiques, éthiques, techniques et aux risques liés à la sécurité et à la confidentialité des données. Cela permet de développer une culture numérique partagée et de limiter les initiatives individuelles non sécurisées.
- **Dialogue social et consultation des représentants du personnel** : Il est fortement recommandé d'impliquer les représentants du personnel dans les discussions pour définir les modalités d'utilisation de l'IA en entreprise. Cela peut se faire via la négociation de chartes ou l'intégration des usages (y compris non officiels) dans les accords collectifs et les chartes internes. Un processus d'information-consultation des instances représentatives est préconisé pour un déploiement transparent.
- **Cartographie des usages réels** : Il est nécessaire de cartographier l'ensemble des usages de l'IA au sein de l'entreprise, y compris ceux qui sont "souterrains".
- **Identifier les causes sous-jacentes** : L'IA clandestine peut être symptomatique de lacunes ou de dysfonctionnements dans les procédures, les outils métiers, ou l'organisation du travail. Elle peut être une opportunité de remettre en question l'organisation et les conditions de travail.
- **Déploiement d'outils internes sécurisés** : Une solution envisagée pour contrer ces pratiques non sécurisées est le déploiement par les entreprises de leurs propres outils d'IA sécurisés.
- **Prudence pour tous les utilisateurs, y compris les entités internes** : Même les acteurs internes, comme les représentants du personnel, doivent faire preuve de prudence lors de l'utilisation d'IA générative, notamment en évitant d'introduire des informations confidentielles dans les "prompts" et en respectant les principes du RGPD (responsabilité de traitement).
- **Conscience de l'impact environnemental** : Il est également pertinent de prendre en compte l'impact environnemental de l'IA, car une simple requête peut consommer significativement plus d'énergie qu'une recherche classique.

En résumé, la prolifération du Shadow AI met en lumière un besoin urgent de structuration, de formation et de dialogue au sein des entreprises pour un déploiement éthique, sécurisé et bénéfique de l'intelligence artificielle.

La **CFTC (Confédération Française des Travailleurs Chrétiens)**, bien que ne mentionnant pas explicitement le terme "**Shadow AI**", adopte une **position claire et implicite** face à ce phénomène, à travers les orientations stratégiques de son document.

Qu'est-ce que la Shadow AI ?

La Shadow AI désigne l'utilisation d'outils d'intelligence artificielle (IA) au sein des entreprises **en dehors des circuits officiels** (non validés par la DSI ou les services conformité). Cela inclut l'usage non encadré de chatbots, de générateurs de texte ou d'outils de data science par les salariés, souvent à des fins d'efficacité ou d'expérimentation. Ces pratiques posent **des risques éthiques, juridiques et sécuritaires**.

Position implicite de la CFTC face à la Shadow AI

1. Encadrer et réguler l'usage de l'IA en entreprise

La CFTC appelle à une **régulation forte** de l'intelligence artificielle, ce qui s'applique pleinement à la Shadow AI :

- Elle insiste sur **l'importance de la transparence** dans l'usage de ces technologies.
- Elle défend une **IA éthique et conforme aux valeurs européennes**, ce qui suppose une **gouvernance claire** des outils utilisés dans les organisations.
- Elle alerte sur les **risques juridiques** d'un usage non contrôlé de l'IA, y compris en matière de vie privée, de biais algorithmiques et de responsabilité.

2. Former les salariés aux bons usages

La CFTC ne cherche pas à interdire l'usage de l'IA, mais à **former les salariés** pour qu'ils :

- Utilisent ces outils de façon **responsable, éclairée et sécurisée**.
- Soient conscients des **enjeux juridiques, éthiques et techniques** liés à l'usage informel de ces outils.
- Contribuent à une **culture numérique partagée**, limitant les initiatives individuelles non sécurisées.

3. Inclure la Shadow AI dans les dialogues sociaux

En promouvant un **projet intersyndical d'identification des tâches automatisables**, la CFTC sous-entend la nécessité de :

- **Cartographier les usages réels de l'IA**, y compris ceux "souterrains" (Shadow AI).
- Impliquer les représentants du personnel dans la **surveillance et l'encadrement** des outils IA déployés.
- **Intégrer l'IA dans les accords collectifs** et les chartes internes, y compris ses usages non officiels.

En résumé la CFTC :

- **Reconnait la réalité de la Shadow AI** comme symptôme d'un manque d'anticipation ou de cadrage officiel.
- **Prône une réponse structurée** : formation, régulation, dialogue social.
- **Rejette les dérives individualistes ou opaques** et milite pour un usage **collectif, éthique et encadré** de l'intelligence artificielle dans les entreprises



La CFDT (Confédération française démocratique du travail)

La démocratisation et l'usage à grande échelle des IA génératives depuis la mise à disposition de ChatGPT au grand public en novembre 2022 ont eu des impacts importants dans le monde du travail. La simplicité d'usage de ces nouveaux outils, qui ne nécessitent pas de matériel, d'infrastructure ou d'investissement conséquents, et leur facilité d'accès, ont conduit de nombreux travailleurs à les utiliser dans le cadre de leur activité professionnelle de manière spontanée. Ces usages, parfois dissimulés, sans prescription ni accompagnement de l'employeur, interpelle la CFDT à plusieurs titres.

D'abord en raison des problèmes de sécurité posés. Les conditions d'utilisation des outils d'IAG sont mal connus et peu explicites pour leurs utilisateurs finaux. Leur usage (rédaction, synthèse de documents, compte rendus) implique d'intégrer des données professionnelles qui devraient être protégées, ce qui peut conduire à des fuites de données confidentielles et à une exposition accrue aux cyberattaques. Ensuite, les conditions d'élaboration de ces outils et leur manque de transparence dans leur fonctionnement peuvent conduire les travailleurs à répliquer à leur insu des biais qui renforcent des préjugés et dégradent la qualité des décisions prises (en matière de recrutement, de finance, ou d'accès aux services publics par exemple). Enfin de manière générale, l'usage non accompagné des IA génératives peut conduire à des violations du RGPD sur le partage et l'exploitation de données personnelles et protégées.

Si les outils d'IA génératives présentent un intérêt et un bénéfice pour la réalisation des activités, les travailleurs doivent pouvoir les utiliser dans un cadre de confiance et de manière transparente. Cela suppose que les règles et les cadres d'usage des outils d'IA fassent l'objet d'un véritable dialogue social dans les entreprises et les administrations.

Il est également indispensable de sécuriser les données mobilisées dans les outils d'IA et de former l'ensemble des acteurs (salariés, agents, managers, prestataires, directions, partenaires sociaux) aux bénéfices, aux risques, aux impacts sociétaux et environnementaux, et aux modalités d'usages respectueuses des droits.

Plus largement la nécessité d'instaurer un dialogue social sur les outils d'IA génératives montre l'importance de développer un dialogue social sur l'ensemble des technologies numériques au travail. L'émergence des IA génératives ne doit en effet pas occulter les autres usages des systèmes algorithmiques dans de nombreuses activités (planification, management, gestion des ressources humaines, évaluation, aide à la décision, etc.) et dont le déploiement n'est pas assez discuté et reste trop souvent absent du dialogue social, alors qu'ils transforment en profondeur les manières de travailler et les relations collectives. Le besoin d'agir et d'encadrer les utilisations spontanées des IA génératives par les travailleurs doit servir de levier, lorsque ce n'est pas le cas, pour faire du déploiement des outils numériques dans le travail un véritable objet de dialogue social.



LA CFE-CGC (Confédération française de l'encadrement - Confédération générale des cadres) : interview de Nicolas Blanc par Mathieu Bahuet

Créé en novembre 2022, l'outil d'intelligence artificielle ChatGPT a fêté ses deux ans. Quel bilan peut-on en tirer de son déploiement ?

ChatGPT a rapidement dépassé son objectif initial de simple projet de démonstration. On ne peut plus l'ignorer. Conçu par l'entreprise américaine OpenAI pour explorer comment les utilisateurs interagiraient avec un outil basé sur le modèle de langage GPT 3.5, ChatGPT a connu une adoption massive, surprenant même ses créateurs. En seulement deux mois, il a rassemblé une base de 100 millions d'utilisateurs ! Ce succès témoigne du potentiel des grands modèles de langage (LLM), qui s'appuient sur l'architecture « transformer » introduite par Google en 2017. Grâce à des mécanismes d'auto-attention, ces modèles traitent et génèrent du texte à la manière humaine et excellent dans des tâches telles que la compréhension du langage naturel.

Et dans l'univers professionnel ?

On voit déjà émerger plein d'usages au travail mais il faut aujourd'hui encadrer ces pratiques qui vont encore se développer avec l'arrivée de Microsoft Copilot, un outil de type ChatGPT qui va être intégré dans sa suite bureautique et de logiciel.

De plus en plus de salariés utiliseraient « discrètement » ChatGPT pour leur travail. Qu'en pensez-vous ?

Ce phénomène de « Shadow AI » (IA clandestine) désigne l'utilisation de solutions d'IA générative telles que ChatGPT ou Google Gemini sans l'approbation de l'entreprise. Par exemple l'utilisation d'une IA sur les appareils personnels du salarié pour corriger du texte, générer du contenu ou rédiger un e-mail à un client. Les cas d'utilisation de l'IA en dépit des politiques mises en place par les organisations se multiplient. Malgré des avantages indéniables, cette utilisation « clandestine » soulève de sérieuses préoccupations concernant la sécurité et la confidentialité des données. Certains GAFAM (Microsoft, Apple, Amazon) ont d'ailleurs interdit à leurs salariés d'utiliser ChatGPT pour un usage professionnel. Cela n'est pas anodin !

En quoi cela constitue un risque pour les entreprises ?

Indépendamment des règles mises en place, le Shadow AI se traduit par des risques bien réels dans la mesure où les solutions d'IA aspirent une grande quantité de données pour fonctionner. Sans contrôle ni supervision, il est possible que des données sensibles de l'entreprise soient exposées ou utilisées de manière inappropriée. Le Shadow AI peut aussi introduire des vulnérabilités dans les systèmes de l'entreprise car les solutions utilisées ne sont pas soumises aux mêmes contrôles de sécurité que celles approuvées par le service IT... La sécurité et la confidentialité de données des entreprises sont donc les deux principaux risques et il ne faut donc pas que les salariés se mettent en péril.

Comment peut-on encadrer ces pratiques ?

L'idée n'est pas de les interdire mais de les accompagner. Pour ce faire, les entreprises doivent avant tout en discuter avec les représentants du personnel afin de reprendre la main et définir des modalités de l'utilisation par leurs salariés de ce type d'outils. Cela peut passer par la négociation d'une charte dans laquelle les salariés s'engagent à respecter les règles négociées par accord. Il faut bien sûr y associer une politique de sensibilisation et de formation des salariés pour rendre la démarche cohérente.

Il est aussi essentiel de fournir une formation adéquate aux employés sur les risques liés à l'utilisation de l'IA et sur les bonnes pratiques à adopter. Cela permettra de les sensibiliser aux enjeux de sécurité et de confidentialité des données ainsi qu'aux limites et aux utilisations appropriées de l'IA générative. Certaines entreprises ont déjà pris les devants en déployant leur propre ChatGPT sécurisé afin de contrer ces pratiques sauvages. Dans tous les cas, nous préconisons dans les entreprises de passer par une

information-consultation pour associer les salariés et leurs représentants dans les conseils économiques et sociaux (CSE). Cela permet un processus de déploiement transparent, en douceur, pour l'ensemble des collaborateurs.

Les représentants du personnel et les syndicats peuvent aussi avoir recours à l'IA générative. Quels doivent être les bons usages ?

Bien sûr, on peut imaginer que de tels outils puissent les aider à préparer des synthèses de réunions de CSE, à rédiger des tracts syndicaux ou à négocier des accords. Là encore, la prudence est de mise. Sans chercher à les interdire, il faut cependant encadrer ces usages et rappeler les bonnes pratiques en matière de confidentialité et de sécurité des données. Les documents remis dans le cadre d'une information-consultation peuvent être confidentiels. Idem pour la négociation d'un accord.

Par ailleurs, l'utilisation de solution payante, par exemple ChatGPT, ne garantit pas une sécurité absolue comme le rappelle OpenAI dans ses FAQ. En effet, il faut aussi ne rien indiquer de confidentiel dans les échanges qu'on a avec l'outil, ce qu'on appelle les « prompts ». Or on le sait rarement... Globalement, il convient donc que les acteurs adoptent un positionnement de responsable de traitements au sens du RGPD, avec les engagements induits.

Comment la CFE-CGC travaille-t-elle tous ces sujets ?

Il y a longtemps que nous investiguons sur l'IA, par exemple avec le projet européen SecoIADeal sur dialogue social et l'IA. Nous allons travailler sur un guide des intelligences artificielles qui rappellera les bonnes pratiques pour nos adhérents et nos militants. Ce sera aussi l'occasion de faire de la pédagogie sur l'utilisation de ces systèmes d'un point de vue environnemental puisque, selon diverses études, une simple requête sur ChatGPT consomme dix fois plus d'énergie qu'une recherche Google.

En parallèle, notre réseau de référents IA dans les fédérations CFE-CGC sont autant de garants pour la diffusion de ces règles au plus près de nos adhérents et de nos militants. Nous travaillons aussi sur la qualification d'un certain nombre d'outils d'IA ou de formation qui respectent un haut niveau d'exigence de sécurité et de confidentialité pour permettre à nos adhérents de profiter sereinement des avantages de l'IA générative.

L'anglicisme « shadow AI », auquel on préférera le terme « IA clandestine », désigne simplement l'utilisation dans un cadre professionnel d'un système IA non fourni par l'employeur et non intégré au système d'information professionnel.


Sur la forme, il s'agit bien d'un phénomène nouveau : les systèmes IA accessibles au grand public, comme les IA génératives, sont récents. Mais sur le fond, il s'agit d'un phénomène ancien, dans la continuité de la pratique de « ramener des dossiers à la maison » et similaire à l'utilisation professionnelle clandestine de matériels ou logiciels non fournis par l'employeur. Cette pratique, qui est apparue avec l'informatique personnelle, a vu son essor avec la démocratisation des ordinateurs portables puis des smartphones, avant de connaître un pic au moment du confinement et du télétravail improvisé en mars 2020. Souvenons-nous également de la vague du BYOD (« Bring Your Own Device ») pendant laquelle il a été question d'abolir la frontière entre matériel professionnel et matériel personnel : cette vague a pris fin quasiment aussi vite qu'elle était apparue, du fait des problématiques de sécurité et de confidentialité qu'elle a soulevées. Si l'IA clandestine s'inscrit également dans ces problématiques, elle ne fait là encore que rejoindre la longue et ancienne liste des logiciels utilisés clandestinement dans un cadre professionnel, notamment les messageries instantanées personnelles.

Est-il possible de mesurer l'utilisation de l'IA clandestine ? En dehors d'une utilisation avec du matériel professionnel ou via le réseau professionnel, non.

D'en empêcher l'utilisation ? Même en contrôlant l'utilisation des logiciels sur le matériel professionnel, ou en empêchant l'utilisation dans les locaux professionnels de matériel personnel, il resterait la possibilité d'un usage domestique de ce même matériel.

L'IA clandestine est-elle révélatrice de la valeur que peut apporter l'IA aux travailleurs et aux organisations ? Pas nécessairement, car elle peut se manifester aussi bien en cas d'absence que d'existence de système IA professionnel et, en cas d'existence, aussi bien si le système IA professionnel est d'usage obligatoire et intégré aux procédures et outils métiers que s'il est d'usage facultatif.

Est-il nécessaire de sortir de l'IA clandestine ? Considérons plutôt la question sous l'angle que l'IA clandestine peut être symptomatique de lacunes ou de dysfonctionnements des procédures ou outils métiers, ou encore de dysfonctionnements de l'organisation du travail et donc de risques organisationnels.



Il faut donc y voir l'occasion de revenir aux fondamentaux et de questionner tant l'organisation du travail que les conditions de travail.

C'est d'ailleurs dans ce questionnement que réside un chemin de sortie qui bénéficiera à la fois aux employés et aux employeurs. La méthodologie pour ce faire est tout simplement la même que celle qui devrait être systématiquement utilisée pour le déploiement d'un système IA professionnel. L'UGICT-CGT préconise donc ses propres travaux sur le sujet, ainsi que les accords cadre européens sur la transformation numérique, le projet intersyndical DIALIA mené par l'IRES, ou encore les travaux du LaborIA

Longtemps cantonnée à des démonstrateurs technologiques peu déployés, l'introduction de l'IA dans les environnements de travail connaît un changement de trajectoire radical avec l'irruption de ChatGPT et de l'IA générative, qui a rendu l'outil directement accessible à des millions d'utilisateurs, sans médiation hiérarchique ou technique.

Loin des cas d'usage industriels pensés par les directions, les usages réels s'apparentent à du bricolage : des ajustements empiriques, personnels, souvent invisibles, portés par des professionnels qui s'approprient l'outil en fonction de leurs contraintes et de leurs tâches. Ces usages informels témoignent d'un fort potentiel d'appropriation et de transformation du travail réel : une commerciale étrangère qui adapte ses e-mails au style français grâce à ChatGPT, un comptable qui l'utilise comme assistant documentaire, un designer qui dialogue créativement avec une IA graphique, ou encore un médecin hospitalier qui rédige des courriers empathiques à ses patients.

Le Shadow AI incarne une tension structurante dans l'introduction de l'intelligence artificielle au travail : il révèle une formidable capacité d'initiative des collaborateurs à s'approprier les outils de manière inventive, parfois pertinente, et adaptée aux contraintes de terrain. Ce sont des professionnels qui expérimentent en direct ce que l'IA peut leur apporter, bien avant que leur organisation n'ait songé à le formaliser. En ce sens, le Shadow AI fonctionne comme un terrain d'expérimentation distribuée, qui alimente des usages concrets et situe l'IA au cœur du travail réel.

Cette dynamique ascendante s'appuie sur des usages dits augmentatifs (qui enrichissent les capacités cognitives) plus que déléгатifs (qui visent à automatiser des tâches). Le recours à l'IA est souvent un levier de gain de temps, de réflexivité, voire d'émancipation — mais toujours situé dans un contexte professionnel précis. Pour autant, ces pratiques informelles ne sont pas sans risque : mésusages profanes, fuites de données sensibles, incompréhension des limites techniques des IA génératives, ou encore isolement des collaborateurs dans des pratiques individuelles et parfois honteuses sur fond de peur pour l'avenir de leur emploi.

Mais cette dynamique se déploie la plupart du temps sans cadre clair, sans discussion collective et sans dispositif de formation structuré. Cela expose les organisations à des risques sérieux : non-conformité réglementaire, déséquilibres dans les collectifs de travail, baisse de qualité ou de traçabilité des productions, voire des effets contre-productifs sur la confiance ou l'autonomie professionnelle. Le Shadow AI révèle donc autant les potentialités d'une IA encapacitante que les vulnérabilités d'un environnement sans gouvernance.

Face à ce phénomène, les organisations pionnières rencontrées ne se contentent pas d'interdire ou de tolérer passivement ces usages spontanés. La deuxième partie du rapport analyse les réponses concrètes apportées par ces employeurs : comment ils identifient les pratiques, évaluent les risques, construisent des cadres d'usage, et surtout, comment ils transforment cette créativité individuelle en stratégie collective. Ces démarches visent à concilier la sécurisation des usages avec l'amplification de la valeur produite par l'IA sur le terrain.

ENCADRER LES BRICOLEURS : LES RÉPONSES DES EMPLOYEURS AUX USAGES INFORMELS⁴⁰

3.1. Méthodologie de notre enquête : à la rencontre d'organisations pionnières

Notre étude vise à comprendre comment les organisations réagissent aux initiatives d'intelligence artificielle générative prises spontanément par leurs salariés. Le protocole combine une investigation qualitative auprès d'acteurs clés évoluant au sein d'« organisations pionnières » et une phase de co-construction collective afin d'aboutir à des guides opérationnels.

Le statut d'« organisation pionnière » doit s'entendre à plusieurs niveaux. Dans un premier temps, il s'agit d'organisations qui ont identifié et adressé des réponses institutionnelles concrètes aux pratiques informelles de leurs collaborateurs. Elles ont en ce sens accepté de partager ces actions avec une équipe de recherche afin qu'elle puisse en publier les principaux enseignements. Dans un deuxième temps, ces organisations ont été sélectionnées pour leur position dans leur champ respectif, leur capital de ressources et leur comportement d'avant-garde identifiés par notre équipe à l'occasion de différentes prises de parole et rencontres.

⁴⁰ Dans cette section, l'usage de l'IA a été important à de nombreux niveaux. Chaque entretien a été enregistré puis transcrit à l'aide d'un logiciel approuvé par le Comité opérationnel d'évaluation des risques légaux et éthique d'Inria en termes de sécurité et de confidentialité. <https://www.css.cnrs.fr/fr/whisper-pour-retranscrire-des-entretiens2/>. Le gain de temps est immédiat mais se réduit face à la nécessité de corriger a posteriori de nombreuses erreurs qu'on ne peut reprendre sans l'appui du fichier source. Ce travail est fastidieux, il s'apparente à la préparation de données. Pour dialoguer avec le recueil des auditions retranscrites, nous avons à nouveau utilisé NotebookLM, après avoir effacé toutes les références aux organisations pionnières, tâche également peu stimulante. Par la suite, l'assistant a vraiment permis une exploitation performante de notre matériau de recherche en renvoyant notamment aux verbatims. Mais l'application ne distinguait pas nos propres questions des réponses. Bien que frustrante, cette limite nous a permis de noter que certaines de nos questions étaient biaisées car trop orientées, insuffisamment ouvertes. Nous avons enfin utilisé plusieurs modèles de langage (Le Chat, Claude, ChatGPT o3) pour améliorer nos modélisations. Nous réalisons la version princeps et l'introduisons dans plusieurs modèles. Nous choisissons le plus satisfaisant puis faisons retravailler tous les modèles sur cette version en suggérant des axes d'amélioration. Plusieurs itérations ont été effectuées avant validation.

Enfin, à ces critères s'ajoutent des considérations de diversité de secteurs (public et privé), de taille d'entreprise et de domaines d'activité.

L'équipe de recherche a mené 14 auditions sous la forme d'entretiens semi-directifs⁴¹ avec les acteurs engagés au sein de leur organisation sur le déploiement de l'IA en général et de l'IA Gen, en particulier. Ce sont en majorité des DSI, des CDO, des DRH ou des chefs de projet IA. Le guide d'entretien s'articulait autour de six axes :

1. Définition du Shadow AI par les acteurs eux-mêmes ;
2. Évaluation quantitative du nombre de collaborateurs concernés par les usages spontanés de l'IA (volume d'utilisateurs, fréquence, outils employés) ;
3. Typologie des pratiques (productivité, créativité, automatisation, etc.) ;
4. Analyse FFOM (forces, faiblesses, opportunités, menaces) pour l'organisation ;
5. Réponses organisationnelles déjà engagées ou envisagées ;
6. Freins et recommandations pour « sortir » du Shadow AI.

es premiers résultats ont servi de socle à une journée d'étude présentielle réunissant représentants d'entreprises, acteurs publics et partenaires sociaux, la majorité ayant été auditionnée lors de la première phase.

Cette rencontre poursuivait trois objectifs :

- Restituer et valider les enseignements des entretiens ;
- Débattre des points sensibles (sécurité, dialogue social, reconnaissance du travail, etc.) ;
- Co-construire des pistes concrètes de sortie du Shadow AI (interdiction temporaire, LLM interne, comptes professionnels agréés, formation & animation de communautés, conventions de confiance...) tout en considérant le capital innovation qu'il constitue.

⁴¹ Les entretiens semi-directifs sont utilisés en sciences humaines et sociales pour explorer en profondeur les représentations, pratiques et logiques d'action des individus, tout en laissant une marge de liberté dans les réponses. Ils permettent de recueillir des discours contextualisés, nuancés et souvent riches, tout en assurant une certaine comparabilité grâce à une trame souple de questions. L'approche par entretiens semi-directifs est particulièrement pertinente pour les terrains d'étude où les phénomènes sont complexes, peu formalisés ou en transformation, et où il est nécessaire de comprendre les logiques d'acteurs "de l'intérieur". Cela inclut notamment les environnements de travail (entreprises, administrations) pour analyser les pratiques professionnelles, les usages informels ou les adaptations aux transformations organisationnelles (comme l'introduction de l'IA).

Quatre thématiques avaient été sélectionnées pour les ateliers qui ont été animés sous la forme d'un « world café » :

- Trouver l'équilibre entre régulation et reconnaissance des initiatives des employés
- Les verrous pour sortir du Shadow AI
- Gérer la tension entre individualisation des pratiques et nécessaire culture collective
- Quelle posture pour les managers de proximité ?

À l'issue de la journée, l'équipe projet a consolidé l'ensemble des matériaux recueillis pour produire un guide incluant :

- Une définition théorique et empirique du Shadow AI ;
- Une quantification croisant études existantes et données de terrain ;
- Une typologie détaillée des usages observés ;
- Une analyse FFOM double (employés / employeurs) ;
- L'identification des verrous (techniques, culturels, sociaux) ;
- Des scénarios et leviers de sortie assortis de recommandations opérationnelles (gouvernance, psychologie du travail, régulation des données, etc.).

Cette enquête croise ainsi observation de terrain et réflexion collective pour éclairer les tensions, les arbitrages et les opportunités liés à l'appropriation spontanée de l'IA générative. En s'appuyant sur l'expérience d'organisations pionnières, elle vise à outiller concrètement les acteurs souhaitant sortir du Shadow AI tout en reconnaissant les dynamiques d'innovation portées par les collaborateurs. Il convient toutefois de préciser que, du point de vue du Shadow AI, notre enquête se limite volontairement à l'audition de membres du top management (DSI, DRH, chefs de projet IA), sans investigation directe auprès des utilisateurs finaux. Par conséquent, il n'est pas possible d'affirmer que les récits recueillis reflètent fidèlement l'expérience vécue par les collaborateurs eux-mêmes. **Notre démarche vise délibérément à documenter le point de vue des directions** : comment elles identifient, interprètent et tentent de répondre aux usages informels d'IA générative au sein de leur organisation. Ce cadrage constitue une première étape, centrée sur les dynamiques institutionnelles en réponse au Shadow AI.

3.2. Enseignements généraux : des initiatives qui accélèrent l'intégration et la gouvernance de l'IA

Depuis l'irruption de ChatGPT à l'automne 2022, l'intelligence artificielle générative s'est invitée dans les organisations leaders par le canal de l'usage individuel, dépourvu de tout parrainage officiel. Cette démocratisation express, rendue possible par des interfaces grand public et gratuites, a introduit une trajectoire d'innovation IA

alternative aux cas d'usage portés les équipes dédiées: elle circule désormais à travers les initiatives des collaborateurs, qui développent leurs propres usages et pratiques "bricolantes".

Les « usages non validés » se multiplient : des collaborateurs recourent à ChatGPT ou à des générateurs d'images sans en référer ni à la DSI ni à la hiérarchie. L'initiative suit un mouvement nettement bottom-up : les premiers cas d'usage sont détectés, testés et raffinés par les équipes de terrain, souvent à l'insu de leur encadrement. Les organisations pionnières se retrouvent alors face à un dilemme : encourager la créativité qui jaillit de la base, levant des verrous sur lesquels les projets IA butaient antérieurement, tout en maintenant la cohérence et la sécurité du système sociotechnique global.

Dans ce contexte, une clandestinité feutrée s'installe. Par crainte de paraître incompetents, d'alourdir la charge IT, ou de franchir des lignes rouges réglementaires, certains dissimulent leurs pratiques. L'ombre protège l'autonomie mais nourrit aussi le risque d'un traitement inapproprié d'informations sensibles. On touche aux enjeux de conformité RGPD, de souveraineté des données et, plus largement, à la confiance, ce capital invisible sans lequel aucune technologie ne prospère durablement.

Cette exploration « hors pistes » stimule l'innovation. À force de bricolage, les employés découvrent des usages métier, documentent empiriquement les limites, se procurent des gisements de productivité auxquels les directions n'avaient pas songé.

La comparaison avec le Shadow IT s'impose : comme lui, le Shadow IA se nourrit de solutions non autorisées. Mais l'IA générative ouvre une dimension supplémentaire : l'automatisation du raisonnement et de la créativité, autrefois réservée aux spécialistes, semble devenir accessible. Le périmètre potentiel du travail « informel » s'élargit donc considérablement.

Interdire purement et simplement ces pratiques s'avère illusoire : nombre de salariés continuent d'y recourir, parfois via des comptes personnels ou des terminaux non contrôlés. Face à cette réalité, les organisations pionnières ont préféré ajuster leur curseur : elles modulent les restrictions en fonction du niveau de risque, s'alignent sur les exigences émergentes (anticipant celles portées par le règlement IA européen) et mettent en place des chartes d'usage plutôt que des interdictions absolues.

À moyen terme, la réponse la plus féconde semble se jouer sur quatre registres : sensibiliser (faire connaître le potentiel comme les limites), réguler (poser des garde-fous clairs et proportionnés), former (donner les compétences nécessaires pour un usage responsable) et échanger (créer des espaces où bonnes pratiques et retours d'expérience circulent). Autrement dit, accompagner plutôt que réprimer, afin de transformer ce Shadow IA en une force de transformation collective, lucide et maîtrisée.



3.3. Une définition élargie du Shadow AI

Les entretiens réalisés permettent d'étendre la définition classique du Shadow IT appliquée à l'intelligence artificielle. Sept types d'usage se dégagent, résumés dans le tableau ci-dessous et détaillés dans les sections suivantes.

FORME REPÉRÉE	DÉFINITION SYNTHÉTIQUE
1 Usage individuel informel & clandestin (solution portée par l'employé – usage personnel)	L'employé recourt à une solution d'IA non validée par l'organisation et garde ses usages totalement secrets.
2 Usage informel légitimé (solution portée par l'employé – usage reconnu)	La solution non validée circule librement : les salariés peuvent mutualiser leurs cas d'usage et leurs « trucs et astuces » avec collègues et managers, sans approbation formelle.
3 Usage informel sécurisé (solution portée par l'employeur – usage reconnu)	L'entreprise met à disposition un outil d'IA validé, mais chaque collaborateur l'exploite librement sans use-case défini ; les retours d'expérience restent éclatés et peu visibles pour le collectif.
4 Usage socialisé (solution portée par l'employeur – usage partagé)	L'outil d'IA est officiellement validé ; les cas d'usage et les bonnes pratiques sont documentés et partagés, l'initiative continue de venir essentiellement des collaborateurs.
5 Usage « tacite » via IA embarquée (solution portée par un fournisseur)	Les employés utilisent une application validée contenant de l'IA sans le savoir restant donc hors du radar des équipes .
6 Usage masquée au client (solution portée conjointement par employeur et employés)	Le personnel exploite un outil d'IA avec l'aval du management mais ne l'indique pas aux parties externes (clients, partenaires)..
7 Shadow learning (auto-formation) (solution portée par un formateur informel)	Les salariés se forment de leur propre initiative aux outils d'IA, créant des « shadow formateurs » internes ou externes et des compétences informelles non cartographiées par l'entreprise.

3.3.1. L'usage individuel informel et clandestin : « ils ont tout intérêt à rester discrets, voire à se cacher »

Ce type d'usage correspond explicitement à la définition convenue du Shadow IT. Il rend compte d'un usage personnel et totalement discret d'une solution d'IA non approuvée par l'organisation, l'employé dissimulant ses pratiques, hors du radar des processus officiels et sans aviser ses collègues et l'encadrement.

“ Je me suis rendu compte qu’il s’en servait vraiment, sur de vrais patients, avec de vraies données, directement, sans aucun filet de sécurité... Il m’a expliqué qu’aux urgences, il faisait leurs comptes-rendus avec ça. ”

Selon les responsables interrogés, une raison majeure de cette clandestinité est l’absence de vision claire et partagée de l’employeur sur l’IA en général, générative en particulier.

“ Quand les gens n’ont pas vraiment compris quelle position leur employeur adopte sur l’IA, ils ont tout intérêt à rester discrets, voire à se cacher. ”

“ Rien n’avait été mis en place. On a laissé faire, et sur le terrain, les équipes se sont emparées du sujet bien plus vite que le national. Résultat : beaucoup d’hétérogénéité... Le directeur général a dû lancer une mission pour enfin donner une vision claire et un cap à tout le monde. ”

Cela a créé le sentiment d’un décalage entre les pratiques des employés et le pilotage des employeurs :

“ On a eu le sentiment qu’il y avait beaucoup de monde dans le train : les employés étaient déjà dedans, mais l’entreprise ne le pilotait pas. ”

Certaines organisations avaient pourtant une stratégie formalisée mais déployée à bas bruit, circonscrite à une communauté d'initiés, craignant les conséquences sociales d'une communication forte autour d'une technologie considérée comme socialement anxiogène par le management.

Selon les responsables, cet usage clandestin peut être motivé par la rétention personnelle des gains de productivité et de la reconnaissance obtenus :

“ La première cause de la clandestinité, c'est souvent l'envie de garder pour soi le gain de productivité. C'est un peu ma petite cuisine perso : je préfère que personne ne vienne me dire quoi faire de plus juste parce que j'ai gagné du temps avec l'IA. ”

“ Il y a aussi l'idée d'entretenir une illusion de performance : en réalité, c'est moi + l'IA, mais si mon manager pense que c'est juste moi, c'est encore mieux ”

Cet usage clandestin pourrait également relever d'un réflexe de préférence pour les outils extérieurs :

“ On sait très bien qu'il y a toujours du Shadow IT : tu peux donner toutes les consignes que tu veux, il y aura quand même du WhatsApp, du Google Drive... et maintenant, du ChatGPT. C'est dans la continuité. ”

“ Certains préfèrent utiliser ChatGPT plutôt que notre 'GPT maison', mais franchement c'est marginal. C'est juste que ce qui vient de l'extérieur paraît toujours mieux que ce qu'on a en interne. ”

L'usage individuel informel et clandestin de l'IA générative est donc d'abord encouragé par le manque de vision claire de l'employeur. Il permet aux employés de conserver à leur seul profit les gains de productivité et l'aura de performance, tout en privilégiant parfois des outils externes jugés plus efficaces que les solutions internes.

3.3.2. L'usage informel légitimé : « On a aussi choisi de faire confiance aux employés »

Ce type d'usage correspond à un déploiement informel d'une solution d'IA non approuvée, dont les employés discutent librement les cas d'usage et s'échangent méthodes et astuces au sein de l'équipe (voire avec leur manager), sans qu'aucune validation officielle n'ait encore eu lieu.

“ On a aussi choisi de faire confiance aux employés. On a expliqué clairement ce que c'était, sans encourager l'usage, mais sans le bloquer non plus. L'idée, c'était de permettre à chacun de découvrir la technologie, de voir concrètement ce que ça fait ”

“ En réunion, ça se passe comme ça : quelqu'un dit 'j'ai nettoyé le son de ma vidéo avec cet outil', un autre découvre une solution, la teste, la recommande, et ça circule. Et à un moment, le manager tranche : 'bon, dans ce cas-là, on prend un abonnement ”

Des espaces de partage plus ou moins formels sont mis en place pour inciter à la désindividualisation des usages, au partage de bonnes pratiques avec les collègues et les services compétents qui les accompagnent.

“ On a fait quelques petites animations... c'était l'occasion de discuter avec les gens, à bâton rompu, et leur dire : “Ah oui, moi j'ai déjà utilisé...” ”

“ On a un fil qui est assez sympa, en interne, où tout le monde peut s’exprimer à tous les niveaux au sujet de l’IA générative ”

Des organisations entreprennent de faciliter cette publicisation des usages, en commençant par les reconnaître et à les valoriser.

“ L’idée, c’est d’être dans une logique de service, pour que les collaborateurs n’aient pas peur de venir nous voir. On essaie de créer un climat de confiance : qu’ils sachent que, quoi qu’il arrive, en passant par nous, ils trouveront une solution pour utiliser l’outil. Pas un refus catégorique. Parce que si on bloque tout d’entrée, ils ne viendront plus nous en parler ”

“ Quand l’employeur manifeste vraiment son intérêt pour l’IA et a été très clair à la fois sur sa stratégie et sa position vis-à-vis du Shadow, cela facilite la libération de la parole, les gens se sentent autorisés ”

“ Il y avait plein de petits cas d’usage. On s’est demandé s’il fallait structurer ça par service — juridique, finance, etc. — mais avec toutes les autres priorités, on a laissé tomber. À la place, on a préféré laisser les équipes tester librement, à condition de les sensibiliser un minimum sur la protection des données. On n’a pas mis de contraintes au départ. L’objectif, c’était aussi de rassurer : permettre un usage décomplexé, sans culpabilité. Il ne s’agit pas de tricher parce qu’on est aidé. Ce serait aussi absurde que de reprocher à un chauffeur de taxi de rouler en voiture sous prétexte qu’autrefois, des gens tiraient les carrioles à la force des bras ou des chevaux. ”

La communication sur les risques de ces pratiques est également une manière indirecte de les reconnaître. Cette méthode a presque systématiquement été déployée par les organisations consultées.

“ La première chose qu'on a faite, c'est rappeler les règles générales qui s'appliquent à tous les outils. On les a réexpliquées, en insistant particulièrement sur les IA de type ChatGPT. La DSI a même envoyé une lettre commune avec le responsable, et dès le début, des messages descendants ont alerté les agents sur les risques, notamment le fait que les données ne sont pas hébergées en France et peuvent être accessibles à certains États. ”

Cette information s'accompagne souvent de deux dispositifs : les chartes et les enquêtes internes sur les pratiques

“ En parallèle de ça, on a adressé tout de suite le deuxième niveau : la charte — la charte d'usage pour nos collaborateurs. C'est un chantier prioritaire : nous avons choisi de ne pas interdire les outils en ligne, mais d'accompagner l'usage, le réglementer, l'encadrer et sensibiliser aux risques, assorti d'une démarche d'acculturation générale. ”

“ On en a profité pour mener une enquête qui était plutôt quantitative, envoyée aux 1 000 cadres et encadrants... Après, on a eu 223 répondants. 36 % des 223 répondants ont officiellement déclaré utiliser l'IA dans un cadre professionnel. ”

“ On a recensé 150 expérimentations, mais la réalité, c'est que... quand vous en recensez 150, vous êtes probablement à 300, 400, 500. ”

Cet usage informel tend, partiellement, à généraliser la culture professionnelle des développeurs informatiques. Ce dirigeant d'une société qui emploie de nombreux informaticiens explique que le « Shadow IT » est culturel dans cette profession. L'enjeu est moins l'interdiction que le partage lorsque la solution expérimentée s'avère performante.

“ En fait, le Shadow IA, c’est souvent une forme de veille technologique personnelle, qui nourrit le cœur du métier. C’est une pratique qui existait déjà avant l’IA générative. Avec tous les nouveaux outils qui sortent chaque semaine, surtout pour les développeurs et data scientists, c’est impossible pour une poignée de personnes — ou pour l’IT déjà très occupée — de tout suivre. Du coup, on s’appuie beaucoup sur les retours d’expérience de chacun dans son domaine ”

Les employés engagent une veille personnelle qui est convertie en capital institutionnel via un chemin déjà structuré.

“ Toutes ces initiatives finissent par remonter au comité technique. Et plus elles circulent par le bouche-à-oreille, plus la demande des équipes grandit — ce qui pousse l’organisation à s’y intéresser davantage ”

“ Dans cette logique, une dynamique de veille active décentralisée a été encouragée : chacun est invité à tester les outils utiles dans son métier et à partager les retours pour en faire profiter le collectif ”

“ On a recensé 150 expérimentations, mais la réalité, c’est que... quand vous en recensez 150, vous êtes probablement à 300, 400, 500. ”

La généralisation de ce modèle au-delà de ce métier semble toutefois limitée dans la mesure où elle repose sur des professionnels avertis concernant les réglementations et les risques.

“ On définit des règles de bonnes pratiques qui laissent aux collaborateurs une certaine liberté pour tester différents outils, tout en étant bien formés à ce qu’il ne faut pas faire : ne pas transmettre de données personnelles ou confidentielles, notamment celles de nos clients. C’est notre cœur de métier, on maîtrise ça parfaitement, c’est presque une règle d’or pour nous ”

L’« usage informel » du Shadow AI marque un début de transition entre la clandestinité solitaire et une expérimentation collective encore hors cadre : les salariés s’autorisent à tester une IA non validée, échangent (ou pas) librement leurs recettes et, ce faisant, esquissent les premiers apprentissages organisationnels. Les organisations qui assument cette zone grise sans la sanctionner, en rappelant les règles de sécurité, en ouvrant des fils de discussion ou des ateliers « prompt », en diffusant chartes et enquêtes internes, commencent à convertir l’informel en laboratoire d’innovation. La reconnaissance explicite de ces pratiques, doublée d’un discours clair sur les risques et la protection des données, a vocation à créer un climat de confiance : la parole se libère, les cas d’usage peuvent potentiellement se mutualiser sans que cela soit particulièrement formalisé.

3.3.3. L’usage informel sécurisé : « des modules développés sans réel accompagnement »

Ce type d’usage correspond explicitement à la définition convenue du Shadow IT. Il rend compte d’un usage personnel et totalement discret d’une solution d’IA non approuvée par l’organisation, l’employé dissimulant ses pratiques, hors du radar des processus officiels et sans aviser ses collègues et l’encadrement.

“ Pas mal de modules ont été développés pour répondre à des problèmes récurrents, mais sans réel accompagnement des métiers. Du coup, les utilisateurs s’en servent de façon opportuniste, en découvrant par eux-mêmes ce qui peut les aider... partiellement. Il n’y a pas vraiment de cadre officiel ni de processus établi ”

La formalisation technique cohabite donc avec des use-cases informels qui restent à l'initiative des collaborateurs. La priorité de cette approche semble être de sécuriser les pratiques des collaborateurs en réduisant le risque de fuite de données sensibles et confidentielles.

“ Pour nos clients dans la banque-assurance, c'est impensable de laisser le Shadow AI se développer librement. C'est pareil dans la défense : zéro Shadow IT. En santé, même logique. En revanche, dans d'autres secteurs, on ouvre un panel d'outils et chacun peut tester dans un cadre défini. ”

Cette formalisation du cadre technique peut être à géométrie variable selon le caractère sensible des données et/ou des clients.

“ On sait que sur les cas d'usage les plus importants, si on met les bons outils entre les mains des gens, avec les bonnes licences, ils continueront à les utiliser comme avant — mais de façon plus sécurisée. Et pour le reste, comme l'innovation fait partie de nos priorités et peut venir de partout, on tient à laisser la possibilité de tester ici ou là. On a une culture assez ouverte, et on sait que les retours viendront naturellement, le jour où un utilisateur dira : « tiens, j'ai trouvé un super outil ”

Enfin, l'usage de solutions validées facilite le suivi de flux, l'usage de métriques qui permettront de mieux identifier les usages à haut potentiel de valeur.

“ On a mis en place un contrat Microsoft et acheté 150 licences M365 Copilot pour les cadres intéressés. Des formations sont prévues, et avec ces outils, ils peuvent utiliser des chatbots en interne, travailler sur leurs mails ou encore accéder à l'espace Copilot dans les environnements documentaires ”

L' « usage informel sécurisé » repose sur des solutions d'IA officiellement homologuées qui minimisent les risques de fuite de données, mais il reste essentiellement initié par les collaborateurs : les cas d'usage sont opportunistes, peu diffusés et ne profitent pas encore pleinement à l'échelle de l'organisation. Cette approche à géométrie variable combine donc un cadre technique et juridique robuste avec une liberté d'expérimentation contrôlée, modulée selon la sensibilité des secteurs et des données.

3.3.4. L'usage socialisé : « présenter ses meilleurs prompts »

Ce type d'usage résulte d'un déploiement collectif d'un outil d'IA homologué par l'employeur ; les collaborateurs documentent et mutualisent activement leurs cas d'usage, scripts et bonnes pratiques, si bien que l'apprentissage se propage « horizontalement ». L'initiative reste néanmoins tirée par le terrain : l'organisation fournit l'infrastructure et le sceau de conformité, mais la diffusion, l'enrichissement fonctionnel et l'acculturation sont animés ou co-animés par les utilisateurs eux-mêmes.

“ On demande à des salariés de présenter, en 15 minutes, leurs meilleurs prompts – ceux qui les ont vraiment aidés à faire quelque chose. Ce qui marche bien, c'est que n'importe qui peut partager : de l'assistante au directeur de marque, tout le monde est concerné. On essaie aussi d'organiser ça au niveau du pays plutôt qu'au niveau corporate, car c'est tout de suite plus proche, dans la langue locale, plus intime. Et ça lève pas mal de freins : les gens se disent 'tiens, je n'y avais pas pensé' ou 'moi aussi j'y ai pensé, mais je n'osais pas le dire'. Et le fait de passer par un cadre officiel aide beaucoup ”

Ce partage des prompts les plus pertinents créés ou identifiés par les métiers peut aussi avoir pour vertu d'harmoniser la valeur de l'IA au sein du collectif du travail, d'éviter le creusement d'un fossé numérique.

“

Chez nous, l'objectif, ce n'est pas qu'un salarié lambda devienne soudain un super collaborateur — il n'y gagnerait pas grand-chose. On n'est pas dans une logique d'avancement administratif, mais en mode collaboratif. L'idée, c'est plutôt que celui qui trouve une bonne idée puisse la faire adopter par tous, et ait envie de la partager

”

L' « usage socialisé » s'appuie sur un outil d'IA homologué et un cadre officiel qui mise sur la dissémination des pratiques individuelles dans un cadre normalisé : les collaborateurs documentent et partagent activement prompts, résultats et retours d'expérience, faisant circuler l'apprentissage de pair à pair. Cette dynamique horizontale diffuse rapidement les bonnes pratiques, harmonise les bénéfices de l'IA dans l'équipe et prévient l'apparition d'un fossé numérique entre utilisateurs.

3.3.5. L'usage « tacite » via IA embarquée : « un logiciel officiel non-maîtrisé »

Dans cette situation, une application fournie par un prestataire intègre subrepticement des briques d'IA. Les collaborateurs se servent de l'outil sans toujours savoir qu'ils activent une IA non homologuée, et l'organisation n'en a qu'une visibilité partielle au risque que les pratiques dévient de sa gouvernance explicite, notamment de ses engagements éthiques. Ce cas de figure n'a été remonté qu'une seule fois mais pourrait potentiellement concerner de nombreuses organisations.

“

Chez nous, l'objectif, ce n'est pas qu'un salarié lambda devienne soudain un super collaborateur — il n'y gagnerait pas grand-chose. On n'est pas dans une logique d'avancement administratif, mais en mode collaboratif. L'idée, c'est plutôt que celui qui trouve une bonne idée puisse la faire adopter par tous, et ait envie de la partager

”

L' « usage tacite » embarquée dans des logiciels fournisseurs révèle un angle mort critique : faute d'information transparente, les équipes utilisent une technologie homologuée dont la brique IA échappe au contrôle de la DSI, fragilise la conformité éthique et expose l'organisation à des risques invisibles. Cet exemple ponctuel souligne donc l'urgence de renforcer la vigilance contractuelle et la traçabilité des fonctions IA intégrées par les prestataires.

3.3.6. L'usage masqué au client : « le client achète surtout du temps humain »

L'usage, approuvé en interne, d'une solution d'IA générative dont la contribution n'est pas révélée aux parties externes (clients, partenaires, fournisseurs). Le contenu livré est présenté comme réalisé entièrement de façon conventionnelle ; l'IA agit en coulisses, souvent pour accélérer la production ou le support, tout en laissant intacte l'apparence de travail humain vis-à-vis du destinataire final. A nouveau, ce cas de figure n'a été révélé qu'une seule fois, mais on peut considérer que la pratique est plus répandue.

“ Ce que le client achète, c'est surtout du temps humain consacré à la production de contenu. Mais avec l'IA, la productivité moyenne des consultants évolue, et les standards de livrables finissent par changer. Dans cet intervalle, il y a clairement une forme d'hystérie. Il m'est déjà arrivé de dire : « Reformule ça, on voit trop que c'est du GPT, ça va se repérer. » Heureusement, on n'est pas une boîte de conseil pure, ça ne représente qu'environ 10 % de nos livrables. On se concentre surtout sur les chiffres. Mais oui, même pour des commentaires en bas de page, il faut rester vigilant ”

Le client pourrait avoir l'impression d'être « escroqué » en identifiant un usage dissimulé de l'IA générative, quand il achète un service qui implique implicitement du temps de production humaine, comme du temps d'analyse par un ingénieur. Masquer l'usage de l'IA devient une bonne pratique pour maintenir des relations harmonieuses avec ses clients : cela élargit la charge morale péjorative des usages de l'IA évoqué en première partie.

Cet « usage masqué au client » mais approuvé en interne maximise la productivité mais déplace le risque : si le subterfuge est découvert, le client peut se sentir trompé, exposant l'entreprise à une rupture de confiance et à une responsabilité morale accrue. **Cette pratique révèle donc la tension entre efficacité et transparence : elle souligne l'urgence d'un positionnement éthique clair sur la divulgation de l'IA dans la chaîne de valeur.**

3.3.7. Le shadow learning : « des espoirs qui ne cadrent pas avec la ligne de l'organisation »

Cette dernière version identifiée dans notre enquête décrit les pratiques d'auto-formation informelles qui se diffusent. En effet, le développement informel des compétences IA par les salariés eux-mêmes, en marge des dispositifs de formation officiels combinés à l'émergence de formateurs et de vidéos en lignes, alimentent des positions d'experts autoproclamés non cartographiés par les organisations. Ces « experts » se construisent une autorité en matière d'IA qui potentiellement concurrencent les équipes en charge de la déployer dans le respect de la stratégie définie par l'organisation.

“ Ce ne sont pas tellement les utilisateurs en Shadow qui m'inquiètent, mais plutôt les Shadow formateurs à l'IA. Ceux qui accompagnent les métiers sans s'inscrire dans une stratégie globale. Résultat : ça diffuse des usages ou des espoirs qui ne cadrent pas avec la ligne de l'organisation, et ça complique la conduite du changement. C'est un peu comme ces influenceurs ou collègues bien intentionnés qui montrent « le prompt magique pour gagner une heure ». Ils font... mais est-ce qu'ils font bien ? ”

Le « Shadow learning » révèle la montée d'une auto-formation informelle autour d'« experts » non référencés ; cette expertise parallèle, libre de tout alignement stratégique, propage des pratiques non validées et peut fragiliser la conduite du changement, d'où l'importance de cartographier et d'encadrer les initiatives d'apprentissage. Cette forme d'initiative, qui accorde une fonction d'expert officieux, diffère sensiblement des démarches de partage de bonnes pratiques décrites dans les modèles deux et trois dans la mesure où elle est plus verticale et échappe davantage à la régulation de l'organisation.

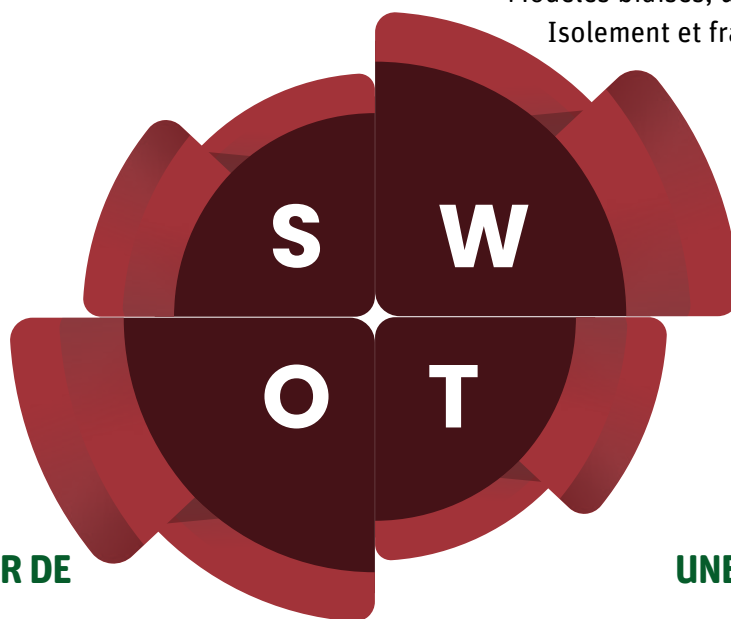
3.4. Trouver l'équilibre entre opportunités et menaces pour les organisations

UNE « ÉTUDE DE MARCHÉ INTERNE »

Accélération de l'adoption et des usages
Exploration préalable aux investissements
Dynamique d'innovation et de créativité individuelle

SÉCURITÉ ET MÉSUSAGES

Risques de non-conformité et de fuite de données
Manque de gouvernance et d'accompagnement
Modèles biaisés, usages mal encadrés
Isolement et fracture des collectifs



ACCÉLÉRATEUR DE STRATÉGIE IA

Identification de cas d'usage métiers
Dialogue et acculturation bottom-up
Formations ciblées et posture partagée
Structuration d'une stratégie IA intégrée

UNE INTEGRATION ANARCHIQUE

Usage de solutions non validées
Failles de cybersécurité et non-conformité légale
Isolement des utilisateurs et désapprentissage
Diffusion de pratiques non maîtrisées

3.4.1. Une étude de marché interne : « qu'est-ce qu'on pourrait faire avec ça ? »

Au-delà de l'engouement initial pour ChatGPT, plusieurs organisations ont transformé ces initiatives en véritable « étude de marché interne » ascendante.

“Aujourd’hui, il se passe plein de choses : les collègues sont curieux, et je trouve ça super de voir que certains, qui étaient assez éloignés de ces sujets, se découvrent une vraie passion pour ces outils. Ça les intrigue, ils testent plein de choses, viennent nous voir en disant : « Qu’est-ce qu’on pourrait faire avec ça ? » Je trouve ça à la fois stimulant et vraiment enthousiasmant

“On a lancé une sorte d’étude de marché interne, en ouvrant une démarche bottom-up pour identifier les besoins et les faire qualifier par les équipes. On avait mis en place un petit CRM avec des notions de business case et de valeur — même si tout cela restait assez incertain, car il est difficile de construire un business case solide juste après l’émergence d’une nouvelle techno sans expérimentation. Certains avaient déjà testé chez eux, lu des articles, fait du benchmarking, et connaissaient bien leurs enjeux, ce qui nous a permis de cartographier les cas d’usage les plus prometteurs. À partir de là, on a défini une stratégie de déploiement par division

Cette même logique d’exploration bottom-up anime des métiers jusqu’alors peu engagés dans l’intégration de l’IA : « ce que ça nous a permis de faire assez rapidement, c’est d’entrer en dialogue avec les métiers des lettres ». Cette accélération par les usages spontanés justifient leur non interdiction.

“Au fond, interdire l’usage de l’IA générative, c’est freiner les opportunités et couper court aux remontées de cas d’usage qui viennent des utilisateurs eux-mêmes. Aujourd’hui, ce qui fonctionne, c’est de laisser les gens dans un espace à la fois libre et responsable — un cadre qui leur donne de l’autonomie tout en les engageant dans une démarche consciente et maîtrisée

En combinant curiosité individuelle et cadrage institutionnel, ces initiatives transforment le Shadow AI en laboratoire, accélérant l'adoption, clarifiant les meilleurs usages et nourrissant une culture d'innovation ouverte au sein de l'entreprise.

“ Il y a une vraie opportunité : faire comprendre aux métiers ce que l'IA peut concrètement leur apporter. Jusqu'ici, c'était un vrai défi – les équipes avaient du mal à expliquer, à faire toucher du doigt l'intérêt de l'IA. Là, on avance vraiment. Ça remet aussi au centre la question : « Quel est mon problème, et comment la donnée peut-elle m'aider à le résoudre ? » Au lieu de fonctionner à l'intuition ou avec des étapes manuelles, un peu au feeling, on commence à se demander comment mieux quantifier et automatiser certaines tâches, et surtout où chacun apporte réellement de la valeur ”

“ On s'est posé une question simple : quels sont les cas d'usage à forte valeur ajoutée, sur lesquels on veut vraiment investir ? On avait déjà une équipe de data scientists et une vingtaine de plateformes, parce qu'on reçoit chaque jour des milliards de lignes de données liées à l'activité de nos clients. Avec une culture data déjà bien ancrée, on s'est demandé où mobiliser au mieux cette équipe pour que ce soit vraiment pertinent. ”

Transformer l'engouement spontané pour l'IA générative en « étude de marché » ascendante a permis de cartographier rapidement les besoins réels, d'identifier les cas d'usage à forte valeur et, en laissant chaque métier expérimenter dans un cadre responsable, de faire du Shadow AI un laboratoire collectif qui accélère l'adoption et renforce la culture d'innovation de l'entreprise.

3.4.2. Des mésusages et des risques : « un vrai défi d'entreprise »

La faiblesse majeure des organisations face à ces pratiques spontanées de leurs collaborateurs réside dans les risques importants en matière de sécurité et la potentielle occurrence de mésusages, principalement liés à la gestion des données.

“ C’est un vrai défi d’entreprise aujourd’hui. Pour l’instant, on s’appuie sur des modèles en lesquels on a confiance, car ils viennent de fournisseurs reconnus avec lesquels on a des contrats. On est donc prescriptifs : on définit les modèles autorisés, même pour les équipes de recherche — pourtant bien outillées et compétentes — on impose un cadre. Pas question d’utiliser un modèle comme DeepSeek, entraîné on ne sait où. Et si, exceptionnellement, un usage devait en être fait, ce serait dans un environnement ultra-sécurisé, totalement isolé. Sur la cybersécurité, on ne transige pas ”

“ Ce qui nous inquiète le plus avec le Shadow AI, c’est le risque d’exposition de données sensibles. Par exemple, si quelqu’un utilise ChatGPT pour reformuler un mail à un client et y entre son adresse, la valeur de son bien ou des détails de contrat, ces données ne sont plus protégées. Et au-delà, il y a aussi une crainte plus large autour du secret industriel : imaginez que des notes destinées à la direction générale soient préparées dans une version publique de ChatGPT... ”

Pour la plupart des organisations auditionnées, le premier rempart reste un contrôle serré des flux et une pédagogie sans relâche. Une organisation assume ainsi une approche de « surveillance bienveillante » :

“ Un employé lambda ne peut pas installer ce qu’il veut : la cybersécurité veille, et on surveille tous les flux de données sortants. On a d’ailleurs envoyé une lettre commune, avec le responsable sécurité, pour alerter sur les risques liés à l’usage de données d’entreprise dans ces outils. On n’a incité personne à les utiliser, mais on n’a pas bloqué l’accès non plus ”

D’autres organisations ménagent la curiosité des équipes tout en balisant strictement le terrain :

“ On a bloqué l'accès aux outils les plus courants via les proxies, tout en précisant : “Si vous voulez expérimenter, des sandbox sont disponibles sur nos fournisseurs cloud.” Même si un collaborateur y mettait des données qu'il ne devrait pas, le risque reste limité, encadré et isolé ”

Les enquêtes internes rappellent toutefois qu'un cadre ne suffit pas si les réflexes ne suivent pas :

“ Le vrai risque, c'est que beaucoup ne se sont pas du tout préoccupés des données personnelles ou confidentielles. Dans notre enquête, seulement 40% disaient y faire un peu attention — donc 60% ne se sont même pas posé la question. C'est inquiétant, surtout quand on sait que certains ont créé des comptes avec leur mail pro, sans qu'on sache s'ils ont sécurisé leur mot de passe... parfois le même que pour leur messagerie perso ”

Pour autant, face aux opportunités que constituent ces pratiques, peu d'organisations choisissent de bloquer l'accès aux solutions grand public : « Il y a des risques juridiques réels si une énorme fuite de données sort un jour, mais interdire serait “se couvrir de ridicule”, comme interdire Google au début des années 2000 ». En outre, certaines organisations qui ont proclamé une interdiction n'ont pas été suivies par leurs collaborateurs.

“ Nous avons donc découvert qu'un nombre important de personnes utilisaient ces outils, parfois en sachant très bien qu'elles n'y étaient pas autorisées. Plus tard, lors d'un échange informel avec OpenAI, ils nous ont mentionné qu'il y avait 1 000 utilisateurs. C'était une information transmise oralement, sans preuve formelle, mais qui illustre bien l'ampleur du phénomène ”

Entre contrôle des accès, offre d'alternatives sécurisées et acculturation continue, la gestion du Shadow AI cherche donc l'équilibre délicat : réduire la surface d'exposition sans casser la dynamique d'innovation, afin d'éviter que la méfiance, ou la clandestinité, n'entrave la confiance collective.

3.4.3. Une intégration anarchique

La menace principale du Shadow AI réside dans le risque d'une intégration anarchique et désordonnée de l'IA dans les processus de travail. L'usage non coordonné et non supervisé des outils externes peut conduire à un manque de « cohérence » et à l'absence d'une « vision globale » de l'IA au sein de l'organisation.

“ Avant 2024, on n'avait pas vraiment de stratégie claire sur l'IA. C'était plutôt une approche du type « que le meilleur gagne », un peu comme les algorithmes génétiques : on semait des initiatives un peu partout, en espérant que certaines prennent. Une approche assez hasardeuse, en somme ”

Cette approche décentralisée rend difficile le contrôle des pratiques, l'assurance de la qualité des productions ou l'intégration robuste dans les systèmes d'information existants.

“ Chacun lance son propre système sur des machines virtuelles un peu partout, sans cohérence ni vision d'ensemble. Or, pour ce type de projet, il faut une approche globale. D'un côté, les besoins en calcul sont importants, vu la taille des modèles et le nombre d'utilisateurs ; de l'autre, il faut aussi une stratégie claire sur les données. Un moteur de prédiction est bien plus utile s'il s'appuie sur des bases de référence communes et validées à l'échelle de l'administration ”

Le risque de « Shadow formateurs » diffusant des pratiques non alignées est également une menace pour la cohérence stratégique.

“

Est-ce qu'il fait dans le cadre éthique et logique qu'on a défini ? Les usages de ceux qui ont développé une expertise, c'est aussi du bottom-up et tout le monde peut devenir expert de l'IA demain, ou presque. En tout cas, c'est improvisé comme ça. Et donc, quelle est la fiabilité ?

”

Une intégration “sans analyse, sans esprit critique ” peut dégrader la qualité du travail produit.

“

Sur les aspects rédactionnels ou techniques, si l'on pousse trop loin l'automatisation, le modèle peut proposer des solutions ou des paquets non standards, incompatibles avec le reste. Cela montre un usage trop passif, sans recul critique. D'où l'importance d'ancrer ces pratiques dans une vraie compréhension métier : ce n'est pas brider les utilisateurs que de leur demander cette ossature de réflexion — c'est justement ce qui leur permettra d'en faire un usage pertinent et autonome

”

“

Il est tentant de demander un avis à ChatGPT et de reprendre sa réponse sans réelle analyse. Mais si tout le monde fait cela, les livrables — qu'il s'agisse d'appels d'offres ou de recommandations — finissent par se ressembler, sans relief. L'enjeu n'est pas d'éviter l'IA générative, mais de l'utiliser avec discernement : savoir chercher les bonnes sources, mener sa propre réflexion, et guider l'outil dans un raisonnement structuré pour produire un résultat à la fois plus riche et plus singulier

”

Le Shadow AI, en représentant l'usage hors cadre, crée un défi pour les organisations souhaitant formaliser, sécuriser et standardiser l'utilisation de l'IA. Les « dérives qui derrière auraient été compliquées à gérer » sont une conséquence potentielle de cette intégration non maîtrisée.

L'intégration anarchique de l'IA fragmente les initiatives, dilue la cohérence technique et éthique : sans vision globale ni esprit critique, l'organisation perd le contrôle de la qualité, des données et des compétences. Seule une gouvernance coordonnée, alliée à une culture d'usage réfléchi de l'IA, peut transformer ces expérimentations dispersées en valeur durable, tant pour les employeurs que pour les employés.

3.4.4. Une accélération de la stratégie IA : « apporter vision et sens »

En réponse à la menace de cette intégration chaotique, l'irruption du Shadow AI a joué le rôle de propulseur : les organisations pionnières ont entrepris de passer du « test sauvage » à une feuille de route explicite et partagée. Le foisonnement bottom-up, « tout le monde doit toucher au sujet », est traduit en priorités métiers :

“ On voit bien que ça peut s'appliquer à plusieurs domaines... On a donc élaboré toute une stratégie de déploiement, et on la confronte maintenant avec les responsables de chaque fonction pour s'assurer qu'on mise sur les cas d'usage les plus prometteurs ”

Cette vague grand public a fait tomber les derniers freins, pour « ne pas rater la dynamique » :

“ Pour les dirigeants, il y a un halo informationnel... du coup c'est devenu une préoccupation au plus haut niveau ; le COMEX a validé un plan global dès avril 2023 »
« Il y a eu une volonté très haute... expérimenter très vite pour ne pas se tromper dans nos feuilles de route et investir là où la valeur est immédiate ”

Une organisation transforme un séminaire d'acculturation en mission nationale dédiée afin d'« apporter vision et sens à toutes les directions métiers ».

Des équipes transverses sont en effet mises en place pour piloter cette transformation, élargissant le portage de l'IA au-delà des expertises purement techniques :

“ Dès février, on a décidé de créer une task force dédiée au sujet, avec un DRH à temps plein dans l'équipe ”

“ On a choisi un pilotage à trois, justement pour éviter que ce soit porté uniquement par la tech. Ce n'est pas un sujet purement technique ; il y a des enjeux humains, sociaux, philosophiques, qu'une approche technocentrée ne saurait couvrir. On ne voulait donc pas laisser la main à la seule direction technique ? ”

“ Le pôle RH et le SI travaillent ensemble pour tester des approches de conduite du changement et organiser des ateliers participatifs. L'objectif, c'est d'aider les équipes à comprendre ce qu'est (et n'est pas) l'IA, et comment elle peut transformer les métiers, avec une vraie dimension d'anticipation. ”

“ La mission de gouvernance IA a mobilisé un large éventail de compétences : experts IA bien sûr, mais aussi profils cloud, cybersécurité (notamment orientée cloud), data governance, data loss prevention, ingénierie, sécurité corporate, ainsi que les fonctions support comme les achats, les RH et le juridique. Le pilotage, soutenu au niveau Comex, a bénéficié d'une forte implication des équipes digital et engineering. Les métiers ont été intégrés progressivement : pas trop tôt, pour laisser aux équipes techniques le temps de poser les bases ; pas trop tard, pour que les cas d'usage soient portés par ceux qui les vivent. Cette montée en puissance a suivi le rythme du déploiement, sans accélération déconnectée du terrain. ”

L'accélération de la stratégie IA s'observe de façon significative dans le déploiement de la formation professionnelle, alors que selon une étude Ifop de 2024, seuls 10% des français ont reçu une formation dans un cadre professionnel. Ainsi, les dispositifs de formation se multiplient et sont parfois obligatoires.

“ Depuis le 1^{er} janvier, il y a un training d'entreprise disponible pour tout le monde ; s'ils n'ont pas fait le training, ils n'auront pas l'accès aux outils. On va couvrir les 200 000 employés du groupe ”

“ Le 1^{er} avril 2024, j'ai créé une formation maison ; on a formé 42 000 personnes en 2024 et notre PDG veut maintenant que 100 % des salariés y passent ”

La formation est appréhendée comme un correctif à l'hétérogénéité des pratiques : « On a développé un module de e-learning obligatoire destiné à tous les collaborateurs ; l'ambition, c'est que chacun atteigne un premier niveau commun de compréhension ». Ces programmes, déclenchés ou accélérés par la découverte d'usages non officiels, traduisent la même conviction : transformer un phénomène incontrôlé en levier d'acculturation généralisée et durable.

Enfin, cette accélération stratégique a pu prendre la forme de « convention » apportant une densité consultative à la feuille de route établie. Après avoir annoncé un moratoire sur ChatGPT, une organisation pionnière a « créé un comité territorial de l'intelligence artificielle auquel nous avons demandé des recommandations ; l'une d'elles a été la mise en place d'une convention citoyenne ». Cette convention doit « discuter de tout ça, élaborer une stratégie données & IA, pérenniser un comité numérique » afin que les futurs choix techniques engagent autant les élus que les habitants.

Une deuxième organisation a choisi la même voie participative : « ... une convention sur le modèle des conventions citoyennes ; on tire au sort 30 salariés et on leur pose la question : « à quelles conditions et selon quelles modalités pouvons-nous intégrer l'IA générative dans nos métiers ? ». Pour la direction, il s'agit avant tout « d'embarquer le corps social dans une transformation qui les inquiète ou les excite, et de montrer que la DG en fait un sujet majeur ». Parmi les points clés soulevés par les panélistes :

- Éviter une surveillance intrusive et favoriser une autonomie réglementée.
- Utiliser l'IA pour réduire la charge cognitive, notamment en automatisant les synthèses d'appels.
- Maintenir un équilibre entre automatisation et charge émotionnelle des téléconseillers.
- Assurer un suivi annuel de l'impact de l'IA sur l'emploi et les conditions de travail en partenariat avec l'Anact (Agence Nationale pour l'Amélioration des Conditions de Travail)

Dans les deux cas, le Shadow AI passe ainsi du statut de pratique informelle et clandestine à celui de laboratoire démocratique : la convention devient l'outil qui convertit la curiosité individuelle en gouvernance collective et renforce la confiance autour de l'IA.

Le Shadow AI a ainsi joué un rôle de révélateur : pour reprendre la main, les organisations pionnières ont entrepris de convertir les pratiques spontanées en capital stratégique, en plans d'action validés, confiés à des task-forces souvent pluridisciplinaires, sortant l'IA du monopole de la tech. Cette formalisation s'est accompagnée d'une massification de la formation – e-learning obligatoire avant d'accéder aux outils – afin d'unifier les pratiques et de sécuriser l'usage. Enfin, plusieurs structures ont ouvert la gouvernance en lançant des conventions qui associent salariés et usagers à la définition des garde-fous et au suivi de l'impact social. Autrement dit, **la pratique informelle et clandestine est devenue un accélérateur : elle a encouragé les organisations à passer de l'expérimentation dispersée à une stratégie IA souvent plus ouverte et négociée.**

3.4.5. Une montée en maturité des organisations provoquée par les initiatives des collaborateurs

Les déferlantes d'usages Shadow ont servi de marchepied vers la maturité IA. Pour objectiver la maturité organisationnelle en matière d'IA, le groupe « Innovation et commercialisation » du Partenariat Mondial pour l'IA⁴² a mis en place un outil, AIMIND 4 SME. Sur le pilier « organisationnel », cinq leviers sont établis :

⁴² Source : <https://ai4sme.cdt.org.rs/aimind-documentation/>. Voir également une version française : <https://packia.fr/aimind/>. Voir également la version française mise en place par le Hub France IA: <https://packia.fr/aimind/>

- Soutien de la direction : de l'absence totale d'engagement à un sponsoring explicite avec budget et feuille de route.
- Niveau de connaissance des employés : part de salariés formés inférieure à 25 % dans les structures « ignorantes », supérieure 75 % dans les « compétentes ».
- Présence de talents IA: aucun expert interne (« ignorantes »), profils maîtrisant modèles, données et/ou intégration dans les environnements de travail (« compétentes »).
- Acceptabilité par les employés : peur d'être remplacés (« ignorantes ») vs perception de l'IA comme outil d'empowerment (« compétentes »).
- Culture de l'expérimentation: tolérance zéro à l'échec (« ignorantes »), encouragement explicite à tester et apprendre (« compétentes »).

La combinaison de ces cinq critères positionne chaque organisation sur un continuum allant d'« ignorante » à « compétente ». En deux ans, tous ces leviers ont connu une montée en puissance remarquable au sein des organisations pionnières. D'abord, les directions ont mis en place des comités de gouvernance pour **piloter** les usages émergents, structurer une stratégie IA transversale et rendre visibles les pratiques spontanées afin d'en évaluer les risques comme les bénéfices. Parallèlement, des parcours de formation obligatoires ou fortement incitatifs ont été lancés, parfois conditionnant l'accès aux outils officiels à la validation d'un e-learning. Dans leur sillage, des communautés d'usage ont vu le jour, illustrant la dynamique de **partage** : les utilisateurs métiers échangent leurs pratiques via des ateliers, des retours d'expérience ou des formations croisées, transformant des expérimentateurs isolés en vivier collectif de compétences.

Cette intelligence distribuée alimente une dynamique de **déploiement** plus structurée : de nombreuses organisations recrutent ou positionnent des experts IA, mettent à disposition des outils sécurisés et accompagnent leur diffusion par une montée en compétences progressive.

Le discours managérial a lui aussi évolué : plutôt que de blâmer l'outil, on valorise désormais les usages maîtrisés, le « travail bien fait avec l'IA », signe d'une acceptabilité croissante dans les équipes. La tolérance initiale vis-à-vis des tests informels – créations de comptes personnels, « semaines de l'IA », ateliers de prompts – a contribué à installer une culture durable de l'expérimentation, dans un cadre où les données sensibles sont mieux cadrées et les usages suivis.

Autrement dit, le foisonnement spontané a fonctionné comme un accélérateur : il a conduit les organisations à avancer sur les dimensions essentielles de la maturité (pilotage stratégique, diffusion des compétences, reconnaissance des talents, acceptabilité sociale et droit d'expérimenter) qui distinguent les organisations encore "ignorantes" de celles désormais qualifiées de "compétentes".

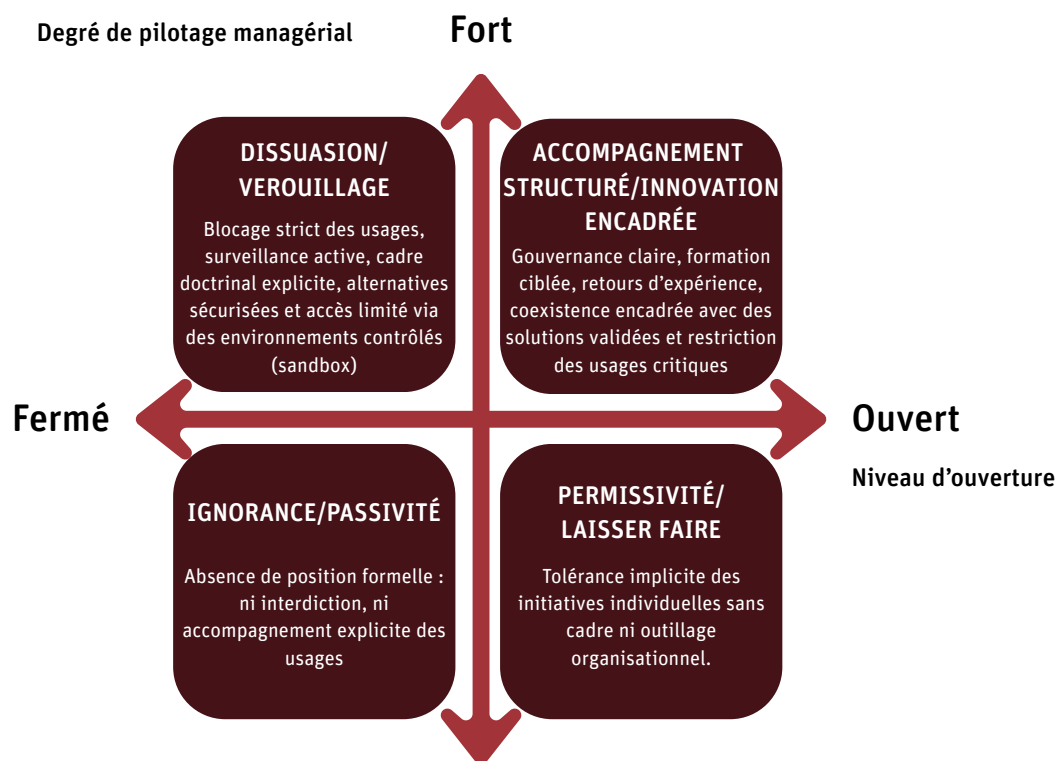
3.5. Quatre postures face aux pratiques spontanées de collaborateurs

3.5.1. Ouverture et pilotage des pratiques informelles

Les organisations pionnières se positionnent face au Shadow AI selon deux critères : le degré d'ouverture aux outils externes et le niveau de gouvernance managériale. Croiser ces axes fait émerger quatre postures, de l'ignorance passive à l'accompagnement structuré, qui servent de boussole pour situer sa pratique et tracer un chemin vers un équilibre entre innovation et maîtrise des risques.

Les postures idéales-typiques adoptées par les organisations pionnières vis-à-vis du Shadow AI se distribuent ainsi autour de deux axes.

- L'horizontale indique le degré d'ouverture vis-à-vis d'outils IA externes ou non officiels.
 - Fermé, on se veut strict : l'entreprise ferme la porte aux solutions non validées.
 - Ouvert, on tolère, voire encourage, l'essai de services extérieurs.
- La verticale représente le niveau de pilotage managérial.
 - Faible, il n'existe ni politique claire ni outillage dédié.
 - Élevé, la gouvernance est formalisée et communiquée : chartes, formations, alternatives "maison", blocage, interdiction.



On obtient quatre zones contrastées qui racontent autant de styles de gestion, leurs risques, leurs atouts et, surtout, les trajectoires possibles d'une organisation donnée. Cette modélisation offre ainsi un point de repère pour situer son état présent, comprendre les tensions qui le traversent et orienter, de façon éclairée, l'évolution vers un équilibre durable entre innovation et maîtrise des risques.

3.5.2. De l'ignorance à l'innovation encadrée

Toutes les organisations pionnières suivent le mouvement de la flèche centrale. Passer progressivement des zones bas-gauche ou bas-droite — où règnent soit le flou, soit la permissivité brute — vers le quadrant haut-droite, où l'organisation capture la valeur de l'IA tout en maîtrisant ses risques. Le temps passé dans chaque zone diffère selon la culture d'entreprise, la sensibilité des données et la pression réglementaire ; l'enjeu est de ne pas s'enliser trop longtemps dans un espace qui compromet soit la sécurité, soit l'innovation. Pour autant, très peu d'organisations pionnières ont adopté la posture « Dissuasion/ Verrouillage » qui semble réservée aux activités critiques et/ou à des cultures organisationnelles traversées par des enjeux de sécurité, de fiabilité et de surveillance particulièrement saillants.

« Il faut créer une solution interne qui répond à 80% des besoins, et fermer l'accès à toutes les alternatives. On choisit la solution d'entreprise, même si elle déplaît parfois. Les équipes sur le terrain souhaitent des outils adaptés à leurs cas spécifiques, mais à une échelle globale, une plateforme commune facilite le partage des connaissances et le support. Ce n'est pas l'optimum local, mais l'optimum global ».

L'objectif le plus communément exprimé est l'accompagnement structuré qui maintient partiellement une certaine tolérance envers les pratiques spontanées dès lors que les collaborateurs sont formés et responsabilisés sur des activités peu sensibles. Cette posture peut être volontaire, quand la direction souhaite conserver les bénéfices de ces explorations, ou subie, lorsqu'elle résulte d'un fatalisme sur l'impossibilité de contrôler totalement les pratiques des collaborateurs.

3.6. Convertir une créativité diffuse en intelligence collective durable et en stratégie robuste

Notre enquête montre que les organisations pionnières n'appréhendent pas le « Shadow AI » comme une anomalie marginale : il constitue la forme première par laquelle l'IA générative pénètre l'entreprise. Les usages se déploient sur un spectre allant de la clandestinité individuelle jusqu'aux pratiques socialisées, en passant par des zones grises, usages informels, sécurisés ou tacites, qui traduisent autant de compromis entre ingéniosité des salariés et inachèvement des cadres officiels. Loin de n'être qu'un risque, ce bricolage révèle des gisements de productivité, d'apprentissages individuels et collectifs que les approches par use-case, circonscrites à des métiers et à des tâches ciblés peinaient à engager. Il expose tout autant l'organisation aux failles de conformité, de souveraineté et de cohérence stratégique.

Face à cette réalité, les réponses ne résident ni dans la prohibition pure ni dans le laisser-faire : les organisations pionnières ont suivi une trajectoire graduée. Elles passent d'abord par un palier de **pilotage** (rendre visibles les pratiques et poser des premières balises normatives pour responsabiliser) ; vient ensuite le **partage** (socialiser les savoir-faire et construire un référentiel opératif commun) ; enfin le palier **sécuriser** (outiller, former et instaurer une boucle de veille qui pérennise l'apprentissage). Cette progression s'appuie sur quatre registres complémentaires : sensibiliser aux enjeux et limites, réguler par des garde-fous proportionnés, former pour développer une compétence responsable, et échanger afin de diffuser les bonnes pratiques.

Ainsi, l'encadrement du bricolage ne vise pas à canaliser une déviance, mais à convertir une créativité diffuse en acculturation et intelligence collective durable et en stratégie robuste. Reconnaître, légitimer puis institutionnaliser les usages réels de l'IA permet de concilier maîtrise des risques et fertilité de l'innovation ; c'est à cette condition que les organisations pionnières estiment pouvoir transformer la ruse individuelle en avantage stratégique partagé.

Cette troisième partie a examiné les réponses concrètes d'un ensemble d'organisations pionnières confrontées aux usages informels de l'IA générative par leurs collaborateurs. Ces réponses s'organisent autour de trois dynamiques : piloter, partager et sécuriser/déployer. D'abord, les organisations cherchent à rendre visibles ces usages spontanés pour mieux en évaluer les risques et les bénéfices. Ensuite, certaines encouragent la mise en commun des pratiques, par le biais d'ateliers internes, de formations croisées ou de communautés d'usage. Enfin, des entreprises amorcent une diffusion plus large d'outils sécurisés et encadrés, associée à une montée en compétence des utilisateurs.

Plusieurs types de postures sont adoptées de façon plus ou moins radicale par les employeurs face au Shadow AI, allant de l'interdiction pure et simple à l'institutionnalisation progressive des usages. Certaines structures passent par des phases d'animation, de légitimation ou de sécurisation, en fonction de leur maturité numérique et de leur culture managériale.

L'ensemble des témoignages montre que ces réponses ne visent pas à brider les usages, mais à les canaliser pour éviter les dérives, — notamment en matière de données sensibles, de traçabilité ou de charge mentale. Elles reconnaissent l'apport des collaborateurs comme premiers expérimentateurs et cherchent à structurer cette intelligence distribuée.

Répondre au Shadow AI ne consiste pas simplement à le supprimer ou à le faire entrer de force dans les cadres existants. Il s'agit, au contraire, de tirer pleinement parti de ce que ces usages spontanés peuvent apporter en termes de valeur, de créativité, de gain de temps ou d'appropriation des outils, tout en réduisant les risques associés à leur informalité.

Ces réponses esquissent des voies nouvelles pour introduire l'IA au travail, moins tournées vers une logique d'innovation technologique descendante que vers une logique d'ancrage dans le travail réel. Elles partent des pratiques effectives, des gestes professionnels déjà modifiés par l'IA, pour accompagner l'outil au plus près des situations vécues.

Là où les démarches classiques d'innovation partent souvent du travail prescrit, en cherchant à standardiser les usages et à augmenter l'organisation dans son ensemble, ces approches plus pragmatiques cherchent au contraire à augmenter les personnes, à renforcer leur autonomie, leur jugement et leur efficacité à partir de ce qu'elles font déjà. Cette bascule d'un pilotage de l'IA par la stratégie à un pilotage par le terrain traduit un changement de paradigme important.

Les enseignements de ces pratiques pionnières ouvrent la voie à un accompagnement plus structuré du Shadow AI, à la fois responsable, inclusif et centré sur le travail. La dernière partie du rapport formule des recommandations concrètes pour aider les organisations à sortir de l'informalité sans casser l'élan des usages. Elle propose aussi des outils et des ressources pratiques pour encadrer, former et faire vivre des usages collectifs, sécurisés et utiles de l'IA générative.

RECOMMANDATIONS ET OUTILS

Après avoir analysé les réponses institutionnelles au déploiement informel de l'IA générative, cette dernière section propose des leviers concrets pour accompagner les organisations dans la régulation, l'appropriation et la structuration de ces pratiques. Plutôt que d'imposer des normes descendantes ou de réprimer les initiatives spontanées, il s'agit ici de dégager un chemin de sortie du « Shadow AI » qui conjugue reconnaissance des usages réels, cadrage collectif et soutien à l'intelligence professionnelle. Ces recommandations s'inscrivent dans une logique de transition : elles visent à faire évoluer des pratiques individuelles et souvent dissimulées vers des usages partagés, maîtrisés et encadrés, dans le respect à la fois des exigences techniques, juridiques et éthiques, et des dynamiques propres à chaque collectif de travail.

4.1 Sortir du Shadow AI ? Des pratiques spontanées à des usages institutionnalisés.

À l'heure où les dispositifs d'intelligence artificielle se disséminent par capillarité dans les organisations, parfois en marge des radars officiels, le phénomène de Shadow IA constitue un observatoire privilégié des tensions entre innovation spontanée et contrôle institutionnel. L'analyse agrégée de nos auditions avec les organisations pionnières permet de baliser un itinéraire en trois paliers et onze étapes du passage d'usages clandestins, parfois perçus comme déviants, vers une gouvernance pleinement assumée. Ces étapes ne forment pas un chemin séquentiel et monolithique. Certaines organisations pionnières ont commencé par sécuriser avant de piloter, d'autres ont mené simultanément les actions.

4.1.1. Piloter : formaliser l'invisible

À ce stade, l'organisation passe d'une multitude de micro-pratiques dispersées à une reconnaissance organisationnelle du phénomène. Un cadre normatif commun se construit : on nomme le risque, on affiche des règles, puis on crée une instance de gouvernance, un cadre cognitif et éthique.

Actions clés	Objectifs
1. Sensibilisation & éducation aux risques	Responsabilisation : Rendre visibles les dangers pour engager un processus de réflexivité individuelle et collective.
2. Communication transparente sur la stratégie IA	Désenclavement symbolique : Lever le voile de la clandestinité en envoyant un signal de légitimité "on peut en parler sans être sanctionné".
3. Équipe de pilotage multi-métier	Institutionnalisation naissante : créer un lieu de traduction entre métiers, data, juridique et RH ; c'est là que se négocie la première "frame" commune de l'IA.

4.1.2. Partager : socialiser les pratiques et les savoirs

Débutent une phase de socialisation et de délibération. Les pratiques se publicisent, elles deviennent des plus légitimes, que l'on partage dans des communautés de pratique. On voit apparaître des rôles d'ambassadeurs qui jouent un rôle de early adopters. L'organisation transforme ainsi une multitude de bricolages individuels en actif expérientiel commun. Plus précisément, l'organisation cherche précisément à créer ce que De Terssac et Chabaud appellent un « référentiel opératif commun » : un socle d'informations et de compréhensions échangées pour synchroniser les actions, anticiper les perturbations et former, le temps du projet, un collectif de travail. En d'autres termes, le partage vise la mise en commun minimale mais indispensable qui permet aux différents acteurs de se coordonner dans un cadre défini, communiqué et intégré.

Actions clés	Objectifs
4. Espaces de partage de pratiques	Communautés de pratiques : Faire circuler les récits d'usage, détecter ce qui crée de la valeur.
5. Reconnaissance des initiatives individuelle	Légitimation symbolique : Sortir les utilisateurs de "l'illégalisme toléré" pour les inscrire dans des carrières reconnues.
6. Groupes d'ambassadeurs & bac à sable	Expérimentation protégée : Un espace liminal où l'on teste, évalue, discute.
7. Dialogue technologique ouvert	Co-production de la confiance : usagers & experts co-définissent les critères d'acceptabilité.

4.1.3. Institutionnaliser et diffuser : Sécuriser

Dernier mouvement : l'institutionnalisation. À ce stade, le Shadow IA cesse d'être un « ailleurs » informel pour devenir un artefact intégré aux routines. La formation massive produit une harmonisation des compétences. Les outils officiels fournissent une infrastructure socio-technique stabilisée. Pour plusieurs organisations, l'objectif est de mettre fin aux différentes pratiques « Shadow », ce qui peut les conduire à bloquer et à surveiller les pratiques des collaborateurs. Outre les enjeux sécuritaires, la qualité du travail est un argument central en faveur d'outils validés considérés comme plus adaptés et plus performants. D'autres organisations acceptent de conserver une part de Shadow, par pragmatisme ou fatalisme, ou pour en conserver les externalités positives.

Actions clés	Objectifs
8. Formation massive	Capacité collective : Homogénéiser les compétences de base pour réduire les asymétries de savoir.
9. Outils corporate	Infrastructure matérielle & normative: Faire migrer les pratiques vers des environnements contrôlés.
10. Accompagnement & support pour la transition	Facilitation : Réduire les coûts de conversion du Shadow vers les outils corporate et maintenir l'engagement dans l'intégration de l'IA.
11. Surveillance & détection des usages	Boucle réflexive permanente : Mesurer, comprendre, ajuster, bloquer; engager le dialogue avec les utilisateurs pour les orienter vers les solutions appropriées.

4.1.4. Un enchaînement cumulatif

Si aucune organisation pionnière n'a suivi scrupuleusement ces onze étapes, les trois paliers se retrouvent systématiquement, avec différents niveaux d'avancement. Certaines organisations sont à la dernière étape, d'autres la visent dans les 12 mois à venir.

Chaque palier est cumulatif :

- Sans gouvernance initiale, le partage reste clandestin ;
- Sans partage, les solutions restent plaquées et inusitées ;
- Sans solutions performantes, le Shadow renaît car les anciens chemins informels restent plus simples.

Ainsi, sortir du Shadow IA n'est pas une simple affaire de conformité technique ; c'est un travail social de traduction, de légitimation et d'institutionnalisation. Ce n'est pas un simple passage de l'illégitimité à la règle ; c'est la construction progressive d'un espace sociotechnique partagé où gouvernance, pratiques et infrastructures se co-constituent. Le palier **Pilotage** rend d'abord visible le phénomène et en fixe les premières balises normatives. Le palier **Partage**, moment charnière, fabrique le « référentiel opératif commun » qui transforme une addition de bricolages individuels en intelligence collective organisée. Enfin, le palier **Sécuriser** installe des dispositifs techniques et des formations qui pérennisent l'apprentissage, tout en maintenant une boucle de veille et d'ajustement. Pris ensemble, ces trois temps montrent que la régulation de l'IA relève moins d'un verrouillage que d'un processus de négociation continu : reconnaître les usages réels, les légitimer, les intégrer sans les figer. L'enjeu n'est donc pas seulement de réduire les risques, mais de cultiver une capacité réflexive durable, condition d'une innovation responsable et d'un collectif de travail toujours capable de se (re)penser face aux évolutions rapides des technologies intelligentes.

Terssac de G., Chabaud, C. (1990), « Référentiel opératif commun et fiabilité », in Leplat, J., de Terssac, G. (dir.), Les facteurs humains de la fiabilité dans les systèmes complexes, Toulouse, Octarès Editions, p. 110-139.

4.2. Libérer les usages par des conflits de qualité

4.2.1. Cartographier les « user-stories »

Chaque participant est invité à décrire une tâche concrète qu'il réalise avec le recours à l'intelligence artificielle générative. Il s'agit de préciser les objectifs assignés à cette tâche, les étapes du processus, les gestes professionnels ordinairement mobilisés, ainsi que les critères implicites ou explicites qui permettent d'en juger la qualité. Cette cartographie ne vise pas uniquement à documenter les pratiques, mais à faire émerger les représentations que chacun se fait du « travail bien fait », révélant ainsi les points de convergence ou de tension entre collègues sur les finalités et les standards du travail.

Questions clés :

1. Quelle tâche précise réalisez-vous avec l'aide de l'intelligence artificielle générative, et quels sont les objectifs que vous cherchez à atteindre ?
2. Quelles sont les étapes et les gestes professionnels que vous mobilisez habituellement pour cette tâche ?
3. À quels critères vous référez-vous, de manière explicite ou implicite, pour juger que le travail est bien fait ?

4.2.2. La valeur recherchée dans le recours à l'IA

Les participants explicitent ensuite les attentes qui ont motivé leur recours à l'IA : cherchent-ils à gagner du temps, à réduire la pénibilité, à stimuler leur créativité, à mieux comprendre des concepts complexes, ou à structurer leur pensée ? Cette étape permet de qualifier les promesses que les outils d'IA font peser sur le travail, en révélant les critères de performance ou de qualité que chacun mobilise dans son jugement d'utilité. Elle engage une réflexion sur les fonctions assignées à l'IA dans l'organisation et les dimensions du travail qu'elle est censée améliorer.

Questions clés :

- Qu'espérez-vous obtenir en ayant recours à l'IA dans cette tâche ? (Objectif principal : gain de temps, réduction de la charge, clarification, stimulation, etc.)
- Quelles dimensions de votre travail souhaitez-vous améliorer grâce à l'IA ? (Ex. : qualité du livrable, efficacité du processus, créativité, compréhension, autonomie...)
- Quels critères utilisez-vous, implicitement ou explicitement, – pour juger que l'IA est utile ou efficace dans ce contexte ? (Ex. : pertinence des réponses, rapidité d'exécution, clarté des idées, confort d'usage...)

4.2.3. La valeur obtenue par le recours à l'IA

Il s'agit ici de confronter les attentes initiales à l'expérience réelle de l'usage de l'IA. Chaque participant analyse ce qu'il a effectivement obtenu : les gains réalisés, les manques ressentis, mais aussi les effets inattendus ou ambivalents. Cette confrontation permet de mesurer l'écart entre valeur recherchée et valeur obtenue, non seulement en termes de résultats, mais aussi de qualité perçue, de charge cognitive ou d'autonomie dans le travail. Ce moment ouvre un espace d'expression sur les désillusions, les ajustements nécessaires ou les satisfactions générées par l'IA.

Questions clés :

- **Qu'avez-vous réellement obtenu en utilisant l'IA pour cette tâche ?** (Gains concrets, apports qualitatifs, améliorations ressenties...)
- **Quelles limites, frustrations ou effets inattendus avez-vous rencontrés ?** (Manques, erreurs, complexité accrue, dépendance, ambivalences...)
- **Quel écart percevez-vous entre ce que vous attendiez et ce que l'IA vous a effectivement apporté ?** (En termes de résultat, d'effort, de qualité ou d'autonomie...)

4.2.4. Les gestes professionnels pour optimiser le recours à l'IA

Les participants décrivent les savoir-faire concrets qu'ils ont mobilisés ou développés pour tirer le meilleur parti de l'IA, en particulier pour réduire l'écart entre le résultat attendu et le résultat obtenu.

Cela inclut des compétences de formulation de requêtes, de sélection ou de reformulation des réponses, mais aussi d'évaluation critique ou d'intégration dans les processus de travail. Il s'agit également d'interroger le coût de coordination induit par l'usage de l'IA : le temps investi pour obtenir un résultat pertinent, les ajustements opérés, la charge mentale associée.

L'analyse de ce « coût-bénéfice » contribue à objectiver les conditions d'un usage professionnel soutenable de ces outils.

Questions clés :

- Quels savoir-faire avez-vous mobilisés ou développés pour améliorer les résultats fournis par l'IA ? (Ex. : formulation de prompts, reformulation des réponses, vérification, intégration au travail...)
- Quels ajustements avez-vous dû opérer pour atteindre un résultat satisfaisant ?
- (Temps passé, itérations nécessaires, détours imprévus...)
- Quel a été le coût (temps, effort, charge mentale) pour obtenir un usage réellement utile et pertinent de l'IA ? (Rapport entre l'investissement consenti et les bénéfices perçus)

4.2.5. La sérendipité : les bénéfices inattendus de l'IA

Les participants sont invités à partager les apports inattendus que l'IA a pu générer dans leur activité : idées nouvelles, effets de clarification, accès à des perspectives ou des formulations inédites.

Cette exploration de la « sérendipité » permet de reconnaître que l'IA peut aussi produire de la valeur hors des objectifs initiaux, en soutenant la créativité, l'exploration ou la reconfiguration des manières de faire. Ces effets positifs, parfois difficiles à anticiper, peuvent nourrir une réflexion collective sur les usages émergents de l'IA et sur les potentialités qu'elle ouvre pour le développement du travail.

Questions clés :

- Quelles idées, perspectives ou solutions nouvelles l'IA vous a-t-elle apportées, sans que vous les ayez recherchées explicitement ? (Ex. : formulations inédites, détournements créatifs, éclairages imprévus...)
- En quoi ces apports inattendus ont-ils modifié votre manière de penser, de faire ou de structurer votre travail ? (Réflexion, posture professionnelle, innovation dans les pratiques...)
- Quels usages émergents de l'IA entrevoyez-vous à partir de ces découvertes ?
- (Nouveaux champs d'application, élargissement des fonctions de l'IA dans l'activité)

4.2.6. Mise en débat collective des référentiels

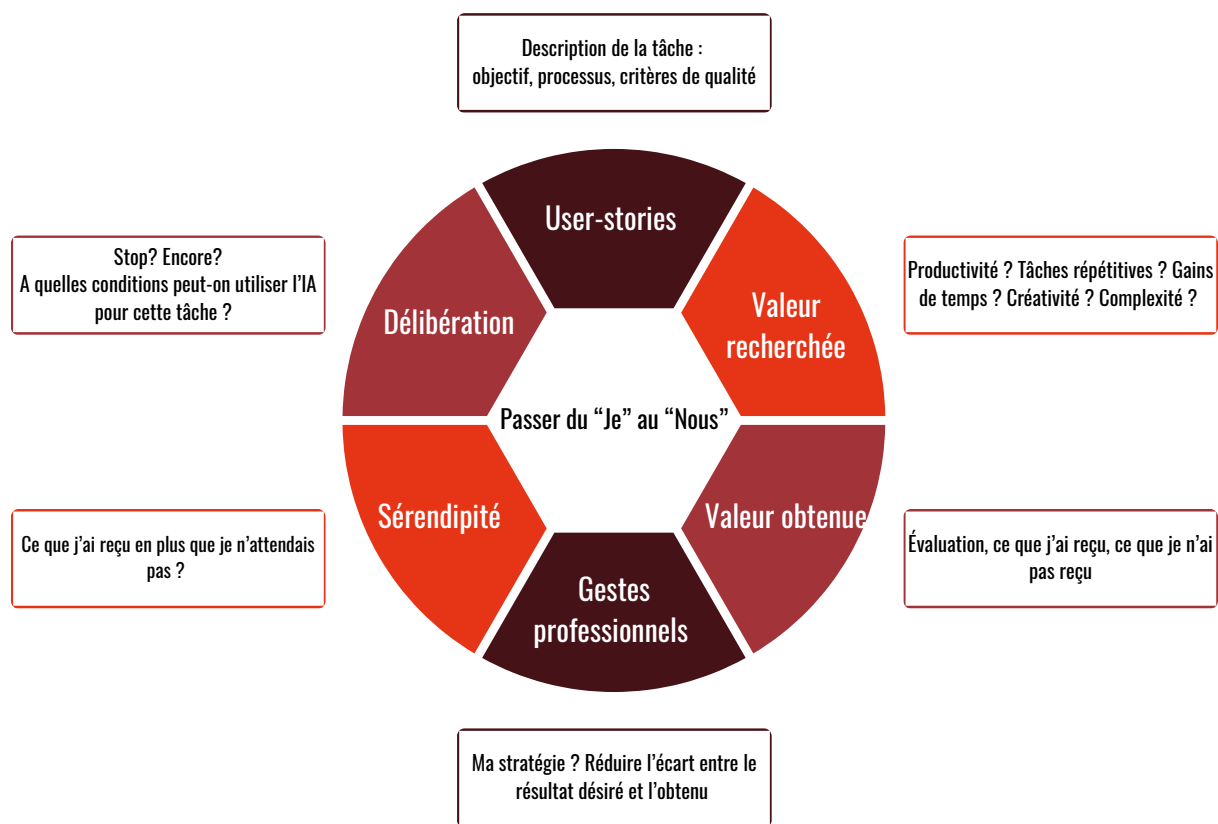
À ce stade de l'atelier, les participants sont invités à confronter leurs points de vue sur les critères de qualité du travail, en s'appuyant sur les tensions exprimées plus tôt. L'objectif est de faire émerger un débat structurant autour de ce qui mérite d'être conservé, transformé ou partagé dans l'usage de l'IA. Cette discussion permet de poser les conditions acceptables d'un usage collectif et légitime de ces outils : dans quelles situations l'IA est-elle utile ou contre-productive ? Quels risques fait-elle courir à la qualité du travail ? Quelles tâches doivent rester à l'humain et pourquoi ? Ce moment vise à construire un langage commun, sans forcément rechercher l'unanimité, mais en valorisant la confrontation argumentée des points de vue.

Sur la base des discussions précédentes, les participants co-rédigent une série de recommandations concrètes pour encadrer ou accompagner l'usage de l'IA dans leur contexte de travail. Ces recommandations peuvent concerner des tâches spécifiques, des conditions d'utilisation (par exemple : supervision humaine, validation, transparence), ou encore des repères sur la qualité attendue. L'enjeu n'est pas seulement de fixer des règles, mais aussi de reconnaître les apports potentiels de l'IA, y compris dans des usages inattendus ou créatifs. Ces recommandations sont autant de jalons pour une gouvernance professionnelle partagée de l'IA au sein de l'équipe.

En rendant visibles les conflits de qualité, cette approche offre un espace pour les verbaliser et les traiter collectivement. Ce processus renforce la capacité des équipes à construire des critères de qualité partagés, à sortir des usages individuels isolés, et à faire de l'intelligence artificielle un levier de coopération plutôt qu'un facteur d'isolement ou de tensions. Il s'agit ainsi d'une étape vers une intelligence collective de l'usage de l'IA, articulée aux réalités du travail bien fait.

Questions clés :

- Quels critères de qualité du travail souhaitez-vous préserver, adapter ou faire évoluer avec l'usage de l'IA ? (Ex. : précision, créativité, autonomie, rigueur, valeur ajoutée humaine...)
- Dans quelles situations l'IA vous semble-t-elle utile, et dans lesquelles devient-elle contre-productive ou problématique ? (Ex. : perte de sens, standardisation excessive, déqualification, perte de temps pour superviser, perte de repères dans l'exécution de la tâche...)
- Quelles règles ou repères collectifs pourriez-vous formuler pour encadrer un usage légitime, éthique et soutenable de l'IA dans votre activité ? (Ex. : tâches réservées à l'humain, supervision systématique, usage collaboratif...)



4.3 Elaborer un cadre de confiance pour un usage responsable de l'IA générative

Plusieurs organisations pionnières nous ont transmis les chartes adressées à leurs collaborateurs en réponse au développement des usages informels de l'IA générative. Ces chartes ont pour vocation de guider un usage éthique, sécurisé et conforme de l'intelligence artificielle générative au sein de l'organisation.

Elle vise à encourager une innovation maîtrisée, garante de la protection des données, la fiabilité des contenus produits, et le respect des réglementations en vigueur, telles que le RGPD et le futur règlement IA. Les utilisateurs sont invités à faire preuve de discernement en validant systématiquement les résultats générés, en protégeant les informations sensibles, et en considérant l'IA comme un outil d'assistance, dont l'utilisation ne saurait remplacer l'expertise humaine. La synthèse de ces documents permet d'élaborer quelques principes généraux applicables par toutes les organisations désireuses de formaliser des principes communs.

4.3.1. Distinguer clairement les usages autorisés et interdits

Les chartes délimitent l'espace où l'utilisation de l'IA générative est tolérée. Sont généralement autorisés, sous réserve de ne traiter aucune donnée confidentielle, les usages visant à tirer parti des capacités de l'IA générative dans un cadre neutre et sans risque pour l'organisation. Il peut s'agir, par exemple, de reformuler ou synthétiser des documents publics afin d'en faciliter la compréhension, de solliciter une aide à la rédaction pour des supports d'information non sensibles, de générer des idées, des plans ou des trames d'entretien en veillant à l'anonymisation complète des situations évoquées, ou encore de créer des contenus fictifs à visée pédagogique ou illustrative.

Afin de préserver la sécurité, la conformité et l'intégrité des activités de l'organisation, les chartes établissent certains cas d'usages de l'IA générative qui sont formellement interdits et doivent être strictement évités, sauf si l'outil a fait l'objet d'une validation par les services compétents. Par exemple, les chartes prescrivent l'interdiction d'utiliser l'IA générative dans des contextes qui exposeraient l'organisation à des risques juridiques, éthiques ou de sécurité. Cela inclut notamment la transmission à un outil d'IA de données personnelles, de santé, stratégiques ou non classifiées, même de manière partielle. Il est également proscrit de produire ou valider, sans relecture humaine, des contenus à portée réglementaire, contractuelle ou engageante. Sont précisés des règles sur l'utilisation de documents internes et le respect des niveaux de classification.

Par ailleurs, les chartes précisent souvent que les contenus générés par IA ne doivent en aucun cas être présentés comme des productions entièrement humaines sans vérification préalable, et toute réutilisation d'œuvres protégées par des droits d'auteur doit faire l'objet d'une autorisation explicite.

Certaines chartes détaillent des exemples concrets se rapportant directement à l'activité les métiers, ainsi que des exemples de formulation de prompts acceptables et inacceptables pour éviter les erreurs courantes.

Les grandes questions à appréhender en vue de définir les usages autorisés et interdits :

1. Quels types d'usages de l'IA générative sont considérés comme sûrs et à faible risque pour l'organisation (ex. : reformulation de contenus publics, aide à la rédaction, génération d'idées anonymisées) ?
2. Quelles sont les catégories de données (personnelles, sensibles, stratégiques, internes non classifiées) dont l'utilisation avec des outils d'IA générative doit être strictement interdite, sauf validation explicite ?
3. Quels critères permettent de déterminer si un contenu généré nécessite une supervision humaine avant diffusion (portée réglementaire, contractuelle, engagement externe...) ?
4. Comment traduire ces principes dans des exemples concrets adaptés aux métiers de l'organisation (usages autorisés/interdits, prompts acceptables et à éviter) ?
5. Quelles règles appliquer à l'utilisation de documents internes ou d'œuvres protégées, et comment s'assurer du respect de la propriété intellectuelle et des niveaux de classification ?

4.3.2. Définir la responsabilité des collaborateurs

Les chartes mettent en avant l'importance de la supervision humaine dans l'usage de l'IA générative : les utilisateurs doivent systématiquement valider les contenus produits, en gardant à l'esprit que l'IA peut générer des informations erronées ou trompeuses sans en comprendre le sens. Cette validation exige un esprit critique et une vigilance professionnelle. Par ailleurs, chaque utilisateur est considéré comme pleinement responsable des instructions qu'il donne à l'IA (prompts) ainsi que de l'utilisation des résultats obtenus, dans le respect des règles de sécurité, de confidentialité et de conformité fixées par l'organisation.

Les grandes questions à appréhender sur les enjeux de la responsabilité des collaborateurs :

1. Comment rappeler que chaque utilisateur reste responsable des contenus générés par l'IA, même lorsqu'ils semblent crédibles ou aboutis ?
2. Quelles règles explicites doivent encadrer la rédaction des prompts (instructions) afin d'éviter toute transmission involontaire d'informations sensibles ou confidentielles ?
3. Quels mécanismes mettre en place pour s'assurer que les utilisateurs valident systématiquement les résultats fournis par l'IA avant leur diffusion ou réutilisation ?
4. Comment encourager l'esprit critique, la vigilance professionnelle et la montée en compétence des collaborateurs face aux limites et aux risques de l'IA générative ?

4.3.3. Prendre en compte les risques spécifiques à l'IA générative

Les chartes alertent sur les risques inhérents à l'usage de l'IA générative qu'il est essentiel d'anticiper et de maîtriser. Les utilisateurs doivent être particulièrement vigilants face aux biais algorithmiques, susceptibles de reproduire des stéréotypes ou de générer des contenus discriminatoires. Le respect de la propriété intellectuelle est également fondamental : il est interdit de créer ou de réutiliser des contenus protégés sans autorisation. L'IA peut aussi être exploitée à des fins de désinformation, rendant nécessaire une vérification croisée des informations produites. Par ailleurs, l'utilisation excessive de ces outils peut entraîner une perte progressive de compétences ou favoriser une paresse cognitive. Enfin, l'impact environnemental lié à l'entraînement et à l'utilisation des modèles d'IA doit être pris en compte, en encourageant un usage mesuré et pertinent des ressources.

Les grandes questions à appréhender sur les risques induits par l'IA générative :

1. Quels dispositifs peut-on mettre en place pour sensibiliser les utilisateurs aux biais algorithmiques et aux risques de générer des contenus stéréotypés ou discriminatoires ?
2. Comment garantir que les contenus générés respectent les règles de propriété intellectuelle, notamment en évitant la réutilisation non autorisée d'œuvres protégées ?
3. Quels principes ou vérifications imposer pour prévenir la diffusion d'informations erronées ou trompeuses générées par l'IA (désinformation) ?
4. Comment prévenir une dépendance excessive à l'IA, qui pourrait affaiblir les compétences professionnelles ou nuire à la réflexion individuelle ?
5. Quelles bonnes pratiques promouvoir pour limiter l'impact environnemental de l'IA, en particulier en encadrant la fréquence et la pertinence des requêtes générées ?

4.3.4 Accompagnement des collaborateurs et évolution continue de la charte

Pour garantir un usage éclairé et sécurisé de l'IA générative, les chartes précisent que les utilisateurs doivent bénéficier d'un accompagnement adapté. Cela passe par des formations obligatoires avant toute utilisation, ainsi que par la mise à disposition de référents internes en cas de doute ou de besoin de clarification. Le rôle des managers est également prépondérant en tant que garant de la mise en œuvre concrète de la charte au sein de son équipe. Il ne s'agit pas seulement de former individuellement les collaborateurs, mais de faire vivre un cadre collectif de confiance, de vigilance partagée et d'apprentissage progressif autour de l'IA générative. Il veille à ce que chaque membre de son équipe comprenne les enjeux éthiques, juridiques et opérationnels liés à l'IA. Il adapte les sessions de sensibilisation aux contextes métiers et favorise les échanges réguliers sur les bonnes pratiques (via par exemple des conflits de qualité, cf. la section précédente). Il coordonne l'évolution des usages de l'IA dans son périmètre, en lien avec les référents techniques et la gouvernance IA. Cela implique de remonter les besoins, les blocages, et les opportunités d'usage identifiés par les équipes, et d'ajuster les pratiques collectives. Pour tout cas d'usage complexe, le manager joue un rôle de filtre, d'alerte et de coordination. Il s'assure, le cas échéant, de la conformité de certaines pratiques, sollicite l'avis de la gouvernance IA, et mobilise les experts concernés. Ce travail inclut l'évaluation des risques, y compris ceux liés aux prestataires ou outils tiers. Le manager instaure un climat propice à la responsabilité individuelle, mais aussi collective, et à la remontée des doutes ou des erreurs. Il valorise les initiatives d'équipe qui visent à documenter, corriger ou améliorer l'usage des outils IA, et fait de la charte un outil partagé, vivant et discuté.

Les grandes questions à appréhender sur les enjeux de l'accompagnement des collaborateurs :

1. Quelles formations initiales et continues doivent être proposées pour garantir que chaque collaborateur utilise l'IA générative de manière éclairée, conforme et sécurisée ?
2. Quels dispositifs d'accompagnement (référents IA, ressources internes, FAQ, glossaire) faut-il mettre en place pour répondre aux doutes et besoins des utilisateurs dans leur quotidien ?
3. Quel rôle concret attribuer aux managers pour qu'ils deviennent des facilitateurs de bonnes pratiques et des animateurs d'un cadre collectif autour de l'IA générative ?
4. Comment organiser, au sein des équipes, une dynamique de veille partagée, de remontée des blocages et de co-construction des pratiques autour de l'IA (ex. : retours d'usage, cas complexes, erreurs identifiées) afin de sortir d'une pure logique de responsabilité individuelle ?

4.3.5. Modalités de pilotage et gouvernance de la charte IA générative

La réalisation d'une charte d'usage de l'IA générative a donné lieu à des approches de pilotage diverses selon les organisations. Certaines ont opté pour une conduite principalement technique, portée par les directions des systèmes d'information et de la cybersécurité, avec une forte attention à la conformité et à la sécurisation des outils. D'autres ont engagé une démarche plus collaborative, interdisciplinaire et multimétier, en associant les fonctions SI, RH, juridique et les métiers, afin d'aligner la charte sur les réalités opérationnelles et les enjeux humains. Quelques organisations sont allées plus loin encore en co-construisant la charte directement avec les collaborateurs, sous forme de conventions salariées, dans une logique de démocratie d'entreprise.

La recommandation qui se dégage des retours d'expérience est de privilégier une approche interdisciplinaire et ouverte, incluant des représentants des usages, sans court-circuiter les instances de représentation du personnel. Celles-ci doivent être informées, consultées et associées tout au long du processus, dans le respect du dialogue social technologique. Enfin, Il est également essentiel de faire de la charte un document vivant : elle doit être diffusée largement dans des formats accessibles (version synthétique, FAQ, glossaire), et faire l'objet de mises à jour régulières afin de rester en phase avec l'évolution rapide des technologies, des usages et des obligations réglementaires.

Les grandes questions à appréhender sur les enjeux du pilotage et de la gouvernance de la charte :

1. Quels acteurs internes (SI, RH, juridique, métiers, cybersécurité...) doivent être mobilisés dès le départ pour garantir une approche interdisciplinaire et équilibrée ?
2. Comment intégrer les collaborateurs dans la démarche de conception, afin qu'ils puissent exprimer leurs besoins, leurs usages et leurs préoccupations liés à l'IA générative ?
3. Quelles modalités prévoir pour associer les instances de représentation du personnel (CSE, CSSCT, etc.) à l'élaboration, à la validation et au suivi de la charte ?
4. Comment articuler les exigences techniques (sécurité, conformité) avec les enjeux humains, organisationnels et métiers dans une logique de co-construction ?
5. Quel dispositif de gouvernance mettre en place pour assurer la transparence, la remontée des usages, et l'actualisation régulière de la charte dans le temps ?

⁴⁸Le projet DialIA (Pour un dialogue social au service des bons usages de l'IA et d'une nouvelle étape de progrès social dans les entreprises et les administrations) élabore un cadre méthodologique permettant de traduire, de manière concrète et opérationnelle, l'accord-cadre européen de 2020 sur la numérisation du travail en levier de transformation, en s'appuyant sur le développement du dialogue social technologique au sein des organisations. Il vise un juste équilibre entre la promotion de ces technologies et l'assurance d'un développement de systèmes fiables. <https://dialia.alwaysdata.net/>

4.4 Une formation holistique à l'intelligence artificielle

Face à la généralisation rapide de l'intelligence artificielle générative dans les milieux professionnels à l'initiative des collaborateurs, les entreprises pionnières ont accéléré et étendu leurs dispositifs de formation à l'IA. Leurs témoignages plaident en faveur d'une formation holistique à l'IA, qui contraste avec la pensée commune, le « Shadow learning » ou même le besoin spontanément exprimé des collaborateurs en matière de prompt engineering — c'est-à-dire l'art de formuler des consignes à des IA génératives. Cette compétence, bien qu'utile pour débiter, démystifier ou stimuler l'intérêt, reste partielle et éphémère. Partielle, car elle ne couvre qu'un minuscule fragment des savoirs et compétences nécessaires à un usage maîtrisé, critique et collectif de l'IA ; éphémère, car les interfaces d'IA intègrent progressivement nativement des fonctions de reformulation automatique et d'assistance contextuelle qui rendent superflue la maîtrise fine des prompts complexes⁴⁹.

Réduire la formation à cette seule dimension, c'est courir deux risques majeurs. D'une part, entretenir l'illusion d'une autonomie technique individuelle là où les usages réels exigent une coopération sociale, une éthique professionnelle et une intelligence collective. D'autre part, contribuer à l'infantilisation numérique des collaborateurs enfermés dans un rôle de simples opérateurs d'interface.

C'est pourquoi il est essentiel de promouvoir une approche holistique de la formation à l'IA, inspirés, par exemple, des cinq axes proposés par le « Artificial Intelligence Competence Evaluation Test »⁵⁰ : savoirs fondamentaux, domaines d'application, méthodes, enjeux juridiques et éthiques, et culture générale de l'IA. Une telle approche outille les collaborateurs pour devenir non seulement des utilisateurs compétents, mais aussi des acteurs critiques, capables de porter des usages responsables et encapacitants, porteurs de sens, et au service de leur propre développement et de la performance robuste de leur organisation.

⁴⁹ «Le Wall Street Journal» titrait même récemment a propos du métier de «prompt ingénieur»: «The Hottest AI Job of 2023 Is Already Obsolete» (avril 2025). Deux arguments étaient développés. D'une part, les modèles d'IA sont désormais capables de mieux interpréter les intentions des utilisateurs et de poser des questions de clarification, ce qui diminue le besoin d'une expertise poussée en rédaction de prompts. D'autre part, les entreprises forment largement leurs équipes à l'usage de ces outils, rendant moins nécessaire le recours à un expert unique.


⁵⁰ Au début de l'année 2024, l'Université de Montpellier, la Métropole de Montpellier ainsi que les entreprises Numalis et BionomeeX ont lancé, en collaboration avec l'Afnor, une spécification visant à encadrer le tout premier test standardisé d'évaluation des compétences en intelligence artificielle. Baptisé AICET (Artificial Intelligence Competence Evaluation Test), ce test a été conçu pour répondre à la demande croissante des entreprises, confrontées à la montée en puissance des IA génératives et à l'entrée en vigueur imminente du règlement européen sur l'IA (AI Act). La spécification officielle a été publiée ce vendredi 4 octobre.

Mais cette formation ne peut être réduite à une simple mise à niveau technique via un test individuel standardisé. Elle doit s'ancrer dans le réel du travail et dans les conflits de qualité que l'IA ne manquera pas de faire émerger : sur le sens des tâches, la répartition des responsabilités, la transparence des décisions, la continuité des collectifs. Loin de les éviter, ces conflits doivent être pris en charge collectivement, dans des espaces de discussion entre collaborateurs, représentants, managers et directions, à l'image de ce que propose déjà le rapport sous forme de cartographies des user-stories et de mise en débat des référentiels dans les conflits de qualité.

Former à l'IA, c'est donc former à un nouveau rapport au savoir, au temps, au collectif et au travail bien fait. Il revient aux collectifs de travail de définir ensemble, à partir de leurs expériences, ce qu'ils attendent de ces outils, ce qu'ils acceptent d'en déléguer, et ce qu'ils souhaitent, au contraire, préserver comme cœur de leur métier. Les organisations qui réussiront ne seront pas celles qui auront le plus vite enseigné à leurs collaborateurs comment parler à une machine, mais celles qui auront permis à leurs équipes de parler ensemble du travail, de ce qu'il devient avec l'IA, et de ce qu'il pourrait être.

Les grandes questions à appréhender pour établir un plan de formation holistique à l'IA :

1. Quelles compétences développer au-delà du simple usage des prompts ? Cela inclut des savoirs fondamentaux sur le fonctionnement des IA, les enjeux juridiques et éthiques, ainsi que la capacité à évaluer la fiabilité des réponses et à exercer un esprit critique.
2. Comment articuler cette formation avec les réalités du travail et les métiers existants ? L'enjeu est de relier les apprentissages à des situations concrètes, à des gestes professionnels réels et aux arbitrages que les collaborateurs rencontrent déjà dans leur quotidien.
3. Quels espaces donnés aux collaborateurs pour débattre collectivement de l'usage de l'IA dans leur activité ? La formation doit s'accompagner d'un espace de délibération où les conflits de qualité du travail avec l'IA peuvent être rendus visibles, discutés et arbitrés.
4. Comment éviter que la formation renforce les inégalités numériques ou générationnelles au sein de nos équipes ? Cela suppose de penser des formats et des rythmes différenciés, qui permettent à chacun, quel que soit son niveau initial, de progresser à son rythme.
5. Comment la formation équilibre-t-elle les schémas délégatifs et augmentatifs : automatiser pour gagner du temps ou augmenter pour enrichir les compétences ? La réponse à cette question oriente la posture de l'organisation et la nature des formations à proposer.



Les éléments présentés dans cette section ne constituent pas des recettes prêtes à l'emploi, mais doivent être compris comme des **paradigmes**, c'est-à-dire des cadres de pensée cohérents pour concevoir collectivement des choix et des modes d'action. Dans le contexte du Shadow AI, ces paradigmes offrent aux organisations des repères pour **bricoler leurs propres schémas régulateurs et stratégies de déploiement**, en tenant compte de leur culture interne, de leur maturité numérique, de leurs équilibres sociaux et de la nature de leurs activités. Ils ne visent ni l'uniformisation des pratiques ni l'imposition d'un modèle unique, mais proposent un espace de réflexion structurant pour penser l'intégration de l'IA générative et de l'IA en général dans le travail réel. Ces outils doivent enfin être mobilisés **en complément d'un dialogue social technologique**, seul à même de garantir que les arbitrages opérés soient à la fois soutenables, négociés et porteurs de sens pour les collectifs concernés.

CONTRIBUTEURS

Merci à l'ensemble des organisations auditionnées :

- Airbus
- Believe
- Caisse Nationale Assurance Maladie
- CHU Montpellier
- Crédit Agricole
- Ekimetrics
- L'Oréal
- MAIF
- Malakoff Humanis
- Ministère des Armées
- Région Ile-de-France
- Skyrock
- Veepee
- Ville de Montpellier

Et aux organisations qui ont participé à la journée de travail du 2 avril, avec les entreprises pionnières auditionnées :

- CFE- CGC
- Dataiku
- Egis
- Paprec
- UCANSS

Conception graphique : Ambre Laubier, Claire Malicorne pour datacraft

Identité visuelle : Jérôme Pidoux - [À propos de Jérôme Pidoux](#)

contact@datacraft.paris