

Les dangers du Shadow AI sur la protection des données

Shadow AI en entreprise : enjeux, risques et stratégies de protection des données

Travail de Bachelor réalisé en vue de l'obtention du Bachelor HES

par :

Samuel Kappel

Conseiller au travail de Bachelor :

Leclère Dessislava, Enseignante à la HEG

Genève, 13.08.2025

Haute École de Gestion de Genève (HEG-GE)

Filière Informatique de Gestion

Déclaration

Ce travail de Bachelor est réalisé dans le cadre de l'examen final de la Haute école de gestion de Genève, en vue de l'obtention du titre Bachelor of Science HES-SO en Informatique de gestion.

L'étudiant a envoyé ce document par email à l'adresse remise par son directeur de mémoire afin qu'il l'analyse à l'aide du logiciel de détection de plagiat COMPILATIO.

L'étudiant accepte, le cas échéant, la clause de confidentialité. L'utilisation des conclusions et recommandations formulées dans le travail de Bachelor, sans préjuger de leur valeur, n'engage ni la responsabilité de l'auteur, ni celle du conseiller au travail de Bachelor, du juré et de la HEG.

« J'atteste avoir réalisé seul le présent travail, sans avoir utilisé des sources autres que celles citées dans la bibliographie. »

Fait à Genève, le 13.08.2025

Samuel Kappel

Remerciements

Je tiens à remercier l'ensemble des entreprises ayant participé à mon enquête, dont les réponses ont grandement enrichi ce travail.

Mes remerciements vont également à ma professeure, Mme Leclère, pour ses conseils et son accompagnement tout au long de ce projet.

Je souhaite également remercier l'outil ChatGPT, développé par OpenAI (versions GPT-4 et GPT-5), qui m'a aidé à reformuler certaines phrases et à structurer mes idées de manière plus claire, tout en veillant à préserver mon style d'écriture et la cohérence globale du document.

Résumé

Ce travail traite du Shadow AI, l'utilisation d'outils d'intelligence artificielle sans validation officielle en entreprise. Le sujet, particulièrement d'actualité, est analysé sous l'angle des risques, des obligations légales et des solutions adaptées, avec un accent sur la réalité des petites et moyennes entreprises (PME).

Après avoir défini les notions clés (données sensibles, LPD et RGPD), l'étude met en évidence les menaces liées au Shadow AI, fuite de données, non-conformité réglementaire, biais, décisions erronées ou perte de contrôle sur l'information. Les tendances récentes, comme le Bring Your Own AI (BYOAI), accentuent ces risques.

Les stratégies de prévention présentées combinent des mesures organisationnelles (chartes, politiques internes, sensibilisation) et techniques (pare-feux nouvelle génération, systèmes DLP, plateformes de gouvernance telles que Microsoft Purview). Le témoignage d'un responsable IT de PME illustre les choix possibles pour limiter ces usages non encadrés.

Une enquête menée auprès de 11 entreprises montre que, si la majorité connaît le Shadow AI, peu disposent d'outils techniques avancés pour s'en prémunir. Les chartes d'usage sont fréquentes, mais leur respect n'est pas garanti, et les solutions techniques restent inégalement mises en place.

En conclusion, la maîtrise du Shadow AI passe par un équilibre entre innovation et sécurité, associant règles claires, sensibilisation continue et outils techniques adaptés aux moyens de chaque entreprise.

Table des matières

Déclaration.....	i
Remerciements	ii
Résumé	iii
Liste des tableaux	v
Liste des figures.....	v
1. Introduction.....	1
2. Données sensibles et Shadow AI.....	2
2.1 Les données sensibles	2
2.2 L'impact de l'AI sur la gestion des données	7
2.3 Le shadow AI.....	10
3. AI et protection des données	13
3.1 LPD, RGPD et autres réglementations.....	13
3.2 Responsabilités des entreprises face aux données personnelles	16
4. Les risques du Shadow AI.....	22
4.1 Risques de fuite et exposition des données sensibles	22
4.2 Conséquences légales et financières	26
5. Sécurisation des données en entreprise.....	30
5.1 Stratégies de prévention et bonnes pratiques	30
5.1.1 Solutions techniques avancées pour la gestion du Shadow AI	33
5.2 Réglementation interne et outils de contrôle	37
5.3 Méthodologie de l'enquête	43
5.4 Analyse des réponses	44
6. Conclusion	52
Bibliographie	53
Annexe 1 : Résultats de l'enquête sur le Shadow AI en entreprise.....	57

Liste des tableaux

Tableau 1 : Différences entre le Shadow IT et le Shadow AI	11
Tableau 2 : Typologie des risques.....	24

Liste des figures

Figure 1 : Classification des données.....	3
Figure 2 : Circulation et protection des données dans l'entreprise (cycle de vie d'une donnée en entreprise)	6
Figure 3: Fonction du DPO.....	18
Figure 4: Chaîne des conséquences d'un incident Shadow AI	28
Figure 5: Le moindre privilège	32
Figure 6: Ensembles des méthodes hybrides	34
Figure 7 : Elaboration d'une politique d'utilisation structurée	38
Figure 8 : Le Next Generation Firewall	41

1. Introduction

Dans le cadre de l'obtention de mon bachelor en informatique de gestion, j'ai choisi d'étudier l'utilisation du Shadow AI en entreprise. Mon intérêt pour la sécurité informatique m'a poussé à explorer ce sujet, car il touche aux données sensibles et aux défis liés à l'intelligence artificielle en milieu professionnel. De plus, le Shadow AI représente une problématique d'actualité que de nombreuses entreprises ne maîtrisent pas encore totalement.

Pour simplifier la lecture, j'ai choisi d'utiliser systématiquement l'acronyme anglais AI, y compris lorsque je parle d'intelligence artificielle de manière générale. Il est important de préciser dès le début que le Shadow AI désigne l'utilisation d'outils d'intelligence artificielle au sein d'une entreprise sans validation ou supervision officielle, souvent à l'insu du service informatique. Sa définition complète et ses enjeux seront détaillés plus loin, mais cette clarification permet déjà de situer le sujet.

Dans ce travail, je commencerai par définir les notions de données sensibles et de Shadow AI. Ensuite, j'aborderai la question de l'intelligence artificielle et de la protection des données, en mettant l'accent sur la LPD (Loi sur la protection des données) en Suisse et le RGPD en Europe. Dans la troisième partie, j'analyserai les risques liés au Shadow AI, notamment les fuites de données et leurs conséquences. Enfin, je proposerai des stratégies de prévention pour limiter ces risques et je réaliserai une enquête auprès d'entreprises afin de comprendre les mesures qu'elles mettent en place pour s'en prémunir.

Pour mener à bien ce travail, j'ai choisi Mme Leclère comme conseillère, en raison de son expertise en protection des données et en législation. J'ai eu l'opportunité de suivre tous ses cours sur ce sujet, ce qui m'a permis d'en apprendre énormément et de développer un réel intérêt pour cette thématique.

2. Données sensibles et Shadow AI

2.1 Les données sensibles

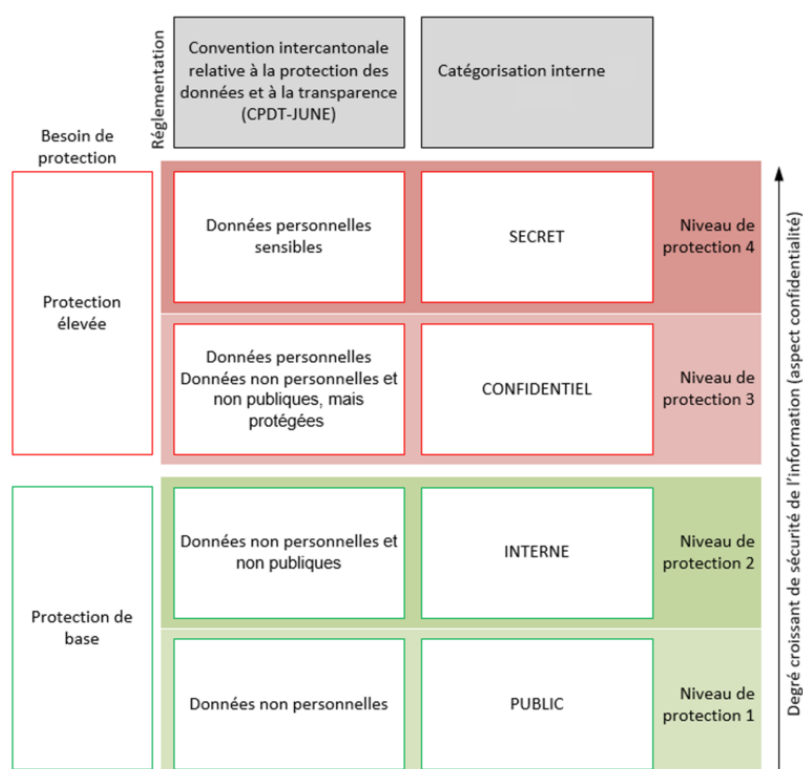
En informatique, les données désignent des informations interprétables et exploitables par des ordinateurs. Elles prennent des formes variées, comme des chiffres, des textes, des images ou encore des mesures collectées par des capteurs. Ces données peuvent être saisies par l'utilisateur, générées par des algorithmes ou issues d'observations. Généralement, elles sont stockées de manière électronique, dans des fichiers ou des bases de données. Pour les entreprises, ces informations représentent un enjeu majeur, car elles servent notamment à la gestion de la clientèle, mais deviennent sensibles dès lors qu'elles contiennent des éléments personnels comme des noms, des adresses ou des images d'identification (LENOVO [sans date]).

Les données dites sensibles sont des données personnelles qui ne sont pas censées être partagées avec n'importe qui. En effet, ces informations relèvent de la sphère intime et privée et peuvent être utilisées à mauvais escient. Parmi celles-ci, on trouve des éléments comme les opinions politiques, les activités religieuses, les informations de santé (telles que les maladies génétiques ou graves) ou encore les poursuites et sanctions pénales. Ces données sont strictement personnelles et ne doivent jamais être rendues publiques, conformément à l'aspect juridique défini par la Loi sur la protection des données (LPD) en Suisse (UNIVERSITÉ DE GENÈVE [sans date]).

Ces informations sont utilisées dans plusieurs secteurs. Par exemple, dans les hôpitaux, les données médicales doivent être particulièrement bien protégées ce qui implique des mesures techniques (sécurisation des systèmes, accès restreint), car toute fuite pourrait avoir de graves conséquences pour le patient, comme une utilisation malveillante sous forme de chantage. Les banques traitent également des données personnelles afin d'identifier leurs clients, notamment en leur demandant leur âge ou leur adresse (UNIVERSITÉ DE GENÈVE [sans date]).

Nous partageons quotidiennement des données personnelles dans notre société, ce qui rend leur protection encore plus essentielle. Il est toutefois possible de partager certaines de ces informations, notamment dans le cadre de recherches ou de sondages, à condition de les anonymiser. L'anonymisation est une mesure technique qui permet de rendre les données non identifiables et donc non exploitables à des fins indésirables (UNIVERSITÉ DE GENÈVE 2018).

Figure 1 : Classification des données



(PPDT [sans date])

Sur cette image, issue du Préposé à la protection des données et à la transparence du canton de Neuchâtel, on distingue clairement la différence entre une protection élevée et une protection de base pour les données. Comme expliqué précédemment, les données anonymisées ou publiques, utilisées par exemple dans les sondages, se situent au niveau de protection 1 ou 2. À l'inverse, plus les informations sont personnelles ou sensibles, plus le niveau de protection exigé augmente, atteignant les catégories « confidentiel » ou « secret » (niveaux 3 et 4). Ce classement visuel permet de comprendre rapidement pourquoi il est crucial d'adapter les mesures de sécurité au degré de confidentialité des données traitées dans l'entreprise.

Une partie des données à ne surtout pas négliger concerne celles que possèdent les entreprises sur leurs projets, leurs stratégies commerciales, leurs analyses ou encore les designs de futurs produits. Ces informations sont primordiales pour le développement d'une entreprise, car elles lui permettent de conserver un avantage concurrentiel. En cas de vol de ces données, les entreprises concurrentes pourraient s'en inspirer pour rattraper leur retard ou, pire encore, dépasser l'entreprise initialement innovante. Dans ce contexte, la protection de ces données va bien au-delà des obligations légales, elle devient un enjeu stratégique pour préserver la compétitivité et l'avenir même de

l'organisation. C'est pourquoi ces données sont considérées comme hautement sensibles, bien qu'elles ne relèvent pas toujours de la sphère des données personnelles (Ettredge, Guo et Li 2018).

Ces données sont privées et ne sont jamais divulguées en dehors de l'entreprise, ce qui les rend très différentes d'un brevet. Dans le cas d'un brevet, toute ressemblance trop marquée peut entraîner des poursuites judiciaires ce qui relève de l'aspect juridique. Mais pour les données confidentielles qui ne sont pas rendues publiques, si elles sont volées, l'entreprise victime n'a souvent que peu de recours, à part porter plainte, ce qui n'empêche pas forcément leur exploitation (Ettredge, Guo et Li 2018).

Leur valeur stratégique attire naturellement l'attention des hackers. Parfois, une simple mention ou un indice sur l'existence d'un avantage concurrentiel suffit à faire de l'entreprise une cible potentielle pour des attaques informatiques (Ettredge, Guo et Li 2018).

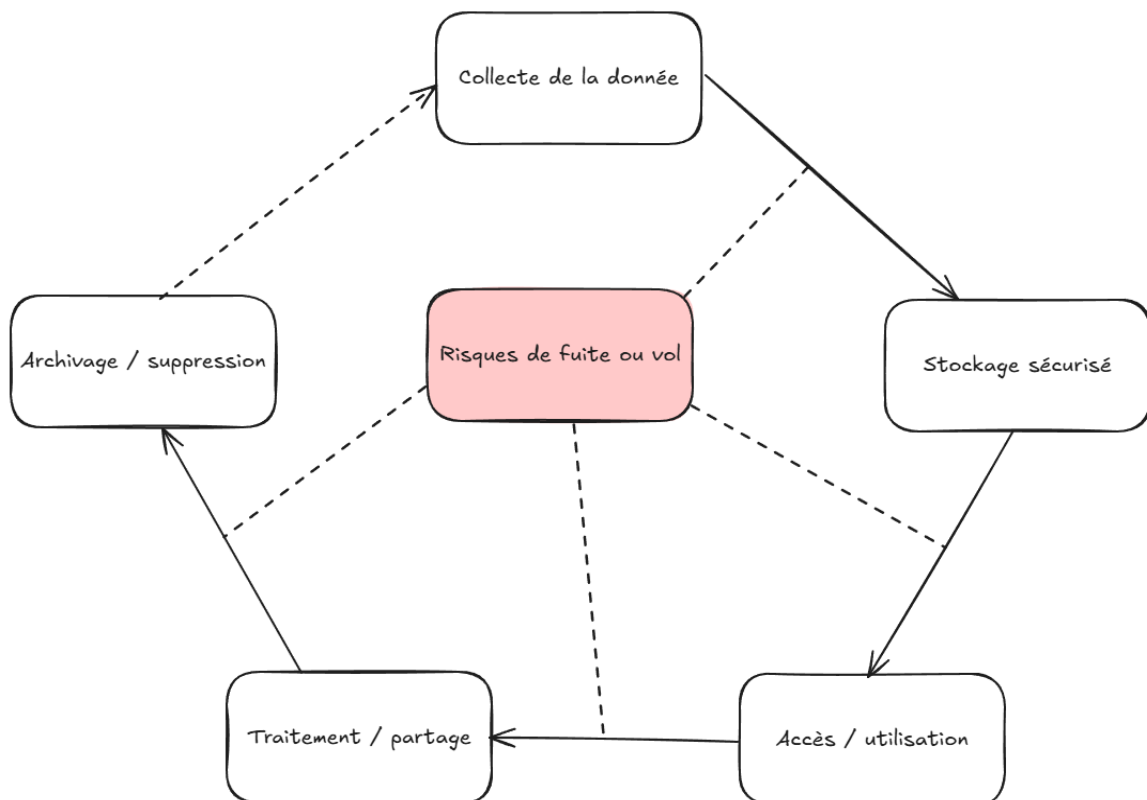
Les secrets industriels représenteraient jusqu'à 80 % de la valeur d'une entreprise, selon une étude citée dans l'article de Ettredge, Guo et Li (2018). Cela montre à quel point ces informations sont précieuses pour les entreprises modernes. Les structures les plus ciblées sont souvent les jeunes entreprises évoluant dans des secteurs très concurrentiels. En effet, disposant de moins de moyens pour protéger leurs systèmes informatiques, elles sont souvent plus vulnérables face aux cyberattaques. De plus, avec l'essor de l'AI, les secrets industriels ne sont plus uniquement en danger à cause des pirates informatiques, mais aussi en raison de cette nouvelle technologie elle-même, notamment lorsqu'elle est utilisée sans contrôle ni encadrement au sein de l'entreprise.

Les pirates informatiques utilisent notamment des techniques comme le phishing, qui consiste à récupérer les identifiants de connexion d'un employé en se faisant passer pour l'entreprise ou un service officiel. Ils peuvent aussi utiliser des malwares pour s'infiltrer dans le réseau informatique de l'entreprise et récupérer des prototypes ou des informations sensibles sur ses projets. Pour éviter cela, les entreprises doivent mettre en place des mesures techniques. Par exemple, l'utilisation d'un système de chiffrement permet de sécuriser les données échangées, même en cas d'interception. Elles doivent aussi instaurer des protocoles d'accès stricts, afin que seules les personnes autorisées puissent consulter ou manipuler des informations confidentielles (ROCHESTON [sans date]).

À cela s'ajoute désormais le risque lié au Shadow AI, lorsqu'un collaborateur utilise sans contrôle des outils d'AI externes, il peut, parfois sans s'en rendre compte, exposer des

secrets industriels ou des données confidentielles à des plateformes dont la sécurité et la confidentialité ne sont pas garanties par l'entreprise. Cette nouvelle menace vient compliquer encore davantage la gestion et la protection des informations stratégiques.

Figure 2 : Circulation et protection des données dans l'entreprise (cycle de vie d'une donnée en entreprise)



Ce schéma illustre le cycle de vie d'une donnée en entreprise, de sa collecte jusqu'à son archivage ou suppression. On constate que le risque de fuite ou de vol est présent à chaque étape du processus :

- Dès la collecte, une faille de sécurité ou une mauvaise manipulation peut exposer les informations sensibles.
- Lors du stockage, si les mesures techniques (chiffrement, accès restreints) ne sont pas conformes aux normes juridiques ou sectorielles, une fuite peut facilement survenir.
- Pendant l'accès ou l'utilisation, une attaque de type phishing ou une erreur humaine peut permettre à un pirate d'obtenir des identifiants et d'accéder aux données.
- Lors du traitement et du partage, le risque augmente si des données sont transférées à des prestataires externes ou via des outils non contrôlés (notamment en cas de Shadow AI).
- Enfin, même à l'archivage ou à la suppression, une récupération non autorisée reste possible, d'où l'importance de vraiment supprimer les données inutiles plutôt que de les conserver indéfiniment.

En résumé, la sécurité des données doit être pensée à chaque étape du parcours, et non uniquement à la fin. Ce schéma met en évidence la vigilance continue nécessaire pour limiter les risques de fuite et protéger le patrimoine informationnel de l'entreprise.

L'analyse des différents types de données sensibles montre à quel point chaque information, qu'elle soit d'ordre personnel ou stratégique, doit rester strictement protégée. Cela nécessite de mettre en œuvre des moyens techniques (chiffrement, contrôle des accès, anonymisation) et organisationnels (procédures internes, sensibilisation des employés) pour éviter toute fuite ou erreur de manipulation.

Les données détenues par les entreprises qu'il s'agisse de propriété technique, de code source ou de projets en développement représentent un enjeu clé dans la concurrence. Si l'intelligence artificielle offre aujourd'hui un soutien précieux pour automatiser certaines tâches ou accélérer la création de nouvelles fonctionnalités, il ne faut pas perdre de vue le principal danger, toute donnée transmise à une AI, en particulier à des outils externes et non contrôlés par l'entreprise, peut potentiellement être stockée, traitée ou même réutilisée à l'insu de l'organisation.

Dans ce contexte, il est indispensable d'adopter une approche responsable lors de l'utilisation de l'AI, chaque utilisateur doit se demander quelles données il intègre dans ses prompts et éviter absolument de transmettre des informations sensibles ou confidentielles. À défaut, les risques de fuite, de perte de propriété intellectuelle ou d'exploitation malveillante deviennent réels, dépassant largement le cadre technique pour engager la responsabilité globale de l'entreprise.

En résumé, la démocratisation du Shadow AI impose de repenser la gestion et la protection des données sensibles, en insistant à la fois sur la rigueur technique, la vigilance humaine et la sensibilisation de tous les collaborateurs.

2.2 L'impact de l'AI sur la gestion des données

Comme évoqué précédemment, la gestion des données en entreprise est primordiale, toute fuite susceptible de porter atteinte à des personnes doit être prise très au sérieux. Il est donc crucial de contrôler avec précision la manière dont les données sont stockées et l'endroit où elles le sont.

Sur le plan technique, ces dernières années, l'intelligence artificielle (AI) s'est largement démocratisée et son usage s'étend rapidement à de nombreux secteurs. Selon Oracle et SAP, l'AI désigne des systèmes capables d'exécuter des tâches qui, jusqu'ici, nécessitaient une intelligence humaine : analyse d'images, compréhension du langage, prise de décision, automatisation de processus, etc. L'objectif de ces technologies est de faciliter la vie, d'automatiser les tâches répétitives et d'optimiser la performance des entreprises. Malgré ses avancées spectaculaires, l'AI évolue encore à grande vitesse et

n'atteint pas toujours la perfection, mais elle ouvre la voie à de nombreuses innovations auparavant inaccessibles (ORACLE 2021; SAP [sans date]).

L'AI est principalement composée de machine learning et de deep learning. Ces deux méthodes de développement sont assez différentes. Le machine learning prend des données, les analyse, et apprend de celles-ci afin d'améliorer ses performances au fil du temps. Par exemple, un algorithme de tri de mails peut apprendre à mieux reconnaître les spams selon les habitudes de l'utilisateur. Le deep learning, lui, s'inspire du fonctionnement du cerveau humain. Il utilise des réseaux de neurones artificiels pour aller plus loin dans l'analyse, notamment en traitement d'images ou de langage. Par exemple, il peut être utilisé pour la reconnaissance faciale ou vocale, ou encore pour prédire des comportements à partir de grandes quantités de données (Robert 2020).

D'un point de vue technique, l'intelligence artificielle est donc un atout majeur pour la gestion des données. Elle permet, entre autres, de collecter, d'organiser et de stocker des données dans une entreprise, de manière plus rapide et plus efficace. Cela fait gagner un temps considérable, surtout lorsque les volumes de données sont très importants. Elle permet aussi de nettoyer les données, c'est-à-dire repérer les erreurs ou doublons et les corriger automatiquement. Par exemple, si une base de données client contient plusieurs fiches avec des informations similaires, l'AI peut les regrouper. Elle peut également traiter les données manquantes, par exemple en devinant une information absente à partir d'autres données similaires (HPE [sans date]).

Mais ce n'est pas tout. L'AI peut aussi repérer les tendances du marché ou du comportement des utilisateurs, ce qui permet aux entreprises d'anticiper les besoins de leurs clients et d'avoir une longueur d'avance sur la concurrence. Elle joue donc un rôle essentiel dans la prise de décision et le pilotage stratégique. Sur le plan organisationnel, l'AI apporte une automatisation encore plus poussée que celle d'un ERP classique, car elle est capable d'apprendre et de s'adapter aux changements. En plus d'optimiser l'utilisation des ressources, elle permet aussi une sécurité renforcée en identifiant les accès anormaux ou les tentatives de fuite de données, ce qui est crucial dans un contexte où les données sont devenues très sensibles (HPE [sans date] ; PWC [sans date]).

Un des grands enjeux liés à l'intelligence artificielle reste la sécurité des données, surtout dans les entreprises. Sur le plan technique et organisationnel l'AI peut rester vulnérable face aux cyberattaques, comme le piratage de bases de données ou l'injection de données falsifiées dans le système. À cela s'ajoute désormais le risque lié au "Shadow AI", c'est-à-dire l'utilisation d'outils d'AI externes par les employés sans contrôle officiel,

qui peut conduire à une exposition involontaire de données sensibles en dehors du cadre défini par l'entreprise. C'est pourquoi les responsables de l'AI et des données doivent sans cesse vérifier les failles possibles, appliquer des mises à jour régulières, et respecter des normes strictes de cybersécurité. Par exemple, une entreprise pourrait utiliser des systèmes de détection d'intrusion pour repérer des activités suspectes, ou encore mettre en place des sauvegardes automatiques pour restaurer rapidement les données en cas d'incident. D'autres aspects doivent également être pris en compte, notamment l'organisation interne, comme la gestion des accès (qui peut accéder à quoi), ou encore la traçabilité des actions, afin de savoir qui a modifié ou consulté des informations sensibles (PWC [sans date]).

Mais au-delà des risques techniques et organisationnels, l'aspect juridique est aussi essentiel, en particulier pour respecter les droits des personnes concernées par le traitement de données. Chaque individu a le droit de savoir comment ses données sont utilisées, et peut exiger des modifications, voire leur suppression. Par exemple, une personne peut demander à ne pas être profilée par une AI à des fins publicitaires, ou bien vouloir connaître les raisons d'un refus automatique de prêt bancaire. Les entreprises doivent donc s'assurer que leurs systèmes d'AI soient compréhensibles, transparents, et permettent aux utilisateurs de faire valoir leurs droits facilement. Cela passe par des explications claires, des interfaces simples pour envoyer une demande, et des procédures internes capables de traiter ces demandes rapidement et de manière équitable (PWC [sans date]).

L'intégration de l'intelligence artificielle dans la gestion des données constitue un atout indéniable pour les entreprises, en termes de gain de temps, d'efficacité et d'amélioration des processus organisationnels. Elle permet de répondre à de nombreux besoins comme par exemple optimisation de l'exploitation des données, aide à la prise de décision ou encore adaptation rapide aux évolutions du marché. Sur ce terrain, le machine learning apparaît comme la technologie la plus répandue, car il se nourrit en permanence de nouvelles données issues du quotidien de l'entreprise : informations clients, développements logiciels, code source, etc. Le deep learning, quant à lui, est surtout mobilisé pour des applications plus spécifiques, nécessitant le traitement de grandes quantités de données, par exemple en reconnaissance d'images ou d'objets.

Toutefois, cette capacité d'apprentissage continu du machine learning peut aussi devenir un vecteur de risque majeur, en intégrant toutes sortes de données, y compris celles fournies directement par les utilisateurs via des prompts ou des intégrations logicielles, il peut conduire, par inadvertance, à des fuites d'informations sensibles. Par

exemple, si du code confidentiel ou des données stratégiques sont injectés dans un outil d'AI externe, rien ne garantit qu'ils ne seront pas stockés, analysés, voire réutilisés sans le contrôle de l'entreprise.

Ainsi, si les bénéfices de l'AI sont incontestables, ils s'accompagnent d'un devoir de vigilance accru, il est essentiel de maîtriser l'alimentation des modèles et de définir précisément quelles données peuvent ou non être exploitées par l'intelligence artificielle. Sans cadre strict, le risque de compromission des données internes voire de perte d'avantage concurrentiel devient une réalité tangible, remettant en cause la sécurité globale de l'entreprise.

En conclusion, il ne suffit pas de se reposer sur les atouts de l'AI, l'enjeu principal est de garantir une gouvernance efficace des données, en alliant innovation, rigueur technique et respect des exigences juridiques.

2.3 Le shadow AI

Depuis fin 2022, les outils d'intelligence artificielle générative comme ChatGPT connaissent un essor fulgurant, aussi bien dans la vie personnelle que professionnelle. Cette technologie, capable de générer du texte, du code ou encore des images, est aujourd'hui utilisée dans de nombreuses entreprises pour automatiser certaines tâches et gagner du temps. Sur le plan organisationnel, un phénomène inquiétant a émergé, de plus en plus de salariés utilisent ces outils sans l'accord de leur service informatique. C'est ce qu'on appelle le Shadow AI, un usage non encadré de l'intelligence artificielle au sein des entreprises. Par exemple, un développeur peut soumettre du code confidentiel à un chatbot pour accélérer un projet, sans mesurer les risques liés à la confidentialité. Selon IBM, plus de la moitié des utilisateurs professionnels auraient déjà recours à ce type d'AI sans validation officielle, faisant du Shadow AI un sujet central dans le débat sur la sécurité numérique (Krantz, Jonker et McGrath 2024; LA REDACTION 2025).

Pour bien comprendre ce qu'est le Shadow AI, il est utile de le comparer à un phénomène plus ancien, le Shadow IT. Ce dernier désigne l'utilisation de logiciels ou services informatiques externes, comme Google Docs ou Dropbox, sans l'autorisation du service informatique. Par exemple, un employé peut y stocker des documents sensibles de l'entreprise, exposant ainsi ces données à des risques de fuite ou de piratage, car ces plateformes ne sont pas toujours contrôlées ni sécurisées par l'entreprise. Le Shadow AI fonctionne de manière similaire, mais avec des outils d'intelligence artificielle, d'un point de vue technique, les salariés y soumettent parfois

des données confidentielles sans validation, ce qui échappe totalement à la gouvernance interne de la cybersécurité (Krantz, Jonker et McGrath 2024).

Tableau 1 : Différences entre le Shadow IT et le Shadow AI

Critères	Shadow IT	Shadow AI
Définition	Usage de logiciels ou services numériques sans validation de la direction informatique (DSI).	Utilisation de modèles ou outils d'intelligence artificielle sans encadrement officiel ni supervision.
Exemples d'outils	Dropbox, Trello, Google Docs	ChatGPT, Notion AI, GitHub Copilot non autorisés, agents AI créés sur mesure.
Visible pour la sécurité informatique	Détection possible via les logs réseaux ou audits, mais parfois partielle.	Très faible, contenu et échanges souvent chiffrés ou transitant hors réseau interne, rendant la surveillance difficile.
Risques	Risques de fuites de données, absence de chiffrement, non-conformité au RGPD ou LPD.	Apprentissage incontrôlé (les données peuvent être réutilisées par l'AI externe), erreurs générées, perte de maîtrise sur les données sensibles.
Impact cybersécurité	Étendu mais relativement bien maîtrisé à ce jour.	Fort développement, encore peu contrôlé, avec un impact potentiel important sur la sécurité globale.

(Frenchweb 2025)

Le tableau ci-dessus, inspiré des éléments présentés par FrenchWeb, met en évidence plusieurs similitudes entre le Shadow IT et le Shadow AI, notamment l'absence de validation auprès de la direction informatique. Il souligne aussi des différences majeures,

comme la difficulté beaucoup plus grande à détecter l'usage du Shadow AI, ou encore la perte de maîtrise totale sur les données soumises à des outils externes.

Ce constat est particulièrement préoccupant pour les PME, qui ne disposent généralement pas de Direction des Systèmes d'Information (DSI) ni d'équipes dédiées à la sécurité. Elles sont donc souvent moins armées pour détecter ou encadrer l'usage de ces outils, ce qui augmente leur vulnérabilité face aux risques liés au Shadow AI. La formation des collaborateurs et la mise en place de règles claires deviennent alors des leviers essentiels pour limiter les incidents dans ce type de structure.

Le Shadow AI représente aujourd'hui une réelle préoccupation pour les entreprises, d'autant qu'il s'inscrit dans la continuité du Shadow IT, dont les risques étaient déjà bien connus. Mais avec l'explosion récente des AI génératives, comme ChatGPT ou Copilot, il devient de plus en plus difficile pour les organisations d'anticiper et de maîtriser tous les nouveaux risques associés. On l'a vu, en l'espace de deux ans, ces outils se sont démocratisés non seulement dans le monde professionnel, mais aussi dans les écoles et les universités. Désormais, beaucoup d'élèves s'appuient sur l'AI pour la rédaction de leurs travaux, la création d'applications, voire pendant certains examens lorsque cela est autorisé.

Cette évolution a un impact direct sur les entreprises, les jeunes diplômés ou nouveaux collaborateurs arrivent avec l'habitude d'utiliser l'AI au quotidien, parfois sans mesurer tous les risques liés à la confidentialité des données ou à la sécurité de l'information. C'est pourquoi l'aspect organisationnel devient crucial, il est indispensable d'informer et de former tous les nouveaux arrivants sur les bonnes pratiques, en précisant clairement quels outils d'AI sont autorisés, dans quels contextes et pour quels types de données.

D'autre part, il faut souligner que les capacités des AI évoluent très rapidement. Là où, il y a deux ans, le code généré par ChatGPT pouvait manquer de qualité, les outils actuels offrent des fonctionnalités bien plus avancées et précises, ce qui peut pousser à leur adoption massive, parfois sans recul critique.

En résumé, le phénomène du Shadow AI n'est pas prêt de disparaître, il s'impose progressivement dans toutes les sphères, de l'éducation à l'entreprise. Les organisations doivent donc impérativement renforcer leur gouvernance, anticiper les usages et sensibiliser en continu leurs équipes pour limiter les risques et préserver la sécurité de leurs données.

3. AI et protection des données

3.1 LPD, RGPD et autres réglementations

En Suisse, la protection des données personnelles s'appuie principalement sur la LPD (Loi fédérale sur la protection des données). Contrairement au RGPD européen, la LPD est pensée pour s'adapter aux réalités locales, mais l'objectif reste le même, protéger la vie privée des personnes et encadrer la façon dont leurs informations sont utilisées par les entreprises (FEDLEX 2023).

Ce qu'il faut bien comprendre, c'est que la LPD s'applique à toutes les entreprises qui traitent des données personnelles de personnes physiques, que ce traitement soit effectué en Suisse ou à l'étranger, dès lors qu'il produit des effets en Suisse. La surveillance de l'application de cette loi est assurée par le Préposé fédéral à la protection des données (PFPDT), qui joue un peu le rôle de "gendarme" pour les questions de vie privée (FEDLEX 2023).

Dans la pratique, la loi impose plusieurs grands principes, il faut toujours informer clairement les gens sur la façon dont leurs données sont utilisées, ne les collecter que pour des raisons valables, et surtout ne pas les conserver plus longtemps que nécessaire. Dès qu'une information n'a plus de raison d'être stockée, il faut la supprimer ou l'anonymiser. Si une personne demande à ce qu'on efface ses données, l'entreprise doit pouvoir répondre à cette demande, à moins qu'il existe une raison valable de les garder (FEDLEX 2023).

Un aspect important de la LPD, c'est le concept de "privacy by design" et "privacy by default". Cela signifie qu'il faut intégrer la sécurité des données dès la conception des systèmes, et prévoir par défaut des réglages qui protègent la vie privée, sans que l'utilisateur ait à tout configurer lui-même. Ce n'est pas seulement un détail technique, c'est une véritable obligation légale (FEDLEX 2023).

Un autre point à garder en tête, si une entreprise travaille avec des sous-traitants (par exemple un hébergeur ou un prestataire IT), elle doit s'assurer que ceux-ci respectent aussi la LPD. On ne peut pas simplement transférer la responsabilité, l'entreprise reste tenue de vérifier que ses partenaires appliquent les mêmes standards de sécurité (FEDLEX 2023).

Pour les entreprises étrangères qui traitent des données personnelles de personnes physiques en Suisse, et dont ce traitement a des effets sur le territoire suisse, la LPD prévoit aussi des obligations spécifiques, comme nommer un représentant local et tenir

un registre des traitements effectués. Ces exigences concernent surtout les grandes plateformes ou prestataires internationaux (FEDLEX 2023).

Côté transfert de données à l'étranger, la LPD impose que ça ne se fasse que vers des pays jugés "sûrs" par la Suisse, ou alors il faut mettre en place des garanties contractuelles supplémentaires (clauses, contrats spécifiques). Quelques exceptions existent (par exemple en cas de consentement explicite, nécessité contractuelle, ou si les données sont déjà publiques), mais dans la pratique, il est recommandé de vérifier systématiquement la conformité juridique avant tout transfert international de données (FEDLEX 2023).

Dans le cas d'une fuite de données importante, la LPD impose à l'entreprise de prévenir rapidement le PFPDT, en expliquant ce qui s'est passé, quelles données sont concernées, et ce qui est fait pour limiter les dégâts. Selon la gravité, il peut aussi être nécessaire de prévenir directement les personnes touchées (FEDLEX 2023).

Les lois peuvent être sectorielles, ce qui signifie que chaque secteur, que ce soit la santé, la finance ou encore les télécommunications est soumis à des règles spécifiques concernant les types de données à protéger et la manière de les sécuriser.

Dans le domaine de la santé, on retrouve par exemple la Loi fédérale sur le dossier électronique du patient (LDEP). Cette loi a pour objectif d'améliorer la qualité des soins médicaux en facilitant l'accès aux informations pertinentes d'un patient grâce à un dossier électronique. Pour créer ce dossier, le consentement écrit du patient est obligatoire (art. 3), et celui-ci doit être clairement informé de l'usage de ses données. Cette loi garantit également la confidentialité des données, en imposant des droits d'accès stricts aux professionnels de la santé. Par exemple, tout accès intentionnel non autorisé à un dossier peut être sanctionné d'une amende pouvant aller jusqu'à 100 000 francs (art. 24) (FEDLEX 2017).

Nous avons aussi la Loi sur le blanchiment d'argent et le financement du terrorisme (LBA). Cette loi encadre la gestion des données financières sensibles, comme les mouvements de fonds, les comptes clients ou les transactions suspectes. Elle impose aux établissements financiers, comme les banques ou les gestionnaires de fortune, des obligations strictes de surveillance et de traçabilité (FEDLEX 2021).

Par exemple, dans l'article 3, alinéa 5, les autorités comme la FINMA ou l'Administration fédérale des douanes ont la responsabilité de définir les seuils financiers jugés comme "importants" pour leur domaine d'activité. Cela veut dire qu'à partir de certains montants, les opérations doivent être signalées, vérifiées et documentées de façon plus rigoureuse.

L'objectif est d'empêcher que des fonds soient utilisés à des fins illégales, comme le financement du terrorisme ou des activités criminelles. Ces obligations impliquent donc une surveillance accrue des données financières et leur conservation pour permettre une traçabilité complète en cas de contrôle (FEDLEX 2021).

Tout comme les deux lois précédentes, on retrouve aussi la Loi sur les télécommunications (LTC). Cette loi vise à protéger les utilisateurs contre certaines pratiques abusives, comme la publicité de masse déloyale ou les services à valeur ajoutée trompeurs. Elle encadre principalement les entreprises de télécommunications pour garantir un traitement équitable des clients, par exemple en assurant un accès à haut débit pendant quatre ans ou en imposant une facturation non discriminatoire (art. 11, al. 1) (FEDLEX 2007).

Mais cette loi ne se limite pas à la relation commerciale. Elle touche aussi directement à la protection des données. Selon l'article 13a, l'Office fédéral de la communication (OFCOM) et la Commission fédérale peuvent avoir accès aux données personnelles dans le cadre de leurs missions, notamment en cas de sanctions ou de litiges. Ils peuvent traiter des données sensibles ainsi que des profils de la personnalité, dans le respect de règles strictes (FEDLEX 2007).

On constate à travers ce chapitre que les lois relatives à la protection des données sont indispensables pour préserver la vie privée et garantir la confiance dans les systèmes numériques. En Suisse, la LPD joue un rôle central, elle encadre précisément le traitement des données personnelles et assure une supervision efficace grâce au PFPDT. Ce cadre juridique permet d'éviter de nombreux abus, comme la revente de données à des sociétés pratiquant le marketing ciblé, ou la divulgation incontrôlée d'informations sensibles.

Pour les entreprises suisses, la LPD constitue un repère concret et pragmatique, parfaitement adapté à la réalité nationale. Il ne faut cependant pas négliger l'importance du RGPD, qui s'applique dès qu'une entreprise suisse traite ou transfère des données vers l'étranger, notamment dans l'Union européenne. Cela rend la compréhension et la conformité à ces deux textes essentiels, surtout dans un contexte économique globalisé où la circulation des données dépasse largement les frontières.

Il est également important de souligner le rôle des lois sectorielles, elles permettent de renforcer la protection dans des domaines particulièrement sensibles, comme la santé ou la finance, où les enjeux sont majeurs pour les individus concernés. Grâce à ce maillage réglementaire, le respect de la vie privée et la sécurité des données deviennent

de véritables priorités pour les entreprises. En définitive, sans ces protections juridiques strictes, nos données pourraient être utilisées ou détournées sans contrôle, ce qui mettrait en péril la confiance dans l'économie numérique.

Il est enfin important de relever qu'à l'heure actuelle, il n'existe aucune réglementation spécifique encadrant le Shadow AI, malgré son essor dans le monde professionnel. Cela ne signifie pas pour autant que ces pratiques échappent au droit, dès lors qu'une AI est utilisée en dehors du cadre officiel, tous les traitements de données restent soumis à la LPD et au RGPD, notamment en cas de transfert de données à l'étranger. Toute fuite ou usage inapproprié de données par l'intermédiaire d'un outil d'AI non validé expose donc l'entreprise aux mêmes risques juridiques que pour tout autre manquement en matière de protection des données. Ce flou réglementaire impose d'autant plus aux entreprises de renforcer leur vigilance et d'appliquer rigoureusement les principes de gouvernance déjà en place.

3.2 Responsabilités des entreprises face aux données personnelles

Après avoir présenté les réglementations comme la LPD et le RGPD, cette partie détaille les actions concrètes que les entreprises doivent mettre en œuvre pour protéger les données personnelles. Elle abordera les mesures internes telles que la sensibilisation du personnel (aspect organisationnel), l'importance d'un Data Protection Officer (aspect organisationnel et juridique), et les risques réels auxquels les entreprises s'exposent en cas de non-conformité (aspect juridique).

Les erreurs humaines sont souvent à l'origine des fuites de données. Parmi les méthodes les plus courantes utilisées par les pirates, on retrouve l'hameçonnage (phishing) ou ransomwares, qui visent à tromper la victime, par exemple avec un faux site demandant de se connecter pour voler ses identifiants. Le Shadow AI constitue également un risque supplémentaire lié à l'erreur humaine, lorsqu'un collaborateur utilise sans précaution un outil d'AI externe, il peut, parfois sans s'en rendre compte, exposer des informations sensibles en dehors du cadre sécurisé de l'entreprise. Pour limiter ces risques, il est essentiel de sensibiliser les employés. Cela passe par une gestion organisationnelle en faisant des campagnes régulières, adaptées à leur métier, pour leur apprendre à repérer les tentatives d'arnaque. Il est aussi utile d'organiser des tests pour vérifier qu'ils ont bien compris les réflexes à adopter (CNIL 2024).

D'après le site gouvernemental [Cybermalveillance.gouv.fr](https://cybermalveillance.gouv.fr), il est recommandé de mettre en place au moins une formation annuelle sur la sécurité des données. Ces formations doivent rappeler les risques encourus par l'entreprise et ses employés en cas de

négligence. Le site propose aussi un kit de sensibilisation complet, incluant des fiches pratiques, des conseils sur les mots de passe, des informations sur l'hameçonnage, ou encore des supports visuels comme des vidéos ou des bandes dessinées. Un autre point important est la séparation entre usages professionnels et personnels. Par exemple, stocker des documents de travail sur des comptes personnels peut représenter un risque et engager la responsabilité de l'employé en cas de fuite (CYBERMALVEILLANCE 2019).

Dans de nombreuses entreprises, les protections techniques sont bien mises en place, mais les mesures humaines appelées « mesures organisationnelles de cyberprotection » sont souvent oubliées. Ces mesures comprennent par exemple la formation des employés, des consignes claires sur l'usage des outils numériques ou encore la prévention des erreurs humaines. Pourtant, ces aspects sont essentiels, selon plusieurs études relayées par Nicole Bögli de La Mobilière, entre 70 % et 95 % des attaques informatiques commencent par une simple erreur humaine. Que ce soit à cause d'une mauvaise utilisation d'un outil d'AI, comme le Shadow AI, ou d'une attaque menée par un pirate, cela montre que les cybercriminels s'appuient surtout sur l'ingénierie sociale, c'est-à-dire sur des techniques de manipulation visant à piéger les employés. Par exemple, un faux message parlant de compte piraté ou de colis bloqué peut facilement pousser un salarié à cliquer sur un lien dangereux (Bögli [sans date]).

Pour limiter ces risques, la Mobilière conseille de désigner une personne de contact spécialisée en sécurité informatique, à qui les employés peuvent poser leurs questions ou signaler un doute. Il est aussi important d'organiser des formations régulières, car les techniques d'attaque évoluent constamment. Ces formations peuvent inclure des vidéos, des quiz ou même des simulations d'attaques pour tester les réactions du personnel. Enfin, il est recommandé de prévoir un budget spécifique pour ces actions, afin de ne pas les négliger. Développer une culture de la cybersécurité au sein de l'entreprise est aujourd'hui aussi crucial que de posséder de bons logiciels de protection (Bögli [sans date]).

Le Data Protection Officer, ou Conseiller à la protection des données en Suisse, est la personne chargée de surveiller si l'entreprise respecte bien les règles sur la protection des données. Contrairement au RGPD européen, la LPD suisse n'oblige pas les entreprises privées à désigner un conseiller. En revanche, c'est obligatoire pour les organes fédéraux, comme l'Office fédéral de la statistique (ADMIN [sans date]).

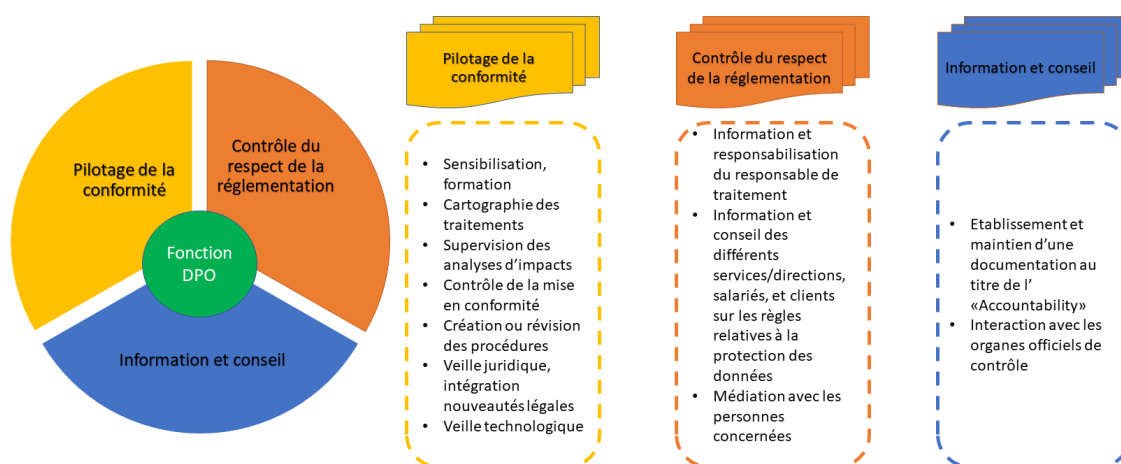
Ce conseiller joue un rôle important à l'intérieur de l'entreprise, mais aussi à l'extérieur, c'est lui qui est le principal contact du PFPDT (Préposé fédéral à la protection des

données et à la transparence). Il participe à la mise en œuvre des conditions d'utilisation et des règles de traitement des données, et il doit s'assurer que tout est conforme à la loi. Il ne consulte pas toutes les données, mais uniquement celles qui lui sont utiles pour remplir ses tâches (ADMIN [sans date]).

Le conseiller doit pouvoir informer directement la direction en cas de problème important. Il est aussi tenu d'agir de manière totalement indépendante, il ne peut pas faire partie de la direction de l'entreprise ni avoir d'autres fonctions qui pourraient créer un conflit d'intérêt. Par exemple, il ne peut pas travailler dans les ressources humaines ou dans l'équipe informatique en même temps (ADMIN [sans date]).

Il doit avoir de bonnes connaissances dans le domaine de la protection des données et de la sécurité informatique. Enfin, si une entreprise privée décide de nommer un conseiller, elle doit publier ses coordonnées et les transmettre via le portail officiel du PFPDT. Cela permet de bénéficier de certains avantages, notamment pour l'analyse d'impact sur la protection des données (ADMIN [sans date]).

Figure 3: Fonction du DPO



(ISEIG [sans date])

Le schéma ci-dessus, adapté des supports de formation de l'ISEIG, met en lumière les missions principales du DPO (Data Protection Officer). Comme on le voit, son rôle ne se limite pas à la conformité juridique, il agit comme un véritable chef d'orchestre, en sensibilisant le personnel, en surveillant la conformité, en prodiguant des conseils et en assurant la médiation en cas de problème. Cette vision transversale est essentielle pour garantir que toutes les étapes de la gestion des données, de la collecte à l'archivage, soient sécurisées et conformes à la législation en vigueur (ISEIG [sans date]).

À noter que dans de nombreuses PME, la fonction de DPO n'est pas toujours formellement présente ou attribuée à une personne dédiée. Cela peut rendre la gestion de la conformité et la sensibilisation plus complexes, d'où l'importance d'adapter les démarches à la taille et aux ressources de l'entreprise.

Lorsqu'une entreprise ne respecte pas la LPD, elle peut être exposée à des sanctions juridiques. D'après le Préposé fédéral à la protection des données et à la transparence (PFPDT), les articles 60 à 63 de la loi définissent quatre types d'infractions. Ces infractions ont trois points communs :

- Elles doivent être commises de manière intentionnelle (les erreurs par négligence ne sont pas sanctionnées pénalement).
- Elles visent d'abord des personnes physiques, c'est-à-dire les individus responsables, et non directement l'entreprise.
- La sanction maximale peut aller jusqu'à 250 000 francs suisses.

Dans certains cas, si l'enquête serait trop coûteuse pour identifier une personne précise, l'entreprise peut tout de même être condamnée à sa place, mais seulement si l'amende ne dépasse pas 50 000 francs (PFPDT [sans date]).

Au début du mois de juin 2025, une cyberattaque a touché le sous-traitant Chain IQ, basé à Zoug, qui travaillait pour plusieurs grandes entreprises suisses, dont UBS et Pictet. Cette attaque a permis aux pirates de voler un grand nombre de données sensibles, notamment celles de 137 000 employés d'UBS (noms, coordonnées personnelles, localisation...) et des fichiers de facturation détaillés pour la banque Pictet (SECURITY 2025).

Même si les banques ont assuré que les données clients n'étaient pas concernées, cette affaire reste grave. Elle soulève une question de confiance, comment garantir la sécurité des données quand un prestataire externe devient le maillon faible de la chaîne ? Cela peut nuire à l'image des entreprises concernées, mais aussi à celle du fournisseur lui-même (SECURITY 2025).

Steven Meyer, directeur de ZENDATA Cybersecurity, alerte sur ce type de situation :

« Trop souvent, une certification est perçue comme une preuve de sécurité. Mais elle ne protège pas face à une attaque réelle » (MEYER, 2025).

Cela montre que la conformité ne suffit pas. Il faut aller plus loin en testant régulièrement les systèmes des sous-traitants et en gardant une vue d'ensemble sur les risques (SECURITY 2025).

En Suisse, les PME sont particulièrement vulnérables face aux cyberattaques. Selon une étude relayée par GS-Global (Murith, 2024), 43 % des attaques ciblent directement ce type d'entreprise, avec un coût moyen estimé à 188'000 francs suisses par incident (Murith 2024).

Plus inquiétant encore, 60 % des PME victimes font faillite dans les six mois suivant l'attaque, ce qui montre bien l'ampleur des conséquences. À cela s'ajoute un temps moyen de détection de la fuite de 197 jours, ce qui laisse aux attaquants beaucoup de temps pour exploiter les données volées (Murith 2024).

Pourtant, la mise en place de mesures de sécurité de base (logiciels, maintenance, formations) représenterait moins de 1 % du chiffre d'affaires annuel d'une PME, ce qui constitue un coût dérisoire comparé aux pertes potentielles (Murith 2024).

En définitive, la responsabilité des entreprises en matière de protection des données personnelles est cruciale et doit être prise très au sérieux. Les conséquences d'une faille de sécurité ou d'un manquement à la réglementation peuvent être lourdes, amendes, poursuites, perte de confiance des clients, voire faillite dans les cas les plus graves. À l'ère de la numérisation et de la généralisation de l'AI, la moindre erreur ou négligence peut rapidement prendre une ampleur considérable.

Pour prévenir ces risques, il ne suffit pas de s'appuyer uniquement sur des dispositifs techniques, il est indispensable de mettre en place des mesures organisationnelles solides, telles que des formations régulières à la cybersécurité et à la gestion des données. Sensibiliser les employés, notamment sur les dangers liés à l'utilisation non maîtrisée des outils d'AI et sur la rédaction des prompts, devient une priorité. Cela permet de limiter les risques de fuite d'informations ou d'erreurs humaines, qui restent la principale cause des incidents de sécurité.

La nomination d'un Data Protection Officer (DPO) ou Conseiller à la protection des données est un levier important pour renforcer la gouvernance interne. Ce responsable veille au respect des réglementations, accompagne l'entreprise dans sa mise en conformité et sert de point de contact avec les autorités en cas d'incident. Il sera donc intéressant de s'interroger sur la présence et le rôle effectif du DPO dans les différentes entreprises, afin d'évaluer à quel point cette fonction est réellement intégrée dans les pratiques.

En résumé, seules les entreprises qui investissent sérieusement dans la prévention, la formation et la gouvernance de leurs données peuvent espérer préserver durablement

la sécurité des informations qui leur sont confiées, et maintenir la confiance de leurs clients, partenaires et collaborateurs.

4. Les risques du Shadow AI

4.1 Risques de fuite et exposition des données sensibles

L'utilisation non encadrée de l'intelligence artificielle représente aujourd'hui un risque important pour les entreprises, tant sur le plan technique (sécurité des données), organisationnel (absence de cadre interne) que juridique (non-conformité aux réglementations).

L'utilisation non encadrée de l'intelligence artificielle représente aujourd'hui un risque organisationnel important pour les entreprises. Environ un tiers des employés reconnaissent utiliser des outils d'AI sans l'accord de leur hiérarchie. Cette pratique, en pleine croissance, soulève plusieurs problèmes sérieux, comme la fuite d'informations sensibles, la non-conformité aux réglementations sur la protection des données, ou encore des atteintes à la réputation des entreprises (Krantz, Jonker et McGrath 2024).

Le risque de violation de données est l'un des plus préoccupants. Faute de contrôle ou de cadre clair, certains collaborateurs partagent involontairement des données confidentielles avec des plateformes d'AI générative. Une enquête menée auprès de responsables de la sécurité informatique (RSSI, pour Responsables de la sécurité des systèmes d'information) révèle qu'au Royaume-Uni, une entreprise sur cinq aurait déjà subi une fuite de données liée à l'utilisation non autorisée de ces outils techniques par ses employés (Krantz, Jonker et McGrath 2024).

En 2023, des employés de Samsung ont commis une erreur organisationnelle majeure en divulguant involontairement des informations confidentielles via ChatGPT. L'entreprise avait autorisé l'utilisation de l'AI pour assister les ingénieurs dans la résolution de problèmes techniques. Cependant, certains ont soumis à l'outil le code source d'un nouveau programme interne ainsi que des données sensibles liées au matériel (hardware). Le problème central est que ces informations sont désormais stockées sur les serveurs d'OpenAI, la société derrière ChatGPT, ce qui signifie que Samsung a perdu le contrôle exclusif sur ses propres secrets industriels. À la suite de cet incident, Samsung a décidé de restreindre fortement l'usage de l'AI en interne et a lancé le développement de sa propre solution d'intelligence artificielle, destinée exclusivement à un usage interne sécurisé (Maddison 2023).

La non-conformité aux réglementations est également un risque juridique majeur lié à l'usage du Shadow AI. En particulier, l'utilisation non autorisée d'outils d'AI peut entraîner des violations en matière de protection des données personnelles. Les conséquences sont principalement financières, avec des sanctions pouvant atteindre des montants très

importants. Par exemple, en cas de non-respect du Règlement général sur la protection des données (RGPD), une entreprise peut être condamnée à une amende allant jusqu'à 20 millions d'euros ou 4 % de son chiffre d'affaires mondial annuel, selon le montant le plus élevé (Krantz, Jonker et McGrath 2024). En Suisse, la Loi fédérale sur la protection des données (LPD) prévoit également des sanctions importantes, notamment des amendes pouvant atteindre 250'000 CHF pour les personnes physiques responsables d'une infraction (FEDLEX 2023). Malgré l'absence actuelle de cas connus spécifiquement liés au Shadow AI dans ce contexte, le risque demeure bien réel et devient de plus en plus probable à mesure que ces technologies se généralisent dans les entreprises.

L'atteinte à la réputation est un risque à ne pas sous-estimer. L'usage d'outils d'AI non encadrés peut produire des résultats qui ne correspondent pas aux attentes de l'entreprise ou à ses valeurs. Par exemple, des données incorrectes ou un modèle mal entraîné peuvent conduire à des décisions inadaptées, suscitant des doutes sur la fiabilité de l'organisation (Krantz, Jonker et McGrath 2024).

L'utilisation non encadrée de l'AI sur les plateformes de livraison peut avoir des effets négatifs sur l'image d'une entreprise. Par exemple, selon The Byte, le restaurant new-yorkais Roma Pizza s'est retrouvé avec des images de plats générées par intelligence artificielle sur son menu Uber Eats. Le problème a été repéré par un utilisateur du réseau X (anciennement Twitter), qui a remarqué des visuels étranges dont une « pizza » ressemblant plutôt à un dessert (Christian 2023).

Interrogé, le personnel du restaurant a affirmé ne pas avoir approuvé ces visuels, suggérant que quelqu'un d'autre les aurait ajoutés à leur place. Cela soulève la question : s'agit-il d'un malentendu ou d'un cas de Shadow AI, où l'outil a été utilisé sans validation de l'entreprise ? Quoi qu'il en soit, l'incident a suscité des moqueries en ligne et illustré à quel point une utilisation mal contrôlée de l'AI peut nuire à la réputation d'un commerce (Christian 2023).

Cet incident illustre parfaitement l'un des dangers du Shadow AI, l'utilisation d'outils d'intelligence artificielle sans validation ni contrôle interne peut entraîner la diffusion de contenus erronés ou inadaptés, impactant directement la réputation de l'entreprise. Lorsqu'aucune règle ou supervision n'encadre ces usages, le risque de perte de maîtrise sur l'image et la communication de la société devient bien réel, même pour des secteurs éloignés de la tech comme la restauration.

Cette section met en évidence les principaux risques liés à l'utilisation de l'intelligence artificielle, qu'ils soient d'ordre organisationnel, technique ou juridique.

Tableau 2 : Typologie des risques

Technique	Organisationnel	Juridique
Fuite de données	Atteinte à la réputation	Non-conformité LPD / RGPD
Mauvais paramétrage des AI externes	Utilisation non validée d'outils AI par les employés	Sanctions administratives ou pénales
Réponse AI biaisée	Diffusion de contenus sensibles	Impossibilité d'identifier la responsabilité

Ce tableau met en évidence la diversité des risques liés au Shadow AI qu'ils soient techniques, organisationnels ou juridiques, chacun nécessite une vigilance particulière pour garantir la sécurité et la conformité des pratiques en entreprise. Les conséquences du Shadow AI sont bien réelles et doivent être prises très au sérieux lors de l'intégration de ces outils dans une entreprise.

Pour Le cas de Samsung, il illustre parfaitement à quel point la méconnaissance ou la sous-estimation des risques peut mettre en péril même les plus grandes organisations. Lorsque l'AI était encore émergente, certains employés n'ont pas mesuré les conséquences potentielles du partage de leur code source avec un service externe, perte d'avantage concurrentiel, fuite d'informations stratégiques, voire un risque de faillite, même pour un groupe de cette envergure. Cela démontre qu'aucune entreprise n'est à l'abri des dérives liées à un usage non contrôlé de l'AI.

À l'heure actuelle, il n'existe pas encore de cas concrets de sanctions juridiques majeures en lien direct avec le Shadow AI, ce qui laisse penser que la prise de conscience est en cours mais que le pire peut être évité si les bonnes pratiques sont rapidement adoptées. Ce vide juridique apparent ne doit cependant pas rassurer, les risques sont bien présents et l'encadrement légal finira par se renforcer.

D'autres exemples, comme l'affaire des images générées par AI dans la restauration, illustrent des situations où il reste difficile de déterminer la responsabilité, s'agit-il d'une

erreur humaine ou d'une mauvaise gestion interne ? Cette incertitude ajoute une complexité supplémentaire à la gouvernance du numérique.

Enfin, l'émergence en 2025 de vidéos et d'images créées par l'AI soulève de nouveaux défis. Des contenus factices peuvent désormais nuire à la réputation d'une entreprise, en la faisant passer pour une organisation peu fiable ou en l'associant à des pratiques qu'elle ne cautionne pas. Ce phénomène est d'autant plus préoccupant pour les publics les plus vulnérables, comme les personnes âgées, qui risquent d'être trompées par des contenus paraissant authentiques.

Une dimension souvent sous-estimée est la multiplication rapide des plateformes d'AI générative. On ne parle plus seulement de ChatGPT, Copilot ou Gemini, en quelques clics, une multitude d'outils sont accessibles, chacun avec ses spécificités, son mode de fonctionnement et ses risques. Sur le site de la Grande École du Numérique, on recense déjà une dizaine de solutions populaires. Cette abondance complique le travail des responsables sécurité et DSI, même lorsqu'un outil est bloqué, un autre apparaît aussitôt ou parvient à contourner les restrictions techniques (GEDN 2024).

Pour les PME, la situation est encore plus délicate : chaque nouvel outil à surveiller demande du temps, des compétences et souvent un budget supplémentaire. Bloquer, tester ou auditer en continu finit par peser lourdement sur les ressources, sans garantir une protection totale.

À cela s'ajoute une tendance récente : le « Bring Your Own AI » (BYOAI), héritier du « Bring Your Own Device » (BYOD). Désormais, certains employés ne se contentent plus d'apporter leur matériel, ils introduisent aussi leurs propres outils d'AI dans leur activité quotidienne, parfois à l'insu de la direction. Selon une étude relayée par Microsoft, près de 78 % des salariés utiliseraient ainsi leurs solutions personnelles, souvent pour pallier les limites des outils internes (SUISSE.ING 2024).

Si cette pratique traduit un besoin réel d'efficacité et d'autonomie, elle fait peser de sérieux risques, fuite de données sensibles, non-conformité réglementaire, failles de sécurité supplémentaires, fragmentation des processus ou perte de maîtrise sur la circulation de l'information. Pour les PME, le BYOAI devient vite un casse-tête, il nécessite un suivi permanent, des formations ciblées et engendre des coûts cachés qui s'accumulent.

En résumé, la tendance du « chacun son AI », sans cadre ni sensibilisation, alourdit la charge organisationnelle et sécuritaire des entreprises, tout en complexifiant le rôle des responsables IT.

En somme, la tendance du « chacun son AI », sans cadre ni sensibilisation, alourdit la charge organisationnelle et sécuritaire des entreprises, tout en complexifiant le rôle des responsables IT.

En résumé, les risques du Shadow AI sont multiples et évoluent rapidement. Les entreprises doivent redoubler de vigilance, adapter en permanence leurs pratiques, et former leurs collaborateurs pour anticiper et limiter l'impact potentiel de ces dérives.

4.2 Conséquences légales et financières

L'usage non contrôlé de l'AI expose les entreprises à des sanctions légales et à des coûts financiers élevés. Que ce soit le juridique avec la LPD ou le RGPD, les amendes peuvent rapidement devenir critiques pour une organisation.

Concernant la LPD, les sanctions directes peuvent atteindre 250'000 CHF pour une personne physique, ou 50'000 CHF si l'auteur n'est pas identifié. Plusieurs types d'infractions sont explicitement mentionnés dans les articles 60 à 64. Il s'agit, par exemple, du refus de collaborer avec le PFPDT ou de la fourniture volontaire de renseignements inexacts lors d'une enquête, de la communication de données à l'étranger sans respecter les conditions légales, ou encore du transfert de données à un sous-traitant sans les garanties prévues par la loi. D'autres cas incluent le non-respect des exigences minimales techniques de sécurité, la révélation de données confidentielles dans le cadre professionnel, ou le non-respect d'une décision du Préposé fédéral à la protection des données et à la transparence. Dans toutes ces situations, l'entreprise ou la personne responsable s'expose à de lourdes amendes en cas de manquement intentionnel à la loi (FEDLEX 2023).

Dans le cas du Shadow AI, une utilisation non contrôlée d'outils d'intelligence artificielle par des collaborateurs peut donc conduire à ce type d'infraction, notamment en cas de fuite de données sensibles ou de transfert de données à l'étranger sans autorisation.

Pour ce qui concerne le RGPD, les sanctions peuvent être bien plus élevées qu'en Suisse. Par exemple, une entreprise peut recevoir une amende allant jusqu'à 10 millions d'euros ou 2 % de son chiffre d'affaires annuel mondial si elle ne respecte pas les principes techniques et organisationnels de privacy by design ou privacy by default. Ce montant peut doubler, c'est-à-dire atteindre 20 millions d'euros ou 4 % du chiffre d'affaires, en cas de violation des principes de traitement des données ou de non-respect des obligations fondamentales (Goldstein 2018).

Une fois un signalement effectué auprès de la CNIL (Commission Nationale de l'Informatique et des Libertés, l'autorité de contrôle en France) ou d'une autre autorité

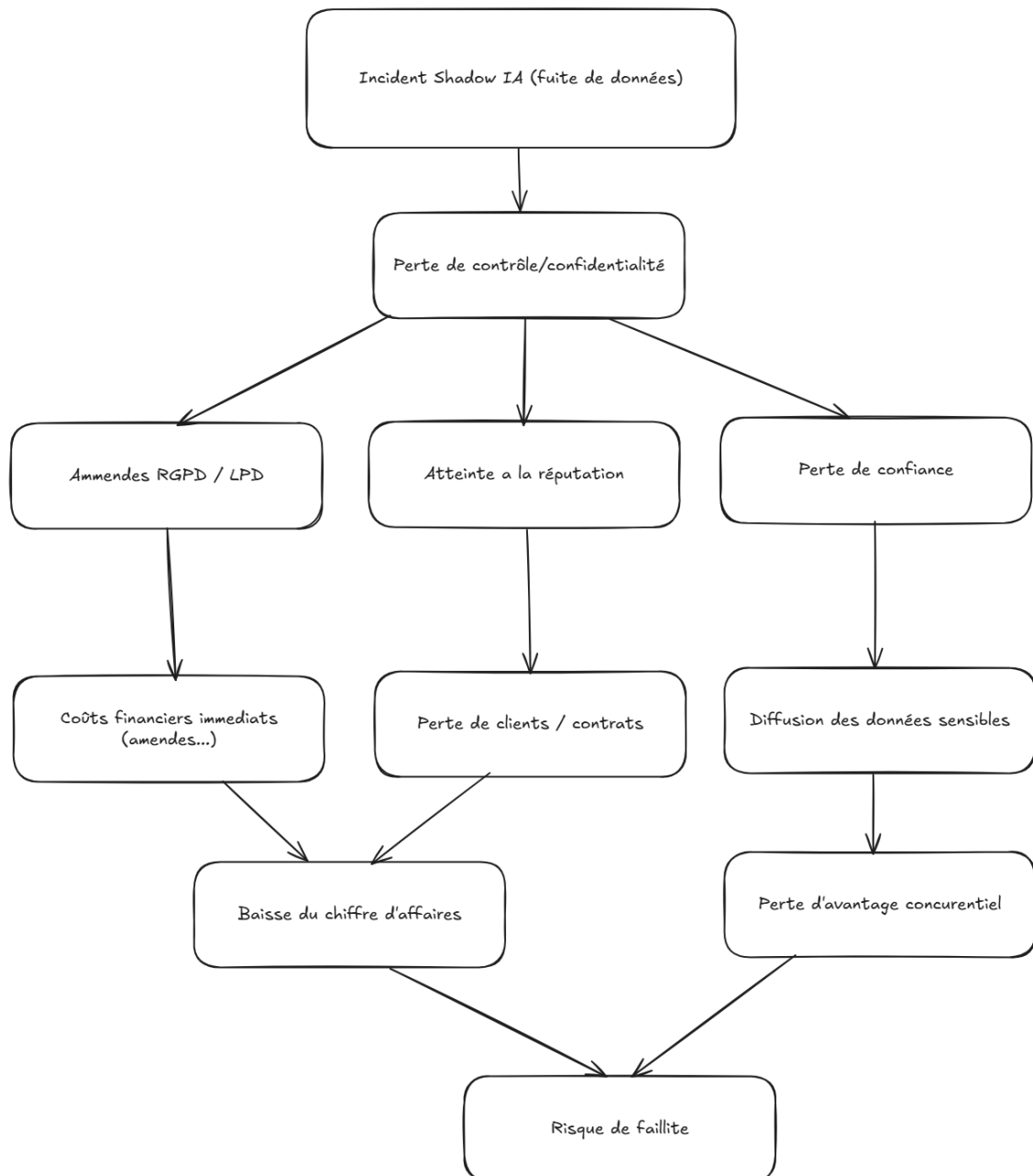
nationale compétente, un contrôle peut être mené et, en cas de non-conformité, différentes sanctions peuvent s'appliquer : sanctions administratives, pénales, et même un impact sur l'image de l'entreprise. Les États membres peuvent également prévoir des peines supplémentaires, pouvant aller jusqu'à 5 ans de prison et 300 000 euros d'amende en cas de détournement de finalité des données (Goldstein 2018).

Au-delà des sanctions légales prévues par la LPD ou le RGPD, une fuite de données liée au Shadow AI peut entraîner des conséquences financières très lourdes pour une entreprise. Selon un rapport d'IBM, le coût moyen d'une fuite de données a atteint 4,88 millions de dollars en 2024. Pour la Suisse, le coût moyen était de 3,7 millions de francs suisses selon une autre étude IBM en 2023. Ces montants incluent non seulement les amendes ou frais judiciaires, mais aussi de nombreux coûts indirects, gestion de crise, indemnisation des victimes, perte de clients et impact sur la réputation. Dans certains cas, il faut également ajouter les frais liés à la gestion des rançongiciels ou à la perte de secrets techniques industriels (Field Effect 2024; SUPPLIERSHIELD [sans date]).

Un autre facteur majeur du coût est lié aux activités post-incident, qui ont fortement augmenté ces dernières années. D'après IBM, ces activités peuvent représenter jusqu'à 1,35 million de dollars du coût total d'une fuite, car elles comprennent le fait d'offrir un service de surveillance aux victimes pour s'assurer qu'aucune fraude n'est commise avec leurs données, ou encore le remboursement de certains frais en cas de problème. À long terme, la perte de confiance des clients et partenaires peut entraîner une baisse durable du chiffre d'affaires et nuire à la valeur de l'entreprise. Ainsi, les conséquences financières d'un incident de Shadow AI dépassent largement le simple montant de l'amende, et peuvent mettre en péril la pérennité même de l'organisation (Field Effect 2024).

Afin de mieux visualiser la cascade de conséquences qu'un incident de Shadow AI peut entraîner, la figure ci-dessous résume la chaîne des impacts possibles, depuis la perte de contrôle des données jusqu'au risque ultime de faillite.

Figure 4: Chaîne des conséquences d'un incident Shadow AI



On constate à travers l'analyse des conséquences légales et financières que les impacts d'un incident lié au Shadow AI peuvent être bien plus lourds qu'il n'y paraît de prime abord. Comme l'illustrent les cas concrets évoqués précédemment, les sanctions juridiques ne sont que la partie émergée de l'iceberg, les entreprises doivent aussi faire

face à des coûts financiers indirects parfois colossaux, qu'il s'agisse de la gestion de crise, de la perte de clients, ou de l'atteinte durable à leur réputation.

Ce qui ressort, c'est que les conséquences d'une fuite de données ou d'un manquement aux réglementations ne se limitent pas à une simple amende, il s'agit bien souvent d'une remise en question profonde de la confiance des clients et partenaires, qui peut mettre en péril la survie de l'entreprise elle-même. Un incident majeur lié au Shadow AI peut entraîner une perte de chiffre d'affaires significative, voire la faillite dans les cas les plus graves, notamment pour les PME qui n'ont pas la capacité d'absorber de telles pertes.

Cela montre l'importance d'une approche globale et préventive, investir dans la sécurité, la formation, et la gouvernance des données n'est plus une option, mais une nécessité vitale pour toutes les organisations, quelle que soit leur taille. Plus que jamais, l'anticipation et la préparation sont les clés pour limiter l'impact de ces risques, garantir la conformité et préserver la pérennité de l'entreprise à long terme.

5. Sécurisation des données en entreprise

5.1 Stratégies de prévention et bonnes pratiques

Avec l'émergence du Shadow AI, les stratégies de prévention et les bonnes pratiques en cybersécurité sont devenues plus importantes que jamais. Ce phénomène, qui consiste à utiliser des outils d'intelligence artificielle sans validation ni contrôle de la part de l'entreprise, fait apparaître de nouveaux risques et oblige les organisations à revoir et renforcer leurs méthodes de protection.

GS-Global, spécialiste en assurance et gestion de patrimoine, recommande aux entreprises de réaliser régulièrement des auto-diagnostic et des audits de cybersécurité afin d'identifier rapidement d'éventuelles failles. Ces démarches relèvent à la fois de l'aspect organisationnel (mise en place de procédures internes, audit des process) et de l'aspect technique (tests d'intrusion, contrôle des mots de passe, mises à jour logicielles) (Murith 2024).

Cependant, évaluer sa vulnérabilité ne suffit pas, il est tout aussi essentiel de mettre en place une véritable stratégie de cybersécurité. Cela passe par des mesures techniques telles que pare-feux, chiffrement des données, sauvegardes régulières et contrôles d'accès stricts, mais aussi des organisationnelles en faisant la formation des employés et l'élaboration de politiques internes claires (Murith 2024).

Enfin, GS-Global insiste sur l'importance de l'assurance cyber, ce type de contrat permet non seulement de couvrir les pertes financières après une attaque (rançons, pertes d'exploitation, restauration de données), mais aussi de gérer la crise (notification des clients, communication, etc.). Bien choisir sa police et réévaluer régulièrement ses besoins est aujourd'hui indispensable pour rester protégé face à l'évolution des menaces (Murith 2024).

Ce point concerne d'autant plus les PME, qui n'ont pas toujours les moyens de supporter seules les conséquences financières d'un incident majeur. L'assurance cyber peut alors représenter une véritable bouée de sauvetage pour la survie de l'entreprise en cas de crise.

Selon Bizz and Buzz, plateforme française dédiée à l'innovation numérique, cinq bonnes pratiques sont recommandées pour encadrer le Shadow AI en entreprise. Parmi celles-ci, deux points méritent une attention particulière ici, la mise en place d'une charte d'usage de l'AI, qui précise quels types de données ne doivent jamais être confiés à des outils d'AI et encadre les usages autorisés, ainsi que l'implication transversale de la DSI,

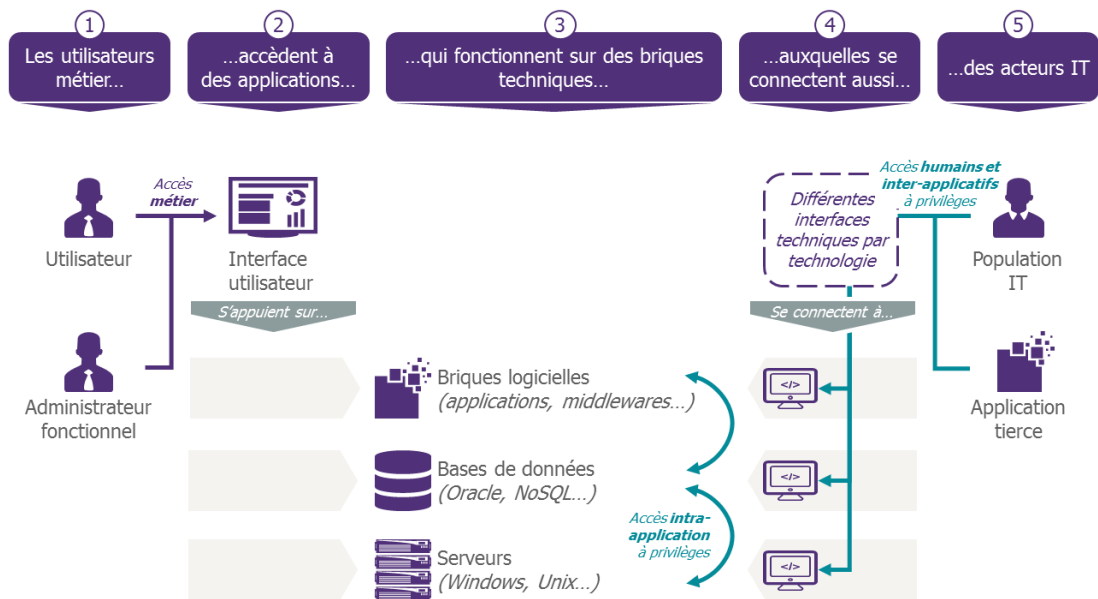
des métiers, du juridique et de la sécurité dans les projets AI. Cette gouvernance partagée garantit que les solutions adoptées répondent à la fois aux impératifs de sécurité et aux besoins opérationnels. L'objectif n'est pas de freiner l'innovation, mais de la canaliser, en sensibilisant et accompagnant les collaborateurs vers un usage responsable et conforme de l'intelligence artificielle (FOEHRENBACHER 2025).

Cette question se pose tout particulièrement pour les PME, qui n'ont pas toujours accès à des experts dédiés à la sécurité ou au juridique. Le fait d'adopter une charte claire et d'associer différents profils internes à la réflexion permet alors de poser un cadre efficace autour de l'AI, même en l'absence de moyens importants.

L'Agence nationale de la sécurité des systèmes d'information (ANSSI) en France propose dix règles de base pour renforcer la cybersécurité au quotidien, aussi bien dans le cadre professionnel que personnel. Parmi ces recommandations, la séparation stricte des usages personnels et professionnels est essentielle, il ne faut jamais utiliser son adresse e-mail d'entreprise pour des inscriptions à titre privé, ni connecter des appareils personnels au réseau de l'entreprise. Ce principe vaut aussi pour l'utilisation d'outils d'AI non autorisés, utiliser une intelligence artificielle externe avec des données professionnelles, même par facilité, expose l'entreprise à des risques de fuites ou d'exploitation des informations sensibles. Cela limite les risques en cas de fuite ou de piratage, car une compromission d'un site personnel pourrait alors permettre à un attaquant de cibler l'organisation par la suite, notamment via des attaques de phishing ou de ransomware (ANSSI 2025).

L'ANSSI insiste également sur l'importance de l'authentification à double facteur, qui rend l'accès aux comptes beaucoup plus difficile pour les pirates, même si un mot de passe est compromis. L'utilisation de réseaux Wi-Fi publics doit être évitée autant que possible, car ces réseaux facilitent l'interception de données et l'installation de logiciels malveillants. Enfin, il est primordial d'appliquer la règle du « moindre privilège », chaque collaborateur ne doit disposer que des accès nécessaires à ses fonctions, afin de réduire le risque d'accès non autorisé à des informations sensibles. L'ensemble de ces bonnes pratiques, simples à appliquer, permettent déjà de se protéger contre la majorité des attaques numériques actuelles (ANSSI 2025).

Figure 5: Le moindre privilège



(Karmann 2017)

L'image extraite du site RiskInsight illustre parfaitement le principe du « moindre privilège » dans la gestion des accès aux systèmes d'information. On y distingue deux grandes catégories d'utilisateurs, d'un côté, les utilisateurs métiers (ressources humaines, comptables, etc.), de l'autre, les acteurs IT. Les utilisateurs non-IT n'ont accès qu'à la partie visible de l'application, c'est-à-dire l'interface utilisateur, ils peuvent utiliser les fonctionnalités prévues sans jamais accéder directement au code, aux bases de données, ni aux serveurs qui font fonctionner l'application. Ce cloisonnement est fondamental, car ces utilisateurs n'ont ni le besoin ni les compétences pour intervenir sur les couches techniques sous-jacentes (Karmann 2017).

Cependant, le Shadow AI introduit un nouveau type de risque, même sans accès technique direct, un utilisateur métier peut, par exemple, copier des données sensibles depuis une application métier et les transmettre involontairement à un outil d'intelligence artificielle externe. Ce simple geste de copier-coller peut suffire à contourner les protections internes et exposer des informations confidentielles à l'extérieur de l'entreprise.

À l'inverse, la population IT bénéficie d'accès spécifiques et étendus, ils interviennent sur les briques logicielles, les bases de données ou les serveurs, et disposent souvent de droits élevés afin de maintenir, mettre à jour, ou sécuriser les systèmes. C'est aussi

eux qui corrigent les éventuelles failles et supervisent les connexions privilégiées, qu'elles soient humaines ou inter-applicatives (Karmann 2017).

Cette séparation stricte limite le risque de voir des utilisateurs non autorisés accéder à des ressources sensibles, conformément à la règle du moindre privilège, chaque personne ne doit disposer que des droits strictement nécessaires à sa fonction. Ce modèle est d'autant plus important que les accès à privilèges sont particulièrement ciblés par les cyberattaques, et peuvent, en cas de mauvaise gestion, ouvrir la porte à des compromissions majeures de tout le système d'information (Karmann 2017).

5.1.1 Solutions techniques avancées pour la gestion du Shadow AI

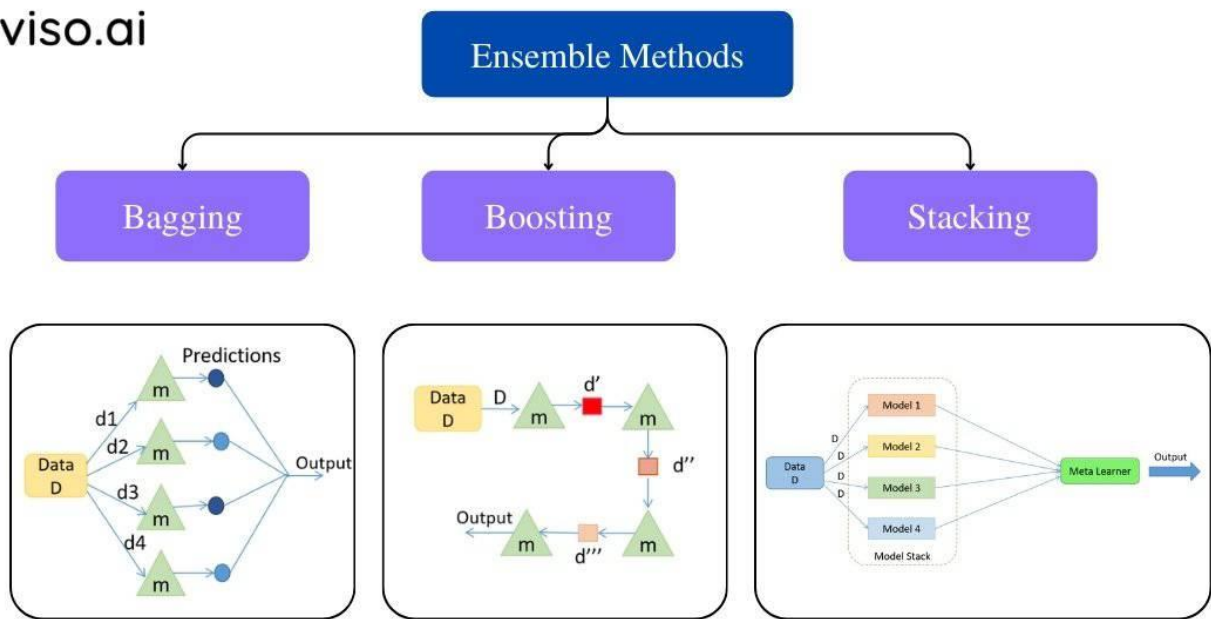
Outre les mesures organisationnelles et les bonnes pratiques classiques, la sécurisation des données face au Shadow AI nécessite également l'adoption de solutions techniques avancées. Parmi celles-ci, les approches hybrides (ou ensemble de modèles), l'Explainable AI (XAI) et l'utilisation de sandbox pour garantir la fiabilité des AI génératives jouent un rôle clé. Ces méthodes permettent non seulement d'augmenter la robustesse et la transparence des systèmes, mais aussi de limiter les risques liés à l'autonomie croissante des intelligences artificielles. Leur intégration progressive dans les entreprises devient aujourd'hui incontournable pour renforcer la sécurité globale et mieux maîtriser les enjeux spécifiques liés à l'AI.

Avant d'aborder en détail ces différents outils, il est essentiel de rappeler ce qu'est un algorithme. Un algorithme correspond à une suite d'étapes permettant de résoudre un problème ou d'obtenir un résultat. Cette notion n'est pas nouvelle, on la retrouve par exemple dans les recettes de cuisine ou les calculs mathématiques. Aujourd'hui, les algorithmes jouent un rôle central dans les AI génératives, qui doivent analyser la demande de l'utilisateur, rechercher des informations pertinentes, puis formuler une réponse adaptée (DATASULTING 2023).

Concernant les approches hybrides, il s'agit d'utiliser non pas un seul modèle d'intelligence artificielle, mais plusieurs en parallèle ou en séquence, afin de croiser et renforcer leurs prédictions. Cette méthode, appelée ensemble learning, fonctionne un peu comme un comité d'experts, chaque AI apporte son point de vue sur la question, puis leurs réponses sont combinées pour obtenir un résultat final plus faible et moins sujet à l'erreur ou au biais individuel (Boesch 2024).

Figure 6: Ensembles des méthodes hybrides

viso.ai



(Boesch 2024)

Trois techniques principales existent dans ce domaine : le bagging, le boosting et le stacking.

Le bagging (Bootstrap Aggregating) consiste à entraîner plusieurs modèles sur des sous-échantillons différents des mêmes données, puis à combiner leurs résultats, ce qui réduit la variance et améliore la stabilité des prédictions.

Le boosting fonctionne de manière séquentielle, chaque nouveau modèle se concentre sur les erreurs commises par les précédents, ce qui permet de corriger progressivement les biais et d'augmenter la précision globale.

Le stacking va encore plus loin, en combinant les prédictions de plusieurs modèles de base pour entraîner un modèle final (appelé méta-modèle) qui apprend comment pondérer et assembler les différentes réponses obtenues (Boesch 2024).

Dans le contexte de l'entreprise, ces approches permettent de limiter le risque d'erreurs ou de résultats biaisés, ce qui est particulièrement important lorsque des outils d'AI y compris des Shadow AI utilisés en dehors du contrôle officiel interviennent dans des décisions sensibles. Même si les méthodes d'ensemble ne remplacent pas une politique de gouvernance et de validation stricte, elles apportent une couche supplémentaire de fiabilité technique et limitent l'impact d'un modèle unique potentiellement vulnérable.

Dans la pratique, de nombreux outils et frameworks d'AI modernes intègrent déjà ces techniques d'ensemble. Par exemple, Random Forest (pour le bagging), XGBoost ou GBM (pour le boosting), ou encore les classes `StackingClassifier` de Scikit-learn (pour le stacking) sont fréquemment utilisés dans des domaines comme la finance, la cybersécurité ou la santé, afin d'améliorer la précision des prédictions et de limiter les risques d'erreur (Boesch 2024; Grootendorst [sans date]).

Cela dit, il revient souvent à l'entreprise de mettre en place ces stratégies de manière adaptée à ses propres besoins, soit en configurant les outils standards, soit en développant des modèles sur mesure pour des problématiques très spécifiques. L'essentiel est que ces techniques sont aujourd'hui accessibles, mais leur efficacité dépendra toujours de leur bonne intégration dans la gouvernance globale des données et de l'AI de l'organisation.

L'intelligence artificielle explicable, ou XAI (Explainable AI), regroupe l'ensemble des techniques qui visent à rendre les décisions d'une AI compréhensibles et transparentes pour l'humain. Au lieu d'accepter les résultats d'un modèle comme une "boîte noire", la XAI permet d'analyser comment et pourquoi l'algorithme a abouti à telle prédiction, de vérifier l'absence de biais et de s'assurer que les décisions restent justifiables. Cette démarche est devenue incontournable pour répondre aux enjeux de conformité, d'audibilité et d'éthique, mais aussi pour renforcer la confiance des utilisateurs et des parties prenantes dans les systèmes automatisés (IBM 2023).

L'AI explicable ne se limite pas à détecter les biais, elle aide aussi à comprendre la logique des modèles, à contrôler l'équité des décisions, à surveiller l'évolution des performances ou encore à gérer les risques réglementaires ou réputationnels. Des techniques comme LIME ou DeepLIFT sont aujourd'hui utilisées pour expliquer localement chaque décision d'un modèle complexe, et des plateformes entières sont conçues pour assurer ce suivi en continu, faciliter l'audit et améliorer l'expérience utilisateur (IBM 2023).

Concrètement, l'AI explicable s'illustre dans des domaines comme la santé (analyse d'images médicales, diagnostic), la finance (approbation de crédit, détection de fraudes) ou la justice pénale (analyse ADN, évaluation des risques), où elle favorise la transparence, la confiance et la conformité réglementaire (IBM 2023).

Dans le contexte du Shadow AI, la question de l'explicabilité prend encore plus d'importance. En effet, un des grands dangers du Shadow AI, c'est l'usage d'outils d'AI externes, non validés par l'entreprise, dont le fonctionnement reste souvent opaque.

Cela peut entraîner des décisions incompréhensibles, des biais non détectés, et mettre en péril la conformité de l'organisation, notamment au regard du RGPD ou de la LPD.

Adopter une politique de prévention basée sur l'AI explicable, c'est imposer ou recommander des outils dont les décisions sont transparentes et vérifiables, et non pas des modèles obscurs impossibles à contrôler. C'est aussi favoriser le développement d'AI internes dont on maîtrise le fonctionnement, ou bien auditer systématiquement tout nouvel outil utilisé par les collaborateurs. En bref, la XAI devient une pièce centrale d'une gouvernance efficace, elle permet à l'entreprise de garder la main sur ses données, de prouver la légitimité de ses décisions automatisées, et de limiter les risques de dérives inhérentes au Shadow AI.

Le concept de sandbox est largement utilisé dans l'innovation technologique pour permettre aux entreprises de tester de nouveaux outils ou méthodes dans un environnement sécurisé, sans exposer leurs systèmes critiques ni leurs données sensibles. En intelligence artificielle, le sandbox fonctionne sur le même principe, il s'agit d'une "zone de test" où un modèle ou une application AI peut-être expérimenté, audité et affiné avant d'être déployé à grande échelle. Cette démarche facilite la collaboration entre développeurs, juristes et métiers, tout en offrant un cadre clair pour respecter la réglementation et l'éthique. Par exemple, l'initiative « Innovation Sandbox for Artificial Intelligence » menée par le canton de Zurich met à disposition un environnement où les entreprises et administrations peuvent tester en toute sécurité leurs projets AI, accéder à de nouvelles sources de données et bénéficier d'un accompagnement sur les aspects réglementaires. Plusieurs projets réalisés dans ce cadre ont permis, par exemple, de simuler la détection de fraudes à l'identité utilisant des deepfakes, ou d'expérimenter l'analyse automatique de dossiers administratifs par l'AI, tout cela sans jamais exposer de données réelles. Cette approche s'avère particulièrement précieuse à l'heure du Shadow AI, où il devient essentiel de contrôler et d'encadrer les expérimentations pour éviter tout risque de fuite ou de non-conformité (Zürich [sans date]).

Dans le contexte du Shadow AI, l'utilisation d'un sandbox prend tout son sens. Quand les collaborateurs se tournent vers des outils d'AI non validés ou non maîtrisés, le risque principal est d'introduire des failles de sécurité, de perdre le contrôle sur les données ou de violer des règles internes et externes sans s'en rendre compte. Mettre en place un sandbox permet justement à l'entreprise d'encadrer et de canaliser l'expérimentation, au lieu de laisser chacun tester des AI de façon non contrôlée (Shadow AI), on centralise les tests dans un espace sécurisé où chaque projet peut être évalué, audité et ajusté avant d'être adopté. Cela limite fortement le risque de fuite de données ou de mauvaise

utilisation de l'AI, tout en gardant une dynamique d'innovation. Le sandbox devient ainsi un outil clé de prévention, il encourage la curiosité et l'initiative, mais dans un cadre maîtrisé qui protège l'organisation et ses actifs.

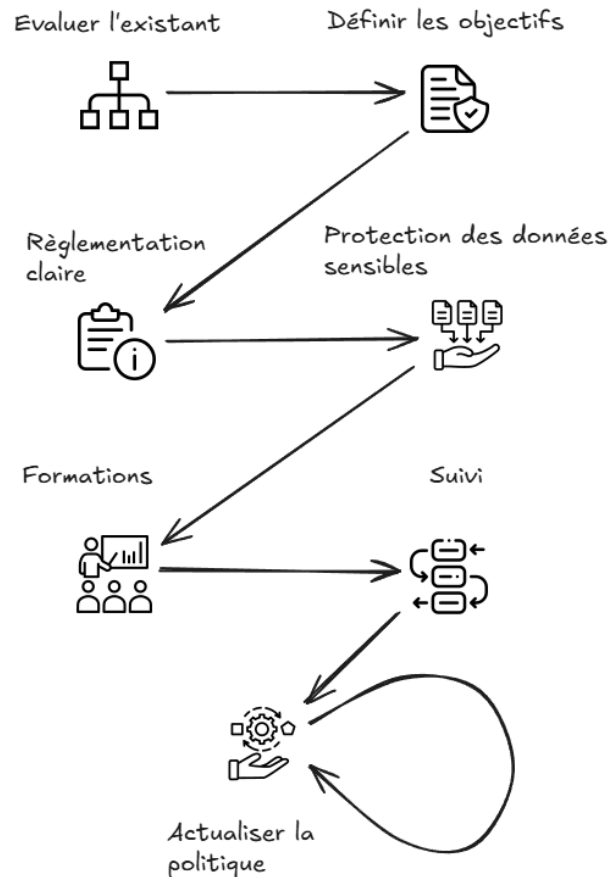
5.2 Réglementation interne et outils de contrôle

Pour limiter les risques liés au Shadow AI, il est indispensable qu'une réglementation interne soit mise en place, cela passe par la mise en place d'une politique claire et de dispositifs adaptés de contrôle et de surveillance. L'usage de l'AI doit ainsi être structuré autour de chartes, procédures et d'outils techniques permettant de détecter et de limiter les usages non autorisés.

En pratique, cela passe par l'élaboration de politiques d'utilisation des AI, la sensibilisation des collaborateurs, mais aussi par le recours à des dispositifs techniques de filtrage, de journalisation et d'audit régulier. Les équipes informatiques et juridiques doivent travailler de concert pour assurer un équilibre entre innovation et sécurité, en veillant à ce que chaque nouvel outil d'AI soit validé, testé en environnement sécurisé, et suivi dans le temps.

Pour encadrer efficacement l'utilisation de l'AI en entreprise, il est essentiel d'élaborer une politique structurée et adaptée. Le site Openmind Tech propose un guide en sept étapes qui permet de poser rapidement des bases solides :

Figure 7 : Elaboration d'une politique d'utilisation structurée



(Openmind 2025)

- **Évaluer l'existant** : Commencez par recenser toutes les utilisations de l'AI, qu'elles soient officielles ou issues d'initiatives personnelles des collaborateurs. Cette première étape donne une vision claire des pratiques et des outils déjà présents dans l'organisation.
- **Définir les objectifs** : Reliez votre politique d'utilisation de l'AI aux objectifs stratégiques de l'entreprise. Il s'agit d'identifier où l'AI peut réellement apporter de la valeur et dans quels domaines son usage doit être limité.
- **Établir des règles claires** : Précisez dans quels cas et comment l'AI peut-être utilisée. Distinguez les tâches autorisées (automatisation de tâches répétitives, analyse de données) et celles qui doivent être évitées, surtout lorsqu'elles touchent à des données sensibles ou confidentielles.
- **Protéger les données sensibles** : Évitez d'utiliser des outils d'AI publics pour traiter des informations critiques. Mettez en place, d'un côté, des protocoles internes clairs qui définissent précisément quels types de données peuvent ou

non être partagés, et dans quelles conditions. De l'autre, renforcez la sécurité technique en appliquant des mesures comme le chiffrement, les accès restreints et la surveillance des transferts.

- Former les équipes : Prévoyez des sessions de formation sur les usages approuvés de l'AI, les risques (biais, erreurs, sécurité) et les meilleures pratiques pour tous les collaborateurs.
- Mettre en place un suivi : Organisez un suivi régulier (audits, contrôles, KPIs) pour vérifier l'application de la politique et ajuster rapidement en cas de dérive ou de nouveauté technique.
- Actualiser la politique : Enfin, adaptez la politique au fil du temps. L'AI évolue très vite : il est donc primordial de réviser, au moins une fois par an, vos règles internes pour rester à jour sur les nouvelles pratiques et technologies.

Cette approche structurée permet non seulement de canaliser l'innovation, mais aussi de garantir la sécurité et la conformité des usages de l'AI en entreprise, tout en responsabilisant chaque collaborateur (Openmind 2025).

Dans la réalité des PME, la mise en place de telles chartes et politiques d'utilisation s'avère particulièrement pertinente. Elles permettent d'encadrer rapidement l'usage de l'AI sans devoir investir massivement dans des outils techniques complexes ou des équipes spécialisées. En optant pour des règles internes claires et une sensibilisation adaptée, les PME peuvent limiter efficacement les risques liés au Shadow AI, tout en maîtrisant leurs coûts. Cette approche pragmatique représente souvent le meilleur compromis, elle structure les usages, responsabilise les employés et évite de devoir mobiliser des ressources importantes pour des solutions de surveillance sophistiquées, qui restent parfois hors de portée des petites structures. En résumé, la charte d'usage est un levier essentiel pour les PME, car elle offre une protection immédiate, modulable et accessible à tous les budgets.

Cependant, il subsiste toujours une part d'incertitude, même avec les meilleures chartes, il n'est jamais possible de garantir que chaque collaborateur respectera systématiquement les règles. Le facteur humain reste un point faible inhérent à toute organisation, et le risque d'écarts ou d'erreurs involontaires ne peut jamais être totalement éliminé. C'est pourquoi la vigilance, la formation continue et l'accompagnement régulier restent indispensables, même dans un cadre réglementaire bien défini.

En complément des outils internes et des chartes d'usage, il est utile de s'appuyer sur les recommandations du NIST (National Institute of Standards and Technology), un organisme public américain de référence en matière de normalisation et de bonnes pratiques technologiques. Reconnu au niveau international, le NIST publie régulièrement des cadres et guides de sécurité qui sont utilisés aussi bien par les administrations que par de nombreuses entreprises privées à travers le monde. Ses recommandations sont particulièrement pertinentes dans le contexte de l'AI, car elles s'appuient sur des travaux de recherche approfondis et sur l'expérience terrain des professionnels de la cybersécurité.

Le NIST insiste notamment sur l'importance d'une gouvernance globale de l'AI en entreprise, mise en place des politiques claires, des audits réguliers et une documentation systématique de tous les usages de l'AI, y compris ceux qui sortent du cadre officiel (Shadow AI). L'idée est de pouvoir repérer rapidement tout usage à risque, d'attribuer des rôles et responsabilités précis (par exemple en désignant une personne ou une équipe chargée de la conformité), et d'impliquer toutes les parties prenantes, notamment lors de la conception et du déploiement des outils d'intelligence artificielle. Le NIST souligne également l'importance d'un canal de feedback pour signaler toute dérive, ainsi qu'une culture d'entreprise qui favorise la remise en question constructive et la formation continue sur ces sujets. Cela permet non seulement de limiter les risques de biais ou de fuite de données, mais aussi d'anticiper les problèmes avant qu'ils ne surviennent (Milovanovic 2025).

Pour les PME, il n'est pas toujours possible d'avoir des équipes dédiées à la sécurité ou à la conformité. Cependant, même avec des moyens limités, il reste essentiel d'adapter ces recommandations, par exemple en centralisant la gestion de l'AI autour d'un référent clairement identifié, en documentant au minimum les outils utilisés et en prévoyant des points de contrôle réguliers, même informels. Cette démarche pragmatique permet de garder un certain contrôle sur les usages de l'AI, sans alourdir l'organisation (Milovanovic 2025).

Grâce à la démocratisation des solutions techniques comme Microsoft Purview ou FortiDLP, même les PME peuvent aujourd'hui surveiller et contrôler efficacement les fuites de données, y compris celles causées par le Shadow AI, et ce, sans forcément disposer d'une équipe de sécurité dédiée. Les outils technologiques prennent ici le relais des approches organisationnelles ou juridiques, pare-feux nouvelle génération, plateformes de prévention des pertes de données (DLP) ou encore suites de gouvernance avancée. Ces solutions sont devenues essentielles pour empêcher la fuite

d'informations sensibles et garder la main sur ce qui circule au sein du système d'information.

Les pare-feux nouvelle génération ajoutent des fonctionnalités avancées aux filtres classiques, inspection approfondie des paquets (DPI), contrôle des applications, détection d'intrusions (IPS) et surveillance du comportement réseau. Par exemple, Fortigate NGFW (Fortinet 2025a) associe des performances réseau élevées à une sécurité renforcée, avec inspection approfondie des paquets, contrôle applicatif et détection d'intrusions optimisés par l'AI FortiGuard. Cette couche d'AI permet d'analyser le trafic en temps réel, repérer des comportements inhabituels et bloquer des tentatives comme l'exfiltration de données via un outil de Shadow AI. Toutefois, comme ces mécanismes reposent sur l'envoi de certaines métadonnées réseau vers les services FortiGuard, il est nécessaire d'évaluer leur conformité avec la politique interne et les obligations légales.

Si Fortinet est une solution reconnue, d'autres acteurs comme Palo Alto Networks, Cisco ou encore des pare-feux open source comme pfSense peuvent aussi répondre aux besoins, parfois avec des modèles de coûts ou de gestion différents. Le choix dépendra donc autant des capacités techniques recherchées que des contraintes budgétaires et réglementaires propres à l'entreprise.

Figure 8 : Le Next Generation Firewall



L'illustration montre comment un pare-feu nouvelle génération (NGFW) agit comme un véritable filtre intelligent entre internet et les ressources internes de l'entreprise.

Contrairement à un pare-feu traditionnel qui ne fait que bloquer ou autoriser le trafic sur la base d'adresses IP ou de ports, le NGFW offre des fonctions avancées comme l'inspection approfondie des paquets (« deep packet filtering »), des politiques d'accès très granulaires et des services de sécurité intégrés. Il permet ainsi de contrôler finement qui accède à quoi, d'assurer une segmentation précise des accès, et d'enregistrer tous les événements pour une traçabilité totale (« logging and reporting »). Résultat : seules les connexions autorisées passent, les menaces sont bloquées, et tout est documenté, ce qui rend la sécurité beaucoup plus réactive et adaptée aux risques actuels comme le Shadow AI (Houghton 2024).

Un système DLP analyse les flux de données (emails, transferts cloud, stockage local) pour identifier et bloquer la diffusion d'informations critiques. Par exemple, FortiGuard DLP (Fortinet 2025b) utilise des modèles prédéfinis pour détecter des schémas sensibles (numéros de carte, données personnelles) et empêche leur sortie non autorisée. Dans le contexte du Shadow AI, il peut éviter qu'un collaborateur transmette involontairement des données à un service d'AI externe. Toutefois, un DLP ne peut pas tout détecter et doit être régulièrement mis à jour pour rester efficace.

Enfin, Des plateformes comme Microsoft Purview permettent de cartographier, classifier et surveiller toutes les données d'une organisation. Elles offrent des outils d'audit, de blocage d'accès à risque et de conformité avec les réglementations (LPD, RGPD). Ce type d'outil est particulièrement intéressant pour centraliser les règles et garder une visibilité globale, mais son intégration peut nécessiter du temps et une adaptation aux ressources de l'entreprise. (Microsoft [sans date]).

Ces solutions techniques restent très utiles pour les PME comme pour les grandes entreprises. À la suite d'une discussion avec le manager IT et Infrastructure d'une PME interrogée, il en ressort que l'entreprise utilise ces outils, jugés très efficaces et fiables. Cependant, il a souligné que deux vraies options sont réellement applicables pour éviter l'utilisation problématique de l'AI en entreprise, bloquer l'accès à certaines AI avec un pare-feu ou contrôler leur usage. Dans ce deuxième cas, le principal obstacle concerne le respect de la vie privée, car la surveillance accrue peut entrer en conflit avec la législation. Il a également insisté sur le fait qu'il n'existe pas vraiment de lois protégeant les employés, notamment lors de la signature du contrat, l'employeur peut inclure une clause précisant que l'employé sera observé. Ainsi, les PME n'ont pas énormément d'options, mais ces solutions restent actuellement les plus efficaces.

5.3 Méthodologie de l'enquête

Objectif de l'enquête

Dans le cadre de ce travail, une enquête est réalisée auprès des entreprises afin de mieux comprendre leurs pratiques en matière de gestion des risques liés au Shadow AI, ainsi que les mesures de sécurisation des données qu'elles mettent en place.

Cible et secteur

L'enquête cible principalement les PME, car ce sont souvent dans ces structures que l'utilisation du Shadow AI est la moins encadrée et où la sensibilisation aux risques liés à l'intelligence artificielle reste limitée. Les personnes interrogées sont principalement des responsables informatiques, de la sécurité, ou des DPO lorsqu'ils existent.

Elaboration du questionnaire

Le questionnaire a été construit à partir des résultats du travail de recherche, des risques identifiés, ainsi qu'en s'appuyant sur le canvas de ChatGPT pour garantir une structuration claire, cohérente et exhaustive des thématiques abordées.

Diffusion de l'enquête

L'enquête est diffusée en ligne via Microsoft Forms, notamment car le HEG utilise principalement des logiciels Microsoft, ce questionnaire sera disponible entre le 28.07.2025 et le 15.08.2025, auprès de contacts professionnels dans le secteur de la sécurité, ainsi que via les sites ou réseaux d'entreprises.

Traitement et analyse des données

Les réponses seront étudiées pour faire ressortir les tendances principales, comprendre les difficultés concrètes des entreprises face au Shadow AI, et formuler des recommandations adaptées à leur réalité.

Introduction à l'analyse des résultats

Pour cette partie, j'ai choisi de présenter uniquement les tendances principales et mon interprétation des réponses. Les résultats bruts détaillés (graphes et répartitions complètes) sont disponibles en annexe 1. L'objectif est de mettre en avant ce qui ressort le plus clairement de l'enquête et de relier ces éléments aux problématiques du Shadow AI abordées dans ce travail.

5.4 Analyse des réponses

Question 1 : Secteur d'activité de l'entreprise

Parmi les 11 entreprises ayant répondu, la majorité évolue dans le secteur informatique/technologie (5 réponses), suivi des services (3 réponses), de l'administration publique (2 réponses) et enfin d'autres secteurs (1 réponse). Cette forte représentation du domaine IT peut influencer les résultats, car ces structures sont généralement plus sensibilisées aux enjeux de cybersécurité et plus exposées aux risques liés au Shadow AI. À l'inverse, des secteurs absents de l'échantillon, comme l'industrie ou le commerce, auraient pu apporter des perspectives différentes, notamment sur des environnements moins digitalisés.

Question 2 : Taille de l'entreprise

Concernant la taille de l'entreprise, nous constatons que surtout les PME ont été sollicitées et qu'elles représentent la majorité des réponses. Seules 18 % des entreprises interrogées sont de grandes structures de plus de 250 employés, tandis que la majorité appartient à la catégorie des petites et moyennes entreprises.

Question 3 : Connaissance du « Shadow AI ».

Pour la question sur la connaissance du « Shadow AI », 82 % des entreprises déclarent en avoir déjà entendu parler. Ce résultat souligne l'intérêt et la pertinence de mon travail, si une large majorité connaît le concept, cela ne signifie pas pour autant qu'elles disposent des moyens ou des stratégies pour s'en protéger efficacement.

Question 4 : Connaissance d'une utilisation d'outils d'AI non validés officiellement (Shadow AI)

Pour cette question, 64 % des répondants affirment avoir connaissance d'outils d'AI non validés officiellement dans leur entreprise. Ce chiffre montre que le phénomène est déjà bien présent dans les organisations interrogées. 27 % déclarent ne pas être concernés, tandis que 9 % ne savent pas, ce qui laisse supposer que certains usages peuvent exister sans être détectés ou communiqués aux responsables.

Question 5 : Principaux risques associés à l'usage d'outils d'AI non encadrés (Shadow AI)

Les résultats montrent que la principale inquiétude des répondants concerne les fuites de données sensibles (41 %), suivies de près par la non-conformité avec la LPD/RGPD (36 %). 23 % évoquent également le risque de décisions erronées basées sur des outils

non validés. Ces réponses confirment que, pour la majorité des entreprises interrogées, les impacts du Shadow AI sont perçus à la fois sur le plan de la sécurité, de la conformité légale et de la fiabilité des décisions.

Question 6 : Le Shadow AI comme source d'innovation ou d'agilité

45 % des répondants estiment que le Shadow AI peut effectivement être une source d'innovation ou d'agilité, tandis que 36 % nuancent en disant que cela dépend des cas. Seuls 18 % rejettent cette idée. Ces résultats montrent qu'une partie importante des entreprises reconnaît un potentiel positif au Shadow AI, tout en restant prudente quant à son usage, ce qui souligne l'importance d'un encadrement adapté pour exploiter ses avantages sans en subir les risques.

Question 7 : Services ou métiers utilisant le plus les outils d'AI

D'après les réponses, ce sont principalement les équipes IT (8 réponses) et les services opérationnels (6 réponses) qui recourent à des outils d'IA non validés. Les départements marketing et finance se situent loin derrière, avec chacun 2 réponses, suivis par les ressources humaines et quelques autres services, cités une seule fois. On constate donc que le Shadow AI se développe surtout dans des environnements techniques et orientés terrain, où les besoins d'automatisation et d'optimisation sont forts. En contrepartie, ces mêmes secteurs manipulent souvent des données sensibles, ce qui augmente mécaniquement les risques en cas de mauvaise utilisation.

Question 8 : Autorisation d'utilisation d'assistants AI génératifs

La grande majorité des répondants (91 %) indique autoriser l'usage d'assistants IA génératifs, contre seulement 9 % qui préfèrent l'interdire. Ce résultat n'a rien de surprenant : ces outils se sont imposés rapidement dans le monde professionnel grâce aux gains de productivité, de créativité et à l'automatisation qu'ils permettent. Cependant, cet engouement doit aller de pair avec des règles claires et un suivi adapté, afin d'éviter les abus ou les dérives liés au Shadow AI.

Question 9 : Les outils autorisés

Parmi les outils autorisés, ChatGPT et autres assistants AI arrivent largement en tête (45 %), suivis par les outils de traduction automatique tels que DeepL ou Google Translate (36 %). Les outils d'analyse ou de résumé automatique représentent 14 % des autorisations, tandis que les générateurs d'images comme Midjourney ou DALL·E ne sont mentionnés qu'une seule fois (5 %). Ce classement montre que les entreprises privilégient principalement les AI textuelles et de traduction, perçues comme plus

directement utiles dans un cadre professionnel et présentant, a priori, moins de risques que certaines AI génératives d'images.

Question 10 : AI développée spécifiquement pour l'entreprise

Une écrasante majorité des répondants (91 %) ne disposent pas d'une intelligence artificielle développée spécifiquement pour leur entreprise, contre seulement 9 % qui en possèdent une. Cela confirme que, pour la plupart des PME interrogées, l'AI utilisée provient de solutions génériques disponibles sur le marché, ce qui peut limiter la personnalisation et le contrôle, mais réduit aussi les coûts de développement et de maintenance.

Question 11 : Autorisation du BYOD et du BYOAI

Plus de la moitié des entreprises interrogées (55 %) n'autorisent ni l'apport de matériel personnel (BYOD), ni l'utilisation d'AI personnelle (BYOAI). Cependant, 18 % permettent le BYOD uniquement, 9 % autorisent le BYOAI, et 18 % acceptent les deux. Ces résultats montrent que la majorité préfère limiter ces pratiques pour réduire les risques de sécurité et de Shadow AI, mais qu'une minorité ouvre cette possibilité, probablement dans un souci de flexibilité ou d'innovation.

Question 12 : Analyse de conformité réglementaire avant l'intégration d'un outil AI externe

Une majorité des entreprises interrogées (64 %) déclarent réaliser une analyse de conformité au RGPD ou à la LPD avant d'intégrer un outil d'AI externe. Toutefois, 36 % ne procèdent à aucune vérification préalable, ce qui laisse entrevoir un risque non négligeable de non-conformité réglementaire. Ce résultat souligne l'importance de sensibiliser davantage les organisations, en particulier les PME, aux obligations légales liées à l'usage de l'AI.

Question 13 : Détail des démarches de conformité

En lien avec la question précédente, les entreprises qui procèdent à une analyse de conformité avant l'intégration d'un outil d'AI décrivent des approches variées. Certaines privilégient la sensibilisation interne, notamment sur les données pouvant être envoyées à un tiers, tandis que d'autres mettent en place des procédures formalisées telles qu'un comité de validation, un guide interne ou une checklist pour évaluer chaque nouvel outil. Certaines structures optent pour une analyse approfondie, identification précise des données personnelles traitées, vérification de la base légale, et évaluation des mesures de sécurité proposées par le fournisseur. Un répondant souligne également qu'aucune

solution ne garantit totalement l'absence de fuite de données, ce qui pousse son organisation à privilégier des solutions locales (« on-prem ») plutôt que des services cloud.

Ces réponses mettent en évidence que, même lorsque la conformité est vérifiée, les méthodes et le niveau de rigueur peuvent varier considérablement d'une entreprise à l'autre.

Question 14 : Évaluation de l'explicabilité et des risques liés aux modèles d'AI

Une grande majorité des répondants (82 %) déclarent analyser l'explicabilité des modèles ou les risques de sécurité associés à leur utilisation. Cela montre une prise de conscience importante sur la nécessité de comprendre le fonctionnement interne des AI utilisées, afin d'identifier d'éventuels biais ou vulnérabilités. Toutefois, 18 % des entreprises interrogées n'effectuent pas ce type d'évaluation, ce qui peut les exposer à des décisions automatisées non maîtrisées et à des failles de sécurité potentielles.

Question 15 : Réaction face à l'apparition d'un nouvel outil d'AI générative

Les réponses montrent une diversité d'approches face à l'arrivée imprévue d'un nouvel outil d'AI. La solution la plus citée (36 %) est un blocage temporaire jusqu'à ce qu'une évaluation complète soit effectuée, ce qui reflète une volonté de prudence avant toute intégration.

En revanche, 27 % privilégieraient une intégration rapide avec contrôle a posteriori, montrant une ouverture à l'innovation mais avec un suivi après coup. 18 % des répondants opteraient pour un usage libre jusqu'à nouvel ordre, ce qui comporte plus de risques si l'outil n'est pas maîtrisé. Enfin, 18 % ont indiqué une approche autre, cela inclut un blocage car interdit dans l'entreprise par exemple.

Question 16 : Existence d'un cadre ou d'une charte d'usage pour les outils d'AI

La majorité des entreprises (73 %) ont mis en place un cadre ou une charte d'utilisation définissant les règles d'usage de l'AI. Cette formalisation permet d'encadrer les pratiques, de réduire les risques et de clarifier les responsabilités. Toutefois, 27 % des répondants déclarent ne pas disposer d'un tel document, ce qui peut laisser place à des usages non contrôlés et donc potentiellement risqués, notamment dans le contexte du Shadow AI.

Question 17 : Politiques internes encadrant l'usage des assistants AI génératifs

Les résultats montrent que la mesure la plus fréquente est la confidentialité et la non-divulgence de données sensibles (82 %), suivie de la définition d'une typologie des

données autorisées ou interdites (73 %). D'autres mesures incluent la gestion des prompts (27 %), la limitation d'accès à certains profils utilisateurs (27 %) et les procédures d'audit ou de suivi des usages (27 %). Dans les réponses ouvertes, on remarque qu'une entreprise interdit totalement l'usage des AI génératives, tandis qu'une autre se repose uniquement sur le bon sens des utilisateurs. Cette diversité illustre bien l'écart de maturité entre organisations, certaines mettent en place des politiques détaillées, tandis que d'autres préfèrent une approche plus souple, voire aucune politique formelle.

Question 18 : Dispositifs proposés autour de l'AI

Près de la moitié des répondants (47 %) indiquent proposer des actions de sensibilisation à l'AI pour leurs équipes, notamment autour de la compréhension des enjeux et des risques. La formation pratique, incluant la rédaction de prompts et la gestion des risques de fuite de données, arrive en deuxième position avec 27 % des réponses. On note également que 27 % des entreprises n'ont mis en place aucun dispositif particulier, ce qui laisse penser que la sensibilisation et la formation ne sont pas encore systématiques, malgré l'essor de l'AI et du Shadow AI.

Question 19 : Présence d'un DPO ou équivalent

La très grande majorité des répondants (91 %) dispose d'un Data Protection Officer (DPO) ou d'un responsable équivalent chargé de la protection des données. Cela montre que, même dans les structures plus petites, la conformité réglementaire est prise au sérieux. Seule une entreprise indique ne pas avoir de personne dédiée à ce rôle, ce qui peut poser des difficultés en cas de contrôle ou d'incident.

Question 20 : Incidents ou fuites de données

De manière surprenante, aucune des entreprises interrogées ne déclare avoir déjà été confrontée à un incident de sécurité ou à une fuite de données. Ce résultat peut traduire une bonne maîtrise de la sécurité... ou simplement un manque de détection ou de déclaration des incidents.

Question 22 : Accès aux données par des prestataires

La majorité des entreprises indiquent ne pas donner accès à leurs données à des prestataires ou sous-traitants, ou seulement de manière rare et ponctuelle. Cette limitation réduit mécaniquement le risque de fuite de données, mais elle n'exclut pas totalement les menaces, notamment lors des accès exceptionnels.

Question 23 : Utilisation d'outils pour détecter ou prévenir les fuites de données

La majorité des répondants (64 %) indiquent ne pas utiliser d'outils spécifiques pour prévenir les fuites de données. Seules 18 % des entreprises déclarent utiliser des solutions non basées sur l'AI, et aucune n'emploie d'AI dédiée à cet objectif. Ce résultat montre qu'il y a encore un grand écart entre la prise de conscience des risques liés au Shadow AI et la mise en place d'outils concrets de prévention.

Question 24 : AI interne vs AI tierce, niveau de sécurité perçu

Les réponses sont partagées et montrent qu'il n'existe pas de consensus clair. Une partie des répondants estime qu'une AI développée en interne offre un meilleur niveau de sécurité, car les données peuvent rester stockées localement, hors d'atteinte d'Internet, et qu'il est possible de mieux contrôler les accès. Certains mettent aussi en avant la logique de limiter les « mauvaises surprises », surtout pour des données sensibles ou envoyées à l'étranger.

D'autres rappellent que tout dépend des moyens et des compétences internes, une AI mal sécurisée peut être plus vulnérable qu'une solution tierce robuste, certifiée et gérée par un fournisseur reconnu. À l'inverse, quelques entreprises considèrent que les AI tierces sont déjà suffisamment sécurisées.

Enfin, deux réponses soulignent que tout dépend du niveau de sécurité appliqué, qu'il soit interne ou externe, et qu'une solution largement utilisée à l'international, avec des certifications, peut présenter un risque faible si elle est bien encadrée.

Question 25 : Problèmes rencontrés avec l'AI (Shadow AI inclus)

91 % des répondants affirment ne pas avoir rencontré de problème majeur lié à l'AI. Seule une entreprise a signalé un incident, lié à un rapport contenant des données erronées transmis directement à la direction. Ce chiffre peut indiquer soit une bonne maîtrise des outils, soit un manque de détection des incidents.

Question 27 : Connaissance ou utilisation de techniques avancées pour l'AI

La grande majorité (64 %) n'a jamais entendu parler ou utilisé de techniques comme les modèles hybrides ou les méthodes d'explicabilité. Seuls quelques répondants citent un usage pour des besoins métiers comme la classification ou la prédiction. Cela confirme que, dans beaucoup de PME, l'usage de l'AI reste basique.

Question 28 : Raisons principales d'utilisation

Les réponses montrent une utilisation ciblée de ces techniques, principalement pour des besoins métiers comme la classification ou la prédiction. Deux entreprises précisent ne pas les utiliser, mais connaître le sujet, ce qui montre une certaine sensibilisation même sans adoption concrète. L'usage reste donc limité et orienté vers des cas précis plutôt qu'une généralisation à l'ensemble des activités.

Question 29 : Utilisation d'environnements de type "sandbox"

Une majorité (73 %) des répondants déclare utiliser ou avoir expérimenté un environnement sécurisé (sandbox) pour tester et évaluer des solutions d'AI avant leur mise en production. Les réponses ouvertes montrent que ces tests sont souvent systématiques avant chaque intégration ou déploiement, afin de limiter les risques. Quelques-uns disposent même d'environnements dédiés et structurés, bien que les moyens techniques puissent varier (certains indiquent des ressources limitées comme l'absence de GPU).

Question 30 : Fréquence et modalités des tests

Les pratiques sont globalement orientées vers la prudence, la majorité teste avant chaque modification, intégration ou déploiement, souvent via un environnement dédié. Certains précisent que les tests sont réalisés par des utilisateurs avancés ou experts externes, d'autres évoquent un processus allégé lorsque les ressources techniques sont limitées.

Question 31 : Commentaires ou expériences

Un répondant précise que l'usage de l'AI est régulier, mais limité à des tâches simples (réponses à des mails, formules Excel), sans traitement de données sensibles. Bien qu'il n'existe pas de procédure formalisée, il semble que les collaborateurs aient intégré la règle implicite de ne jamais insérer d'informations critiques dans les requêtes.

Conclusion de l'analyse des résultats

L'enquête met en lumière que, même si la majorité des entreprises interrogées, principalement des PME du secteur IT connaissent le Shadow AI et en perçoivent les risques, la mise en place de mesures concrètes reste inégale. Les réponses montrent un décalage entre la prise de conscience et l'action, certaines entreprises disposent déjà de chartes, de procédures d'évaluation et de tests réguliers, tandis que d'autres fonctionnent encore sans cadre formel, laissant la place à des usages non maîtrisés. Si l'AI est globalement vue comme un levier d'innovation, les résultats confirment la nécessité d'un encadrement plus systématique pour éviter les dérives, surtout dans les

structures aux ressources limitées. En somme, la gestion du Shadow AI reste un défi constant, où la vigilance, la formation et la clarté des règles apparaissent comme les clés pour allier innovation et sécurité.

6. Conclusion

L'étude menée sur le Shadow AI montre que ce phénomène, bien que porteur d'opportunités, représente un risque réel pour la sécurité et la conformité des entreprises, en particulier pour les PME. Les mesures préventives combinant chartes internes, sensibilisation et outils techniques modernes comme les pare-feux nouvelle génération ou les systèmes DLP apparaissent comme les meilleures options actuellement disponibles.

Cependant, l'entretien réalisé avec un manager IT de PME met en lumière une réalité plus nuancée, en pratique, deux stratégies principales s'imposent pour limiter les usages problématiques de l'AI, bloquer purement et simplement certains outils ou en contrôler l'utilisation. Or, la seconde option soulève des enjeux majeurs de respect de la vie privée et se heurte à un cadre légal encore flou, où l'employeur peut, par clause contractuelle, mettre en place une surveillance que peu de textes protègent réellement du point de vue des employés.

Ainsi, si les solutions techniques présentées restent pertinentes et efficaces, elles ne sauraient à elles seules résoudre les défis éthiques et juridiques posés par le Shadow AI. À défaut d'un cadre réglementaire clair, les PME devront trouver un équilibre délicat entre innovation, sécurité et respect des droits des collaborateurs.

Avec l'évolution rapide des AI génératives et l'apparition constante de nouveaux outils, le Shadow AI sera probablement de plus en plus difficile à détecter et à encadrer. Les réglementations devront s'adapter, mais ce processus prendra du temps. En attendant, les entreprises devront miser sur la veille technologique, la formation continue et l'adaptation régulière de leurs politiques internes pour rester réactives face aux menaces émergentes.

En ce qui concerne des recommandations auquel les PME devraient faire :

- Formaliser une charte claire sur l'usage des AI, avec des exemples concrets.
- Mettre en place un environnement sécurisé ("sandbox") pour tester tout nouvel outil avant adoption.
- Sensibiliser régulièrement les équipes aux risques du Shadow AI et aux bonnes pratiques.
- Adopter une combinaison de mesures organisationnelles et techniques (contrôle d'accès, journalisation, chiffrement).
- Surveiller de près l'évolution légale et technologique pour anticiper les ajustements nécessaires.

Bibliographie

- ADMIN, [sans date]. Conseiller à la protection des données. *Admin.ch* [en ligne]. [sans date]. Disponible à l'adresse : <https://www.edoeb.admin.ch/fr/conseiller-a-la-protection-des-donnees> [consulté le 23 juin 2025].
- ANSSI, 2025. Appliquer les dix règles d'or préventives. *cyber.gouv.fr* [en ligne]. 25 juin 2025. Disponible à l'adresse : <https://cyber.gouv.fr/appliquer-les-dix-regles-dor-preventives> [consulté le 1 juillet 2025].
- BOESCH, Gaudenz, 2024. Ensemble Learning: A Combined Prediction Model Guide. *viso.ai* [en ligne]. 7 octobre 2024. Disponible à l'adresse : <https://viso.ai/deep-learning/ensemble-learning/> [consulté le 8 juillet 2025].
- BÖGLI, Nicole, [sans date]. Sensibiliser le personnel pour renforcer la cyberprotection. *La mobilière* [en ligne]. [sans date]. Disponible à l'adresse : <https://www.mobiliere.ch/etude/teletravail-cybersecurite-pme-suisse/sensibilisation-personnel> [consulté le 23 juin 2025].
- CHRISTIAN, Jon, 2023. Uber Eats Mocked for Grotesque AI-Generated Food Pictures on Menu. *Futurism* [en ligne]. 12 février 2023. Disponible à l'adresse : <https://futurism.com/the-byte/uber-eats-ai-generated-food-menu> [consulté le 26 juin 2025].
- CNIL, 2024. Sécurité : Impliquer et former les utilisateurs. *CNIL* [en ligne]. 14 mars 2024. Disponible à l'adresse : <https://www.cnil.fr/fr/securite-impliquer-et-former-les-utilisateurs> [consulté le 23 juin 2025].
- CYBERMALVEILLANCE, 2019. Comment protéger vos données en sensibilisant vos collaborateurs ? *Assistance aux victimes de cybermalveillance* [en ligne]. 25 décembre 2019. Disponible à l'adresse : <https://www.cybermalveillance.gouv.fr/tous-nos-contenus/actualites/comment-protoger-vos-donnees-en-sensibilisant-vos-collaborateurs> [consulté le 23 juin 2025].
- DATASULTING, 2023. Algorithme, Intelligence artificielle, machine learning... différences. *DATASULTING* [en ligne]. 22 juin 2023. Disponible à l'adresse : <https://www.datasulting.com/articles/algorithme-intelligence-artificielle-machine-learning-deep-learning-quelle-difference/> [consulté le 8 juillet 2025].
- ETTREDGE, Michael, GUO, Feng et LI, Yijun, 2018. Trade secrets and cyber security breaches. *Journal of Accounting and Public Policy*. Vol. 37, no 6, pp. 564-585. DOI [10.1016/j.jaccpubpol.2018.10.006](https://doi.org/10.1016/j.jaccpubpol.2018.10.006). [consulté le 14 mai 2025].
- FEDLEX, 2007. RO 2007 921 - Loi sur les télécommunications (LTC). [en ligne]. 27 mars 2007. Disponible à l'adresse : <https://www.fedlex.admin.ch/eli/oc/2007/165/fr> [consulté le 22 mai 2025].
- FEDLEX, 2017. RO 2017 2201 - Loi fédérale sur le dossier électronique d... [en ligne]. 4 avril 2017. Disponible à l'adresse : <https://www.fedlex.admin.ch/eli/oc/2017/203/fr> [consulté le 22 mai 2025].
- FEDLEX, 2021. RO 2021 656 - Loi fédérale concernant la lutte contre le ... [en ligne]. 10 novembre 2021. Disponible à l'adresse : <https://www.fedlex.admin.ch/eli/oc/2021/656/fr> [consulté le 22 mai 2025].
- FEDLEX, 2023. RS 235.1 - Loi fédérale sur la protection des données. [en ligne]. 1 septembre 2023. Disponible à l'adresse : <https://www.fedlex.admin.ch/eli/cc/2022/491/fr> [consulté le 19 juin 2025].

FIELD EFFECT, 2024. The real cost of a data breach in 2024. *Field Effect* [en ligne]. 5 août 2024. Disponible à l'adresse : <https://fieldeffect.com/blog/real-cost-data-breach> [consulté le 27 juin 2025].

FOEHRENBACHER, Daniel, 2025. Shadow AI en entreprise : comprendre, prévenir et encadrer ce phénomène croissant. *Bizz and Buzz* [en ligne]. 16 avril 2025. Disponible à l'adresse : <https://www.bizzandbuzz.alsace/2025/04/16/shadow-ai-en-entreprise-comprendre-prevenir-et-encadrer-ce-phenomene-croissant/> [consulté le 1 juillet 2025].

FORTINET, 2025. Pare-feu de nouvelle génération (NGFW). *Fortinet* [en ligne]. 2025. Disponible à l'adresse : <https://www.fortinet.com/fr/products/next-generation-firewall.html> [consulté le 29 juillet 2025].

FORTINET, 2025. Protection des pertes de données : services FortiGuard DLP. *Fortinet* [en ligne]. 2025. Disponible à l'adresse : <https://www.fortinet.com/fr/support/support-services/fortiguard-security-subscriptions/data-loss-prevention.html> [consulté le 29 juillet 2025].

FRENCHWEB.FR, LA REDACTION, 2025. Shadow AI : la nouvelle zone grise de la cybersécurité. *FRENCHWEB.FR* [en ligne]. 14 avril 2025. Disponible à l'adresse : <https://www.frenchweb.fr/shadow-ai-la-nouvelle-zone-grise-de-la-cybersecurite/453014> [consulté le 7 mai 2025].

GEDN, 2024. Quels sont les principaux outils d'AI générative ? *La Grande Ecole du Numérique* [en ligne]. 12 novembre 2024. Disponible à l'adresse : <https://www.grandeecolenumérique.fr/actualites/quels-sont-les-principaux-outils-dai-generative> [consulté le 22 juillet 2025].

GOLDSTEIN, Samuel, 2018. Sanctions et amende en cas de non-respect du RGPD. *LegalPlace* [en ligne]. 23 mars 2018. Disponible à l'adresse : <https://www.legalplace.fr/guides/rgpd-sanction/> [consulté le 26 juin 2025].

GROOTENDORST, Maarten, [sans date]. Exploring Language Models | Maarten Grootendorst | Substack. *Maarten Grootendorst* [en ligne]. [sans date]. Disponible à l'adresse : <https://www.maartengrootendorst.com/blog/stacking/> [consulté le 8 juillet 2025].

HOUGHTON, Sean, 2024. Next-Generation Firewall (NGFW) vs Traditional Firewall. *Aztech* [en ligne]. 25 septembre 2024. Disponible à l'adresse : <https://www.aztechit.co.uk/blog/next-generation-firewall-ngfw-vs-traditional-firewall> [consulté le 29 juillet 2025].

HPE, [sans date]. Qu'est-ce que la gestion de données AI ? *HPE* [en ligne]. Disponible à l'adresse : <https://www.hpe.com/ch/fr/what-is/ai-data-management.html> [consulté le 25 avril 2025].

IBM, 2023. Qu'est-ce que l'AI explicable (XAI) ? | IBM. *IBM* [en ligne]. 29 mars 2023. Disponible à l'adresse : <https://www.ibm.com/fr-fr/think/topics/explainable-ai> [consulté le 8 juillet 2025].

ISEIG, [sans date]. DPO - Data Protection Officer. *ISEIG - Institut Suisse d'Enseignement de l'Informatique de Gestion* [en ligne]. [sans date]. Disponible à l'adresse : https://iseig.ch/iseig_cours/dpo-data-protection-officer/ [consulté le 9 juillet 2025].

KARMANN, Thomas, 2017. Accès à privilèges : la face sombre de l'AIM. *RiskInsight* [en ligne]. 31 mars 2017. Disponible à l'adresse : <https://www.riskinsight-wavestone.com/2017/03/acces-privileges-la-face-sombre-de-laim/> [consulté le 1 juillet 2025].

KRANTZ, Tom, JONKER, Alexandra et MCGRATH, Amanda, 2024. What Is Shadow AI? *IBM* [en ligne]. 25 octobre 2024. Disponible à l'adresse : <https://www.ibm.com/think/topics/shadow-ai> [consulté le 7 mai 2025].

LENOVO, [sans date]. Qu'est-ce que les données en informatique ? | Comment les ordinateurs utilisent-ils les données ? | Lenovo Switzerland. *Lenovo* [en ligne]. [sans date]. Disponible à l'adresse : <https://www.lenovo.com/ch/fr/glossary/data/> [consulté le 6 juillet 2025].

MADDISON, Lewis, 2023. Samsung workers made a major error by using ChatGPT. *TechRadar* [en ligne]. 4 avril 2023. Disponible à l'adresse : <https://www.techradar.com/news/samsung-workers-leaked-company-secrets-by-using-chatgpt> [consulté le 24 juin 2025].

MICROSOFT, [sans date]. Sécurité et gestion des données Microsoft Purview | Sécurité Microsoft. *Microsoft* [en ligne]. [sans date]. Disponible à l'adresse : <https://www.microsoft.com/fr-fr/security/business/microsoft-purview> [consulté le 29 juillet 2025].

MILOVANOVIC, Miroslav, 2025. Detect and Control: Shadow AI in the Enterprise. *Knostic* [en ligne]. 23 juillet 2025. Disponible à l'adresse : <https://www.knostic.ai/blog/shadow-ai> [consulté le 29 juillet 2025].

MURITH, Lionel, 2024. Guide complet : Cybersécurité pour les PME. *gs-global* [en ligne]. 10 septembre 2024. Disponible à l'adresse : https://gs-global.ch/blog_entreprise/hello-world-xqqvy [consulté le 24 juin 2025].

OPENMIND, 2025. Guide en 7 étapes : Mettre en place une politique d'utilisation de l'IA pour votre entreprise | Openmind. *Openmind* [en ligne]. 15 janvier 2025. Disponible à l'adresse : <https://www.openmindt.com/fr/ressources/ai-usage-policy-implementation-guide/> [consulté le 22 juillet 2025].

ORACLE, 2021. Qu'est-ce que l'intelligence artificielle (AI)? *ORACLE* [en ligne]. 15 mai 2021. Disponible à l'adresse : <https://www.oracle.com/ch-fr/artificial-intelligence/what-is-ai/> [consulté le 25 avril 2025].

PFPDT, [sans date]. Dispositions pénales. *Admin.ch* [en ligne]. [sans date]. Disponible à l'adresse : <https://www.edoeb.admin.ch/fr/dispositions-penales> [consulté le 24 juin 2025].

PPDT, [sans date]. Préposé à la protection des données et transparence Jura-Neuchâtel. *PPDT* [en ligne]. [sans date]. Disponible à l'adresse : <https://www.ppd-t-june.ch/fr/Documentation/Guides-pratiques/Classification-des-donnees/Classification-des-donnees.html> [consulté le 9 juillet 2025].

PRICEWATERHOUSECOOPERS (PWC). Protection des données et intelligence artificielle. *PwC* [en ligne]. Disponible à l'adresse : <https://www.pwc.ch/fr/insights/regulation/protection-des-donnees-et-intelligence-artificielle.html> [consulté le 25 avril 2025].

ROBERT, Jérémy, 2020. Machine Learning vs Deep Learning : Quelles différences ? *DataScientest* [en ligne]. 6 juillet 2020. Disponible à l'adresse : <https://datascientest.com/quelle-difference-entre-le-machine-learning-et-deep-learning> [consulté le 25 avril 2025].

ROCHESTON, [sans date]. The Importance of Cybersecurity in Protecting Trade Secrets -. [en ligne]. [sans date]. Disponible à l'adresse : <https://u.rocheston.com/the-importance-of-cybersecurity-in-protecting-trade-secrets/> [consulté le 14 mai 2025].

SAP, [sans date]. Qu'est-ce que l'AI ? | Intelligence artificielle en entreprise. *SAP* [en ligne]. [sans date]. Disponible à l'adresse :

<https://www.sap.com/suisse/products/artificial-intelligence/what-is-artificial-intelligence.html> [consulté le 6 juillet 2025].

SECURITY, ZENDATA CYBER, 2025. Fuite massive de données chez UBS et Pictet | ZENDATA. *ZENDATA Cybersecurity* [en ligne]. 18 juin 2025. Disponible à l'adresse : <https://zendata.security/fr/2025/06/18/fuite-massive-de-donnees-chez-ubs-et-pictet-via-un-fournisseur/> [consulté le 24 juin 2025].

SUISSE.ING, 2024. suisse.ing «Bring Your Own AI». *suisse.ing* [en ligne]. 22 novembre 2024. Disponible à l'adresse : <https://www.suisse.ing/fr/blog/bring-your-own-ai> [consulté le 22 juillet 2025].

SUPPLIERSHIELD, [sans date]. The hidden costs of Inadequate TPRM: A Swiss perspective. *suppliershield* [en ligne]. [sans date]. Disponible à l'adresse : https://www.suppliershield.com/de/post/the-hidden-costs-of-inadequate-tprm-a-swiss-perspective?utm_source=chatgpt.com [consulté le 24 juillet 2025].

UNIVERSITÉ DE GENÈVE, [sans date]. Qu'est-il entendu par données personnelles sensibles ou profil de personnalité au sens de la LIPAD ? *Cureg* [en ligne]. [sans date]. Disponible à l'adresse : <https://cureg.unige.ch/faq/14-quest-il-entendu-par-donnees-personnelles-sensibles-ou-profil-de-personnalite-au-sens-de-la-lipad/> [consulté le 3 avril 2025].

UNIVERSITÉ DE GENÈVE, 2018. Données personnelles ou sensibles ? *Unige* [en ligne]. 11 décembre 2018. Disponible à l'adresse : <https://www.unige.ch/researchdata/fr/stocker/personnelles-sensibles/> [consulté le 3 avril 2025]. Modifié le 2021-09-13

ZÜRICH, [sans date]. Innovation Sandbox for AI. *Kanton Zürich* [en ligne]. [sans date]. Disponible à l'adresse : <https://www.zh.ch/en/wirtschaft-arbeit/wirtschaftsstandort/innovation-sandbox.html> [consulté le 8 juillet 2025].

Annexe 1 : Résultats de l'enquête sur le Shadow AI en entreprise

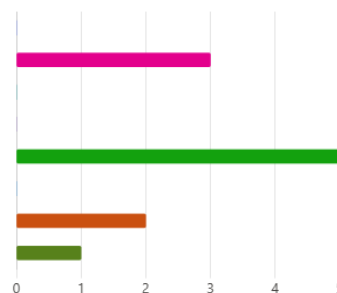
Lien de l'enquête :

https://forms.office.com/Pages/AnalysisPage.aspx?AnalyzerToken=TSRq6RKgwIjR95N5af5p9RtTGIMrudLP&id=JPdYo7LAoE6r-w64xvhOQEdApQbVkl5LINV_pJLr7fZUNDISSTFTRVdESkdIOFIzVDNKRkJPU09aRC4u

1. Secteur d'activité de l'entreprise

[En savoir plus](#)

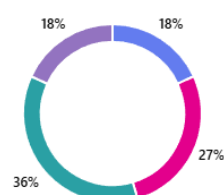
Industrie	0
Services	3
Commerce	0
Santé	0
Informatique/Technologie	5
Éducation	0
Administration publique	2
Autre	1



2. Quelle est la taille de votre entreprise ?

[En savoir plus](#)

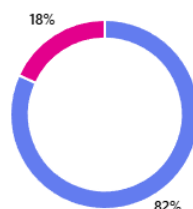
Moins de 10 employés à plein temps	2
Entre 10 et 50 employés	3
De 50 à 250 employés	4
Plus de 250 employés	2



3. Connaissez-vous le terme "Shadow AI" ?

[En savoir plus](#)

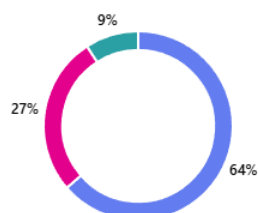
Oui	9
Non	2



4. Avez-vous connaissance d'une utilisation d'outils d'intelligence (IA) non validés officiellement (Shadow AI) dans votre entreprise ?

[En savoir plus](#)

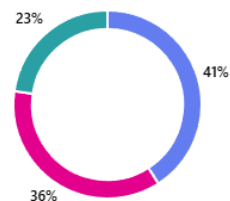
Oui	7
Non	3
Je ne sais pas	1



5. Quels sont selon vous les principaux risques associés à l'usage d'outils d'IA non encadrés (Shadow AI) ?

[En savoir plus](#)

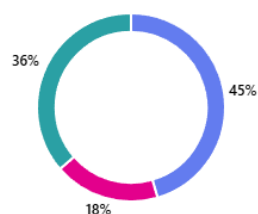
Fuite de données sensibles	9
Non-conformité LPD/RGPD (Lois fédérale / européenne sur la protection des données)	8
Décisions erronées basées sur les outils non validés	5
Aucun	0
Autre	0



6. Pensez-vous que le Shadow AI peut aussi être une source d'innovation ou d'agilité ?

[En savoir plus](#)

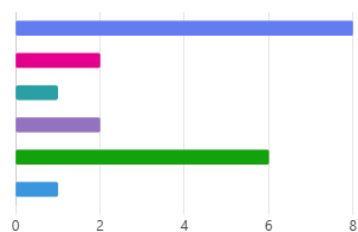
Oui	5
Non	2
Cela dépend des cas	4
Autre	0



7. Par quels services ou métiers ces outils sont-ils le plus utilisés ?

[En savoir plus](#)

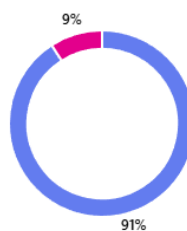
IT	8
Marketing	2
RH	1
Finance	2
Opérations	6
Autre	1



8. Autorisez-vous l'utilisation d'assistants IA génératifs ?

[En savoir plus](#)

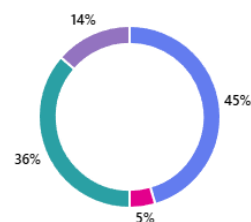
Oui	10
Non	1



9. lesquels sont autorisés ?

[En savoir plus](#)

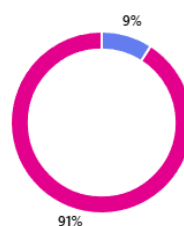
● ChatGPT ou autres assistants IA	10
● Générateurs d'images (ex : midjourney, DALL E)	1
● Outils de traduction automatique (DeepL, Google Translate)	8
● Outils d'analyse ou résumé automatique	3
● Autre	0



10. Disposez-vous d'une IA développée spécifiquement pour les besoins de votre entreprise ?

[En savoir plus](#)

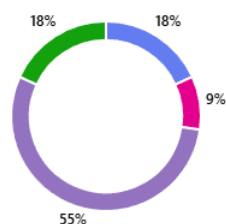
● Oui	1
● Non	10



11. Votre entreprise autorise-t-elle d'apporter son propre matériel informatique (Bring Your Own Device) ou d'utiliser sa propre IA (Bring Your Own Artificial Intelligence) ?

[En savoir plus](#)

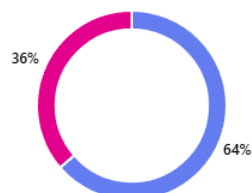
● Oui, BYOD autorisé	2
● Oui, BYOAI autorisé	1
● Oui, les deux sont autorisés	0
● Non	6
● Autre	2



12. Procédez-vous à une analyse de conformité réglementaire (RGPD/LPD) avant l'intégration d'un outil IA externe ?

[En savoir plus](#)

● Oui	7
● Non	4



13. Merci de préciser votre démarche

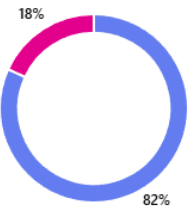
7 Réponses

ID ↑	Nom	Réponses
1	anonymous	Oui, en sensibilisant sur les données qui peuvent être envoyées à un tiers
2	anonymous	Globalement, aucun outil ne permet à l'heure actuelle de garantir l'absence de fuite de données. Nos recherches se tournent quasi exclusivement vers des solutions on-prem plutôt que des services cloud.
3	anonymous	Via comité
4	anonymous	Identification des données personnelles traitées, vérification de la base légale, évaluation des mesures de sécurité proposées par le fournisseur.
5	anonymous	Nous avons un guide interne a suivre pour chaque intégration
6	anonymous	L'analyse est primordiale avant chaque intégration, qu'il s'agisse d'outils d'IA ou d'autres solutions.
7	anonymous	Nous avons une checklist qui permet de savoir si oui ou non nous pouvons l'utilisé

14. Évaluez-vous également l'explicabilité des modèles ou les risques de sécurité ?

[En savoir plus](#)

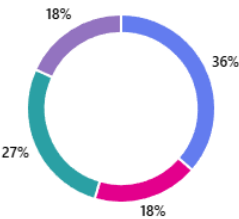
- Oui 9
- Non 2



15. En cas d'apparition soudaine d'un nouvel outil IA génératif dans votre entreprise, quelle serait la réaction la plus probable ?

[En savoir plus](#)

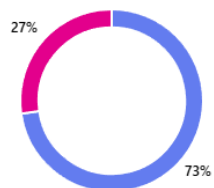
- Blocage temporaire jusqu'à évaluation 4
- Usage libre jusqu'à nouvel ordre 2
- Intégration rapide avec contrôle a posteriori 3
- Autre 2



16. Un cadre d'utilisation ou une charte d'usage a-t-il été défini pour ces outils ?

[En savoir plus](#)

- Oui 8
- Non 3

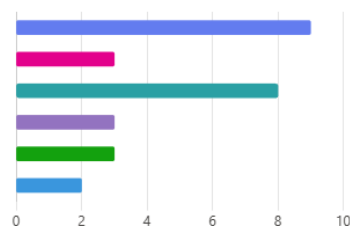


17. Quelles politiques internes encadrent l'usage des assistants IA génératifs dans votre organisation ? (cochez tout ce qui s'applique)

[En savoir plus](#)

- Confidentialité et non-divulgence de données sensibles (Exemple : il est interdit d'insérer des...
- Gestion des prompts (instructions envoyées aux IA) (Exemple : règles sur la façon de formuler les...
- Typologie des données autorisées/interdites (Exemple : certaines données (financières, RH,...
- Limitation d'accès à certains profils utilisateurs (Exemple : seuls certains métiers ou niveaux...
- Procédures d'audit et de suivi des utilisations (Exemple : contrôle régulier des usages pour...
- Autre

9
3
8
3
3
2

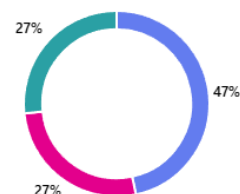


18. Proposez-vous des dispositifs de :

[En savoir plus](#)

- Sensibilisation à l'IA pour vos équipes (compréhension des enjeux, des risques) ?
- Formation à l'IA (bonnes pratiques pour la rédaction de prompts, gestion des risques de fuite de donnée...
- Aucun dispositif
- Autre

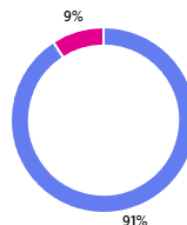
7
4
4
0



19. Votre entreprise dispose-t-elle d'un Protection Officer (DPO) ou d'un responsable équivalent de la protection des données ?

[En savoir plus](#)

- Oui 10
- Non 1



20. Votre entreprise a-t-elle déjà été confrontée à des incidents de sécurité ou à des fuites de données ?

[En savoir plus](#)

● Oui 0
● Non 11



21. Certains de ces incidents étaient-ils spécifiquement liés à l'utilisation de l'IA ou du Shadow AI ?

0 réponses envoyées

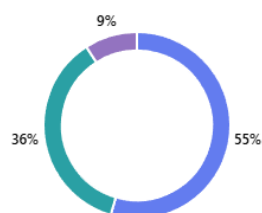
0
Réponses



22. Avez-vous recours à des prestataires ou sous-traitants ayant accès à vos données ?

[En savoir plus](#)

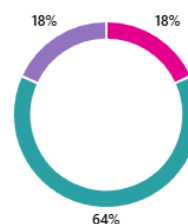
● Accès rare/ponctuel 6
● Accès systématique 0
● Aucun 4
● Autre 1



23. Utilisez-vous des outils pour détecter ou prévenir les fuites de données ?

[En savoir plus](#)

● Oui - IA 0
● Oui - autres outils 2
● Non 7
● Je ne sais pas 2



24. À votre avis, les IA développées en interne offrent-elles un meilleur niveau de sécurité que les solutions IA tierces ? (Merci d'expliquer votre réponse)

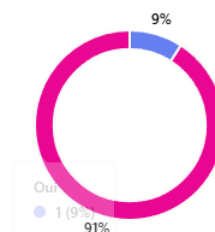
11 Réponses

ID ↑	Nom	Réponses
1	anonymous	Oui de façon différente. En interne on peut avoir des failles, mais l'infrastructure limite l'impact. En externe, ils peuvent offrir plus de sécurité, mais on a aucun contrôle
2	anonymous	Bien sûr. Une IA sur une infrastructure on-prem peut être totalement isolée d'un réseau externe et donc incapable d'exfiltrer la moindre donnée. Le problème rencontré est que cette IA ne va pas avoir toutes les fonctionnalités intéressantes de recherche qui viennent avec un service capable de communiquer sur internet, par exemple.
3	anonymous	Oui car je pense que les données peuvent rester stocker en interne. Le problème est surtout l'utilisation de données potentiellement sensibles lors de l'utilisation de IA externes.
4	anonymous	Non. Le développement des IA tierces sont déjà suffisamment évolué pour ne pas à avoir à créer une IA et la sécurité qui va avec dans la société.
5	anonymous	Pas de point de vue
6	anonymous	Oui, l'utilisation d'une IA développée en interne permet de ne pas avoir de mauvaises surprises surtout pour les données vers l'étranger
7	anonymous	Oui logiquement la sécurité est primordiale
8	anonymous	Le niveau de sécurité dépend surtout des ressources et des compétences disponibles. Une IA interne mal sécurisée peut être plus vulnérable qu'une solution tierce robuste, développée et maintenue par un fournisseur reconnu avec des certifications de sécurité.
9	anonymous	Logiquement Oui, car cela évite d'envoyer des données a n'importe qui.
10	anonymous	Cela dépend du niveau de sécurité mis en place.
11	anonymous	Cela dépend de l'utilisation de l'IA. Une solution externe reconnue et largement utilisée à l'échelle internationale, disposant de certifications et d'audits de sécurité, présente généralement un faible risque de fuite de données.

25. Avez-vous rencontré des cas problématiques liés à l'IA (Shadow AI comprise) tels que : fuite d'information, biais algorithmique, décision non expliquée, incident réglementaire... ?

[En savoir plus](#)

● Oui 1
● Non 10



26. Merci de préciser si possible

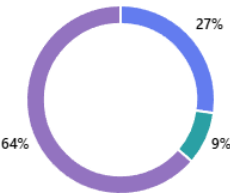
1 Réponses

ID ↑	Nom	Réponses
1	anonymous	Rapport donnant des données erronées et donné directement à la direction.

27. Avez-vous déjà entendu parler ou utilisé dans votre organisation au moins une des techniques suivantes ?

[En savoir plus](#)

Techniques d'ensemble ou hybrides pour l'IA (par exemple : stacking, Random Forest, XGBoost...)	3
Outils ou méthodes d'explicabilité et d'audit des modèles (Explainable AI, LIME, rapports d'audit, etc.)	0
Les deux	1
Non aucune	7
Autre	0



28. Pour quelles raisons principales les utilisez-vous ?

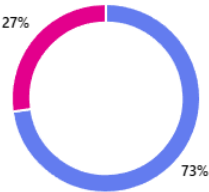
4 Réponses

ID ↑	Nom	Réponses
1	anonymous	Nous ne les utilisons pas, mais je sais de quoi il s'agit.
2	anonymous	Pour des besoins métiers (classification, prédictions)
3	anonymous	Pour la prédiction
4	anonymous	Nous ne l'utilisons pas mais connaissons le sujet

29. Utilisez-vous, ou avez-vous déjà expérimenté, un environnement de type "sandbox" ou des tests sécurisés pour déployer et évaluer des solutions IA avant leur mise en production ?

[En savoir plus](#)

Oui	8
Non	3



30. Merci de préciser comment et/ou à quelle fréquence

8 Réponses

ID ↑	Nom	Réponses
1	anonymous	Tous nouveaux services est d'abord testé dans un environnement de recette par des utilisateurs avancés et des experts externes
2	anonymous	Nous avons un environnement de test limité (sans GPU) pour évaluer les possibilités techniques de mise en place, particulièrement pour la gestion du back-end et la faisabilité technique. Nous testions régulièrement des modèles open-source lorsqu'une percée se produit, mais les modèles testés restent très légers.
3	anonymous	Un environnement dédiés aux tests et à la pré-production afin de valider les résultats qui sert de sandbox avant de passer en prod
4	anonymous	Ne préfère pas commenter
5	anonymous	A chaque modifications nous évaluons et testons sur un environnement dédié
6	anonymous	Avant chaque déploiement nous devons faire des tests d'intégration
7	anonymous	Avant chaque intégration il faut obligatoirement faire des tests
8	anonymous	A chaque déploiement IA nous faisons une série de tests

31. Souhaitez-vous ajouter un commentaire ou partager une expérience liée à l'usage de l'IA dans votre entreprise ?

1 Réponses

ID ↑	Nom	Réponses
1	anonymous	L'usage est régulier, mais sur des tâches simples (réponses de mail, formule excel etc.), mais nous n'utilisons jamais les données sensibles. Nous n'avons pas de procédure en place mais nous savons que nous ne devons pas insérer des données sensibles lors de requêtes.