POURQUOI LES ENTREPRISES ONT-ELLES BESOIN D'UN ESPACE DE TRAVAIL IA PRIVÉ AVANT QUE LES DONNÉES SENSIBLES NE QUITTENT L'ORGANISATION

Préparé par cogify

Mettant en avant l'implémentation, la maintenance et le conseil en cybersécurité de l'espace de travail IA privé alimenté par IVY.

COGIFY

RÉSUMÉ EXÉCUTIF

L'intelligence artificielle fait déjà partie du travail quotidien.

Les employés utilisent des outils d'IA publics pour résumer des documents, reformuler des e-mails, traduire du contenu, préparer des présentations, analyser des rapports, soutenir les ventes, examiner des contrats, écrire du code et accélérer les tâches répétitives. Cela n'est pas automatiquement un problème. L'IA peut améliorer la productivité, réduire le travail manuel et aider les équipes à avancer plus rapidement.

Le problème commence lorsque les entreprises ne savent pas quels outils d'IA leurs employés utilisent, quelles données sont saisies, où ces données sont traitées, si les droits d'accès sont respectés et si **les règles de sécurité de l'entreprise** sont suivies.

Cette utilisation non contrôlée de l'IA est souvent appelée **« shadow AI » (IA fantôme)**.

Le shadow AI crée un risque de **cybersécurité et de protection des données**, car des informations sensibles de l'entreprise peuvent quitter l'organisation par de simples comportements des employés : copier des documents dans des outils d'IA publics, télécharger des fichiers pour les résumer, coller du code source dans des chatbots ou demander à l'IA de reformuler du **contenu confidentiel lié à des clients**.

Pour les équipes techniques, le risque peut être encore plus grave. Les développeurs et les employés informatiques peuvent coller des journaux d'erreurs, des clés API, des jetons d'accès, des variables d'environnement, des détails de bases de données, des architectures système, des **noms d'utilisateurs techniques et des mots de passe** dans des outils d'IA publics, dans le but de déboguer, documenter ou comprendre un problème plus rapidement.

Cela n'est généralement pas dû à de mauvaises intentions. Les employés cherchent simplement à être efficaces. Mais du point de vue de la cybersécurité, le résultat peut tout de même entraîner **une exposition grave des données**.

La solution n'est pas d'interdire complètement l'IA. Les employés utilisent l'IA parce qu'elle les aide à travailler plus vite. Si les entreprises ne fournissent pas une alternative sûre et utile, les employés peuvent continuer à utiliser des outils publics via des comptes personnels, des outils de navigateur, des applications mobiles ou des flux de travail non gérés.

Une meilleure approche est **l'adoption sécurisée de l'IA** grâce à un espace de travail IA privé.

cogify aide les entreprises à implémenter, exploiter et maintenir un espace de travail IA privé où les employés peuvent utiliser l'IA pour les connaissances de l'entreprise, la recherche de documents, la synthèse et le travail quotidien sans envoyer d'informations sensibles vers des outils d'IA publics. L'espace de travail peut être alimenté par IVY, la couche technologique d'IA privée de cogify. La solution orientée client est simple : **un espace de travail IA privé** où les données de l'entreprise restent à l'intérieur de l'entreprise, dans un environnement contrôlé.

Cet environnement peut être hébergé sur **l'infrastructure suisse** de cogify ou installé sur le **serveur propre du client** ou dans son environnement privé, en fonction des besoins du client en matière de sécurité, de protection des données et d'exploitation.

À l'intérieur de l'espace de travail, les données de l'entreprise sont contrôlées grâce à des **espaces de travail séparés**, des droits d'accès basés sur les rôles, des sources de connaissances approuvées et des restrictions claires. **Les employés accèdent** uniquement aux informations qu'ils sont autorisés à voir.

Les documents RH peuvent être restreints au service RH. Les documents financiers peuvent être restreints à la finance et à la direction. Les données des projets clients peuvent être séparées par client ou par équipe projet. La documentation technique peut être limitée aux développeurs et à l'informatique. Les documents du conseil d'administration et de stratégie peuvent être limités à la direction.

Cela donne aux entreprises une alternative pratique à l'utilisation non contrôlée de l'IA publique. **Les employés obtiennent l'IA pour le travail quotidien.**

La direction conserve le contrôle sur les données, les droits d'accès, l'hébergement et les risques de cybersécurité. Les connaissances de l'entreprise restent dans un environnement contrôlé de l'entreprise, garantissant que les données ne quittent jamais l'entreprise.

SOMMAIRE

1. LA NOUVELLE RÉALITÉ : L'IA EST DÉJÀ PRÉSENTE AU LIEU DE TRAVAIL

L'IA N'EST PLUS UNE TENDANCE D'AVENIR. ELLE FAIT DÉJÀ PARTIE DU QUOTIDIEN DES ENTREPRISES.

Les employés utilisent l'IA pour résumer des rapports, réécrire des e-mails, rédiger des propositions, traduire des documents, analyser des feuilles de calcul, relire des contrats, expliquer des concepts techniques, déboguer du code, préparer des réponses clients, générer des comptes-rendus et optimiser les contenus marketing ou commerciaux.

La plupart des employés agissent en toute bonne foi. Ils ne cherchent pas à nuire à l'entreprise. Ils veulent gagner du temps, améliorer la qualité et répondre aux attentes.

Pourtant, d'un point de vue cybersécurité, l'utilisation non maîtrisée de l'IA crée un angle mort critique.

L'entreprise peut ignorer :

- quels outils d'IA sont utilisés
- si les comptes utilisés sont personnels ou professionnels
- si des documents confidentiels sont copiés dans des outils publics
- si des données clients ou employés sont partagées
- si du code source, des identifiants ou l'architecture technique sont exposés
- où les prompts et les fichiers importés sont stockés
- si les droits d'accès sont respectés
- si les réponses générées par l'IA sont fiables et sécurisées
- si l'usage de l'IA respecte les politiques internes

Le défi principal est simple :

La productivité évolue plus vite que la gouvernance.

2. QU'EST-CE QUE LE SHADOW AI ?

L'IA fantôme désigne l'utilisation d'outils d'IA au sein d'une organisation sans l'approbation, la gouvernance, la visibilité ni le contrôle des équipes IT, de la sécurité ou de la direction.

Elle s'apparente à l'IT fantôme, où les employés utilisent des outils ou services cloud non autorisés pour gagner en efficacité. L'IA fantôme est toutefois plus critique. En effet, les outils d'IA ne se contentent pas de stocker des données. Ils les traitent, les résument, les réécrivent, en tirent des déductions et génèrent de nouveaux contenus.

Concrètement, un simple copier-coller peut exposer des données normalement sécurisées au sein des systèmes de l'entreprise.

LES DONNÉES SENSIBLES POUVANT ÊTRE SAISIES DANS DES OUTILS D'IA NON MAÎTRISÉS INCLUENT :

- documents clients
- contrats et dossiers juridiques
- rapports internes
- documents du conseil d'administration
- code source
- clés API et jetons d'accès
- noms d'utilisateur et mots de passe
- identifiants de base de données
- variables d'environnement
- journaux d'erreurs et traces système
- URL internes
- configuration des serveurs et du cloud
- données employés
- documents RH
- données financières
- pipelines commerciaux
- feuilles de route produit
- stratégie d'entreprise
- tickets de support
- procédures de sécurité
- communications clients
- documentation de projet
- informations fournisseurs
- documents liés au secteur public

Dans de nombreux cas, ces informations ne sont pas exposées par une cyberattaque. Elles le sont par les pratiques quotidiennes des employés.

Un collaborateur veut un résumé plus rapide. Un manager veut un e-mail plus clair. Un développeur cherche de l'aide sur son code. Un commercial veut réécrire une proposition. Un employé des finances veut de l'aide pour expliquer un tableur.
L'objectif est la productivité.

Le résultat peut tout de même être une fuite de données.

3. POURQUOI L'IA FANTÔME EST UN ENJEU DE CYBERSÉCURITÉ

L'IA fantôme n'est pas qu'une question d'innovation. C'est aussi un enjeu de cybersécurité. La cybersécurité classique protège les systèmes, fichiers, droits d'accès, e-mails, réseaux et terminaux. Or, les outils d'IA publics offrent une nouvelle voie d'exfiltration pour les données sensibles.

Un document confidentiel peut être protégé au sein du système interne. Mais si un employé en copie le contenu dans un outil d'IA externe, le contrôle d'accès initial n'a plus aucun effet.

L'IA fantôme engendre plusieurs risques.

FUITE DE DONNÉES

Des données métier confidentielles peuvent être partagées avec un outil non approuvé, non supervisé ou non encadré par contrat.

PERTE DE VISIBILITÉ

Les équipes IT et sécurité peuvent ignorer quelles données ont été partagées, par qui, avec quel outil et dans quel but.

PERTE DE CONTRÔLE DES ACCÈS

Des informations issues de documents restreints peuvent être copiées dans un chat IA, puis partagées avec des personnes non autorisées à accéder au contenu original.

RISQUE DE CONFORMITÉ ET DE PROTECTION DES DONNÉES

Les données personnelles, clients, employés ou les informations réglementées peuvent être traitées en dehors des environnements autorisés.

EXPOSITION DE LA PROPRIÉTÉ INTELLECTUELLE

Le code source, l'architecture technique, les idées de produits, la recherche, les algorithmes propriétaires et le savoir-faire métier peuvent être exposés via des prompts ou des fichiers importés.

RÉSULTATS NON FIABLES

Les réponses générées par l'IA peuvent sembler professionnelles, mais contenir des erreurs, des affirmations non fondées, du code vulnérable ou de fausses hypothèses.

SURFACE D'ATTAQUE ÉLARGIE

Les workflows, chatbots et agents IA peuvent être manipulés s'ils sont connectés à des actions sensibles sans vérification, contrôle d'accès, traçabilité ni supervision humaine adéquats.

EXPOSITION DES IDENTIFIANTS ET DE L'INFRASTRUCTURE

Les employés peuvent accidentellement coller des noms d'utilisateur techniques, mots de passe, clés API, jetons d'accès, identifiants de base de données, variables d'environnement, logs serveur, URL internes, configurations cloud ou l'architecture système dans des outils d'IA publics.

Cela est particulièrement dangereux, car ces informations techniques sont directement exploitables par les attaquants. Même un court fichier de log ou un message d'erreur peut révéler des identifiants, des jetons, des noms de systèmes, des URL d'endpoints, des chemins de base de données ou des détails sur l'architecture et l'accès aux systèmes de l'entreprise.

4. SIGNAUX D'ALERTE DU MARCHÉ

L'objectif des exemples du marché n'est pas de dire que chaque incident IA est identique. Certains sont des fuites de données causées par les employés. D'autres concernent l'exploitation de workflows IA. D'autres relèvent de problématiques plus larges de sécurité IA. Ensemble, ils illustrent une leçon claire :

Les entreprises ont besoin de gouvernance, de sécurité et d'environnements contrôlés pour l'usage de l'IA.

Les cas en entreprise montrent que les employés peuvent saisir des informations internes sensibles, comme du code source, dans des outils d'IA publics. Les cas du secteur public prouvent que même les environnements sensibilisés peinent à réagir face au téléchargement de documents sensibles sur des plateformes d'IA publiques. Enfin, les workflows IA démontrent que l'automatisation d'actions sensibles exige une vérification, un contrôle d'accès et une supervision rigoureux.

La leçon n'est pas d'éviter l'IA.

La leçon est que l'IA doit être déployée en toute sécurité.

Les entreprises ont besoin de règles claires, d'outils approuvés, de droits d'accès contrôlés, de modèles d'hébergement sécurisés, de directives pour les employés et d'un suivi continu. Sans cela, l'adoption de l'IA se fait de façon informelle. Des données sensibles peuvent alors quitter l'organisation avant même que la direction n'ait identifié le problème.

5. CONTEXTE SUISSE : POURQUOI L'IA FANTÔME EST CRUCIALE POUR LES ENTREPRISES SUISSES

Actuellement, aucun cas public largement confirmé en Suisse ne montre qu'une entreprise a été piratée spécifiquement à cause de l'usage non contrôlé d'outils d'IA par ses employés. Pourtant, la Suisse fait déjà face à un risque cybersécurité bien réel. Les organisations suisses sont confrontées aux rançongiciels, au phishing, à l'extorsion de données, aux vulnérabilités logicielles, aux attaques de la chaîne d'approvisionnement et aux fuites de données.

L'affaire Xplain a aussi démontré la gravité de l'exposition des données lorsque des informations sensibles liées aux autorités publiques et aux fournisseurs sont divulguées. Bien que ce ne soit pas un incident lié à l'IA, il rappelle fortement qu'en Suisse, les fuites de données, les risques fournisseurs et le manque de contrôle sur les informations sensibles peuvent avoir de lourdes conséquences.

L'IA fantôme doit donc être considérée comme une couche de risque supplémentaire.

Elle ne remplace pas les risques de cybersécurité traditionnels. Elle crée une nouvelle voie d'exfiltration pour les données sensibles, via les pratiques quotidiennes des employés :

- ⚠ copier-coller des documents dans des outils d'IA publics
- ⚠ importer des fichiers pour les résumer
- ⚠ traduire du contenu confidentiel
- ⚠ demander à l'IA d'analyser des données clients
- ⚠ utiliser l'IA pour examiner du code source
- ⚠ coller des données financières dans les prompts de chat
- ⚠ coller des noms d'utilisateur, mots de passe, jetons ou clés API dans les prompts
- ⚠ générer des contenus à partir d'informations internes d'accès restreint

Pour les PME suisses, les fiduciaires, les associations, les sociétés de logiciels, les organisations de santé, les fournisseurs du secteur public et les cabinets de conseil, l'enjeu est crucial : de nombreuses équipes manipulent des données sensibles au quotidien.

Un espace de travail IA privé offre aux collaborateurs une alternative plus sûre aux outils d'IA publics non maîtrisés. Il permet de conserver le savoir de l'entreprise, les droits d'accès et la résidence des données sous strict contrôle.

6. SCÉNARIOS DE RISQUES CONCRETS EN SUISSE

Les scénarios suivants ne sont pas présentés comme des incidents de piratage avérés en Suisse, causés par l'IA. Ils illustrent comment le risque de l'IA fantôme peut émerger dans les entreprises suisses au quotidien.

CABINET FIDUCIAIRE OU COMPTABLE

Les équipes d'une fiduciaire peuvent utiliser l'IA pour résumer des documents clients, rédiger des e-mails, expliquer des données financières ou préparer des notes internes.

Sans espace de travail IA contrôlé, les employés risquent de coller des documents financiers clients, des données de paie, des éléments fiscaux, des factures ou des données personnelles identifiables dans des outils d'IA publics.

Un espace de travail IA privé permet d'isoler le savoir général de l'entreprise, les documents clients, les données de paie, les dossiers fiscaux et les informations de direction dans des espaces restreints. Les utilisateurs n'ont accès qu'aux espaces correspondant à leur rôle et à leurs responsabilités.

Pour des exigences de confidentialité client plus strictes, cet espace peut être déployé sur les serveurs du cabinet ou dans un environnement privé.

PME SUISSE AVEC DES PROJETS CLIENTS

Une entreprise de services suisse peut utiliser l'IA pour rédiger des e-mails clients, résumer des documents de projet, préparer des offres et traduire du contenu interne.

Le risque est que les noms de clients, les tarifs, les détails contractuels, les risques du projet et la stratégie interne soient copiés dans des outils d'IA non maîtrisés.

Un espace de travail IA privé peut être segmenté par département, client, projet ou niveau de sensibilité. Les commerciaux peuvent travailler avec des propositions approuvées. Les équipes projet peuvent accéder à leurs propres documents clients. La direction peut restreindre l'accès à la stratégie, aux tarifs et aux documents du conseil d'administration.

Cela offre aux collaborateurs un espace de travail IA utile, sans pour autant ouvrir l'accès à l'ensemble des informations de l'entreprise.

ENTREPRISE DE LOGICIELS OU DE TECHNOLOGIE

Les développeurs et équipes techniques peuvent utiliser des outils d'IA publics pour déboguer du code, expliquer des erreurs, examiner l'architecture, rédiger de la documentation, générer des cas de test ou comprendre plus rapidement le comportement du système.

Cela peut exposer le code source, l'architecture système, les structures d'API, les logs, les noms d'utilisateur, mots de passe, clés API, jetons, identifiants de base de données, variables d'environnement, URL internes, configurations cloud ou informations CI/CD. Dans ces entreprises, l'IA fantôme dépasse le simple risque documentaire. Elle menace directement l'infrastructure, les accès et la propriété intellectuelle.

Un espace de travail IA privé peut mettre à disposition des développeurs une documentation technique approuvée, des connaissances sur l'architecture, des supports d'intégration, des standards de codage et du contenu d'assistance interne.

Les accès peuvent être restreints par rôle, équipe ou projet. Un développeur sur un produit donné n'a pas forcément besoin d'accéder à tous les environnements clients ni à tous les documents d'architecture.

Mots de passe, clés API, identifiants de production et jetons d'accès doivent rester dans des gestionnaires de mots de passe, des coffres-forts ou des systèmes de gestion des secrets approuvés. L'espace de travail IA peut assister le travail technique, mais ne doit jamais remplacer une gestion sécurisée des identifiants.

ASSOCIATION OU ORGANISATION À BUT NON LUCRATIF

Une association peut gérer des dossiers de membres, des factures fournisseurs, des documents du conseil d'administration, des listes de participants à des événements et des dossiers administratifs. Les employés peuvent utiliser des outils d'IA publics pour résumer des comptes-rendus de réunion, traduire des communications aux membres ou réécrire des documents du conseil.

Un espace de travail IA privé permet de segmenter les informations en espaces dédiés : administration générale, communication aux membres, finances, conseil d'administration et événements. Les documents du conseil restent ainsi accessibles uniquement à ses membres. Les données financières sont réservées à l'équipe financière et à la direction autorisée. Les documents d'événements ne sont visibles que par les organisateurs.

Cela permet aux petites équipes de tirer parti de l'IA sans mélanger tous les documents dans un environnement ouvert.

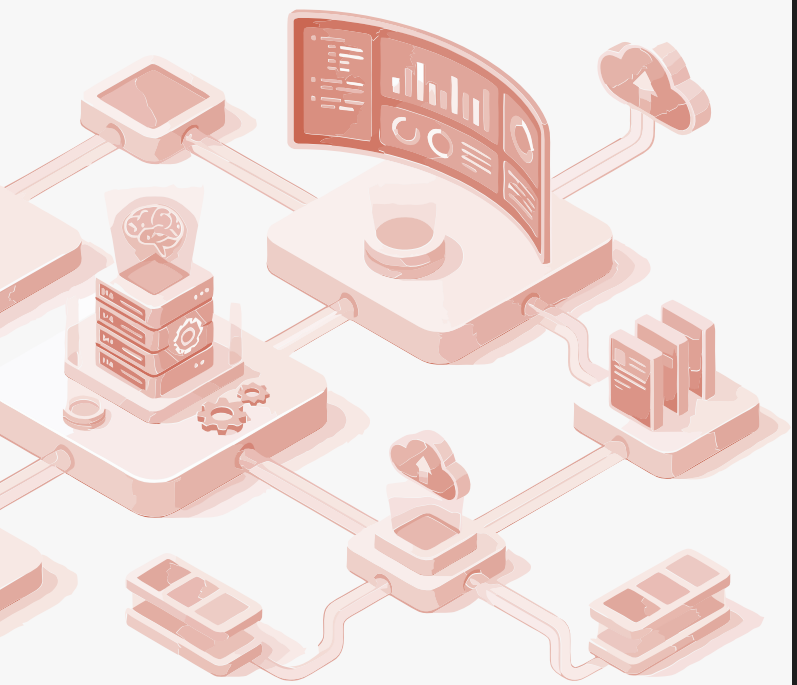
FOURNISSEUR DU SECTEUR PUBLIC OU ORGANISATION LIÉE AUX AUTORITÉS

Les organisations suisses collaborant avec les autorités publiques manipulent souvent des contrats, des documents de projet, des informations relatives aux citoyens, des dossiers techniques ou des données opérationnelles.

Ces informations ne sont pas toujours classifiées. Elles peuvent néanmoins être sensibles, protégées par contrat, cruciales pour la réputation ou stratégiques pour la sécurité.

Un espace de travail IA privé peut être structuré par autorité, client, projet, contrat ou niveau de sensibilité. Cela limite le risque de mélanger les données publiques avec le contenu général de l'entreprise, ou de les rendre accessibles à des utilisateurs non autorisés.

Pour les missions liées aux autorités ou hautement sensibles, cet espace peut être déployé sur les serveurs du client ou sur une infrastructure privée.



ORGANISATION RH, DE RECRUTEMENT OU DU SECTEUR DE LA SANTÉ

Les équipes manipulant des données de santé, RH ou de recrutement peuvent utiliser l'IA pour résumer des CV, rédiger des communications internes, préparer des rapports ou structurer des notes d'entretien.

Ces workflows peuvent impliquer des données personnelles, des informations employés sensibles ou des évaluations confidentielles.

Un espace de travail IA privé permet d'isoler le savoir RH général, les documents de recrutement, la documentation employés, les contenus de la direction et les politiques d'intégration.

Les recruteurs accèdent aux dossiers candidats. Les RH consultent la documentation sur les politiques et processus employés. La direction accède aux documents de décision restreints. Les autres employés n'accèdent qu'aux politiques RH approuvées, aux guides d'intégration ou aux FAQ internes.

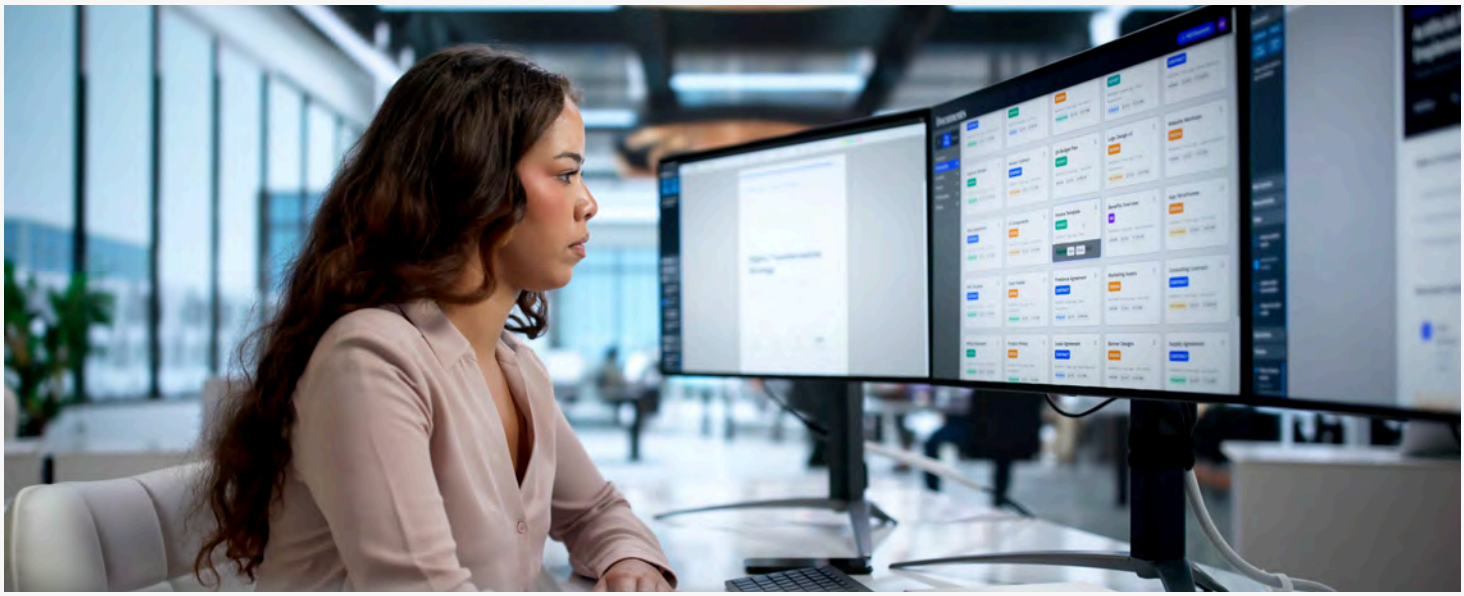
L'objectif n'est pas d'automatiser les décisions humaines sensibles. Il s'agit de sécuriser le travail des équipes, tout en empêchant les données personnelles et internes de finir dans des outils d'IA non maîtrisés.

7. POURQUOI « SIMPLEMENT INTERDIRE L'IA » NE SUFFIT PAS

Certaines organisations tentent de réduire les risques en interdisant les outils d'IA publics. Cette mesure peut être nécessaire dans certains cas, mais elle ne suffit généralement pas comme stratégie à long terme.

Les collaborateurs utilisent l'IA pour travailler plus vite, gagner du temps, réduire les tâches répétitives, améliorer la qualité de leurs écrits, assimiler des informations complexes et rester compétitifs.

Si l'entreprise ne propose pas d'alternative sûre et pratique, les employés risquent de continuer à utiliser l'IA via des comptes personnels, des extensions de navigateur, des applications mobiles ou des plateformes non maîtrisées.



Une interdiction totale peut créer un faux sentiment de sécurité. La direction peut penser maîtriser l'usage de l'IA, alors que les employés continuent d'utiliser discrètement des outils externes.

Une meilleure approche est l'adoption sécurisée de l'IA.

Cela consiste à mettre à disposition des employés un environnement d'IA approuvé par l'entreprise, avec des règles claires, des sources de connaissances maîtrisées, des droits d'accès définis et un accompagnement cybersécurité.

L'objectif n'est pas d'empêcher les collaborateurs d'utiliser l'IA. L'objectif est de leur offrir un espace plus sûr pour l'utiliser.

8. LES ALTERNATIVES NÉCESSAIRES POUR LES ENTREPRISES

Le problème des outils d'IA publics est simple :
Les employés copient souvent des données de l'entreprise dans des systèmes qu'elle ne contrôle pas. cogify résout ce problème en déployant un espace de travail IA privé. Les données sensibles n'ont ainsi plus à quitter l'environnement contrôlé de l'entreprise.

Cet espace peut être alimenté par IVY, la couche technologique d'IA privée de cogify. L'objectif n'est pas de vendre un modèle d'IA. Il s'agit de fournir un espace de travail privé et sécurisé, réellement utilisable au quotidien par les collaborateurs.

Un espace de travail IA privé peut prendre en charge :

- recherche de connaissances internes
- résumé de documents
- questions-réponses internes
- rédaction d'e-mails, de rapports, de propositions et de présentations
- comptes-rendus de réunion et documentation métier
- assistance à la documentation technique
- intégration et partage des connaissances internes
- travail assisté par l'IA et contrôlé entre les services

L'espace de travail est conçu autour de l'environnement réel du client :

- leurs documents
- leurs sources de connaissances
- leurs droits d'accès
- leurs exigences de sécurité
- leur infrastructure
- leurs règles internes
- leurs besoins en protection des données

Selon les exigences du client, l'espace de travail IA privé peut être :

- hébergé sur l'infrastructure suisse de cogify
- installé sur les serveurs du client
- déployé dans un environnement privé défini conjointement avec le client

Le résultat est simple : **les employés disposent de l'IA au quotidien**. Le savoir de l'entreprise reste sous son contrôle.

Les données sensibles n'ont plus à quitter l'environnement de l'entreprise. La direction dispose ainsi d'une alternative plus sûre et mieux maîtrisée que les outils d'IA publics.

9. COMMENT LES DONNÉES SONT CONTRÔLÉES AU SEIN DE L'ENTREPRISE

Un espace de travail IA privé n'est utile que si l'entreprise maîtrise les informations disponibles, les droits d'accès et le lieu de traitement des données. C'est pourquoi cet espace doit intégrer des restrictions dès sa conception.

HÉBERGEMENT CONTRÔLÉ OU DÉPLOIEMENT CÔTÉ CLIENT

L'espace de travail IA privé peut s'exécuter sur l'infrastructure suisse de cogify, ou directement sur les serveurs du client dans un environnement privé. Ainsi, les documents, les prompts, les sources de connaissances et les workflows assistés par l'IA ne transitent plus par des outils d'IA publics.

Le modèle de déploiement est choisi selon les exigences du client :

- Hébergement suisse opéré par cogify
- Installation sur les serveurs du client
- Déploiement dans un environnement privé

Configuration restreinte pour les secteurs sensibles ou les missions liées aux autorités. L'objectif est de conserver les données dans un environnement contrôlé, plutôt que de les disperser sur des comptes d'IA publics non maîtrisés.

ESPACES DE TRAVAIL CLOISONNÉS

Un espace de travail IA privé ne doit pas être un environnement ouvert regroupant toutes les données de l'entreprise. Il doit être structuré en espaces cloisonnés, selon l'organisation, les niveaux de sensibilité et les processus de travail réels.

Exemples :

- Espace général de l'entreprise
- Espace RH
- Espace Finances
- Espace Commercial
- Espace Projets Clients
- Espace Développement
- Espace IT Interne
- Espace Direction
- Espace Conseil d'administration
- Espace Secteur public ou lié aux autorités

Chaque espace dispose de ses propres documents approuvés, utilisateurs et règles d'accès. Cela évite de mélanger les informations sensibles avec le savoir général.

DROITS D'ACCÈS PAR RÔLE

L'accès à chaque espace de travail doit reposer sur le rôle de l'utilisateur et ses besoins métier.

Par exemple :

- tous les employés peuvent accéder aux politiques de l'entreprise et aux FAQ internes approuvées
- les RH peuvent accéder aux processus RH, aux modèles de documents et aux supports de recrutement
- le service financier peut accéder aux procédures financières et à la documentation de facturation
- les développeurs peuvent accéder à la documentation technique et aux standards de codage
- les équipes projet peuvent accéder uniquement à leur propre espace de travail pour les projets clients
- la direction peut accéder à la stratégie, aux documents du conseil d'administration et aux décisions restreintes
- les membres du conseil peuvent accéder uniquement aux documents approuvés par le conseil d'administration

Le principe est simple : Les employés ne doivent voir que ce qu'ils sont autorisés à voir. L'IA ne doit pas devenir un moyen de contourner les droits d'accès existants.

SOURCES DE CONNAISSANCES APPROUVÉES

L'espace de travail IA privé doit uniquement utiliser des sources de connaissances approuvées.

Avant de connecter les documents, l'entreprise doit définir lesquels :

- sont sûrs à utiliser
- sont d'accès restreint
- nécessitent une validation
- doivent être exclus
- doivent être anonymisés ou nettoyés
- contiennent des données personnelles
- contiennent des secrets techniques
- appartiennent à un client ou un projet spécifique

Cela garantit que l'espace IA respecte les règles internes de l'entreprise.

RÉPONSES RESPECTUEUSES DES DROITS D'ACCÈS

L'espace de travail doit respecter les droits d'accès lors de la génération des réponses. Un utilisateur ne doit obtenir des réponses qu'à partir des documents et sources de connaissances auxquels il a accès.

Par exemple :

- le marketing ne doit pas recevoir de réponses basées sur les documents salariaux RH
- les développeurs ne doivent pas recevoir de réponses basées sur les documents stratégiques du conseil
- une équipe projet ne doit pas recevoir de réponses basées sur les fichiers restreints d'un autre client
- les employés ne doivent pas recevoir de réponses basées sur les documents financiers ou juridiques

C'est l'une des différences majeures entre un espace IA privé contrôlé et l'usage non maîtrisé d'outils d'IA publics. Les outils d'IA publics ignorent les droits d'accès de l'entreprise. Un espace IA privé, lui, peut être structuré autour d'eux.

SECRETS TECHNIQUES ET IDENTIFIANTS

Les secrets techniques exigent une gestion rigoureuse. Un espace IA privé ne doit pas servir à stocker sans contrôle des mots de passe, clés API, identifiants de production ou jetons d'accès. Ils doivent rester dans des gestionnaires de mots de passe, coffres-forts ou systèmes de gestion des secrets approuvés.

Pour les équipes de développement et l'IT, l'entreprise doit définir des règles claires :

CE QU'IL NE FAUT PAS FAIRE :

- coller des mots de passe dans les prompts IA
- coller des clés API dans les prompts IA
- coller des identifiants de production dans les prompts IA
- importer des logs non vérifiés susceptibles de contenir des secret
- importer des fichiers d'environnement sans validation

CE QU'IL FAUT FAIRE :

- nettoyer la documentation technique avant de la connecter à l'espace de travail
- conserver les secrets dans des outils dédiés de gestion des secrets

L'espace de travail IA peut assister le travail technique, mais il ne doit pas remplacer une gestion sécurisée des identifiants.

10. COMMENT COGIFY AIDE LES ENTREPRISES

cogify aide les entreprises à adopter l'IA en toute sécurité en combinant déploiement d'espaces de travail IA privés, conseil en cybersécurité, hébergement suisse et support technique à long terme.

IMPLÉMENTATION D'UN ESPACE DE TRAVAIL IA PRIVÉ

cogify conçoit et déploie un espace de travail IA privé. Il s'adapte aux besoins métier, à la sensibilité des données, à l'infrastructure et aux exigences de sécurité du client.

Cet espace peut être alimenté par IVY, la couche technologique d'IA privée de cogify.

Cela inclut l'installation, la configuration, la connexion des sources de connaissances, la structuration des espaces, la logique des droits d'accès et la planification du déploiement.

HÉBERGEMENT SÉCURISÉ OU DÉPLOIEMENT CÔTÉ CLIENT

L'espace de travail IA privé peut être hébergé sur l'infrastructure suisse de cogify ou installé sur les serveurs du client dans un environnement privé.

Le modèle de déploiement est choisi selon les exigences du client en matière de protection des données, de cybersécurité, de conformité et d'opérations.

CONCEPTION DES ESPACES ET DES DROITS D'ACCÈS

cogify aide ses clients à définir la séparation des données au sein de l'espace de travail. Cela inclut les espaces par département, par projet, par client, pour la direction, ainsi que les espaces restreints pour les contenus sensibles.

L'objectif est de garantir que les utilisateurs n'accèdent qu'aux informations qu'ils sont autorisés à consulter.

iVY

CONSEIL EN CYBERSÉCURITÉ

cogify aide les entreprises à réduire les risques de l'IA fantôme via le conseil en cybersécurité, la révision des accès, la cartographie des données sensibles, l'élaboration de politiques, la sensibilisation des collaborateurs et l'audit de sécurité des workflows IA.

Cela inclut la définition de règles pour les données confidentielles, les informations clients, les données RH, les données financières, le code source, les identifiants, les logs et les secrets techniques.

MAINTENANCE ET AMÉLIORATION CONTINUE

L'adoption sécurisée de l'IA ne s'arrête pas à l'installation initiale. cogify peut accompagner l'exploitation continue de l'espace de travail IA privé via :

- mises à jour techniques
- mises à jour de sécurité
- ajustements des droits d'accès
- accompagnement des nouveaux utilisateurs ou services
- révision des sources de connaissances connectées
- amélioration de la structure des documents
- analyse des retours utilisateurs
- optimisation des workflows IA
- audit périodique de cybersécurité

Cela est crucial car les cas d'usage de l'IA évoluent rapidement. De nouveaux services demandent des accès, de nouveaux documents sont créés, de nouveaux risques émergent et les politiques internes évoluent. Grâce à un accompagnement continu, l'espace de travail reste utile, sécurisé et parfaitement aligné avec la réalité métier.

11. CHECKLIST DE SÉCURITÉ IA POUR LA DIRECTION

Les entreprises doivent se demander :

- Savons-nous quels outils d'IA les employés utilisent ?
- Les employés utilisent-ils des comptes IA personnels pour travailler ?
- Savons-nous quelles données sont collées dans les prompts d'IA ?
- Avons-nous des règles pour les données clients, RH, financières et le code source ?
- Avons-nous des règles pour les noms d'utilisateur, mots de passe, clés API et logs ?
- Disposons-nous d'un espace de travail IA approuvé ?
- L'espace de travail est-il segmenté en zones restreintes ?
- Les sources de connaissances de l'entreprise sont-elles régies par des droits d'accès ?
- Les utilisateurs accèdent-ils uniquement aux espaces liés à leur rôle ?
- Savons-nous où sont hébergées les données liées à l'IA ?
- Formons-nous les collaborateurs à une utilisation sécurisée de l'IA ?
- Surveillons-nous et évaluons-nous régulièrement les risques liés à l'IA ?
- Les sorties générées par l'IA sont-elles vérifiées avant d'être intégrées dans des workflows sensibles ?
- Savons-nous quels outils d'IA sont connectés aux systèmes de l'entreprise ?
- La responsabilité de la gouvernance de l'IA est-elle clairement définie ?
- Notre espace de travail IA bénéficie-t-il de maintenance et de support ?

Si la réponse est « non » à plusieurs de ces questions, l'entreprise est peut-être déjà exposée au risque d'IA fantôme.

12. PREMIÈRES ÉTAPES RECOMMANDÉES

Pour les entreprises concernées par l'IA fantôme, cogify recommande les premières étapes suivantes.

ÉTAPE 1: IDENTIFIER LES USAGES ACTUELS DE L'IA

Comprendre quels outils d'IA les employés utilisent déjà, quels services les utilisent et dans quel but.

ÉTAPE 2: CLASSIFIER LES DONNÉES SENSIBLES

Définir quelles données ne doivent jamais être saisies dans des outils d'IA publics ou non maîtrisés. Cela doit inclure les données clients, les données employés, les données financières, le code source, les identifiants, les noms d'utilisateur, les mots de passe, les clés API, les jetons d'accès, les logs et les détails de configuration technique.

ÉTAPE 3: RÉVISER LES DROITS D'ACCÈS

Déterminer qui doit avoir accès à quelles sources de connaissances internes.

ÉTAPE 4: CONCEVOIR LES RESTRICTIONS DES ESPACES DE TRAVAIL

Définir les espaces de travail nécessaires. Par exemple :

- espace général de l'entreprise
- espace RH
- espace Finances
- espace Développement
- espace Projets Clients
- espace Direction
- espace Conseil d'administration

Pour chaque espace, définir qui peut y accéder, quels documents y sont inclus et quelles données doivent rester restreintes.

ÉTAPE 5: CHOISIR LE BON MODÈLE DE DÉPLOIEMENT

Décider si l'espace de travail IA privé doit s'exécuter sur l'infrastructure suisse de cogify, sur les serveurs du client ou dans un autre environnement privé.

ÉTAPE 6: IMPLÉMENTER, FORMER ET GOUVERNER

Configurer l'espace de travail, connecter les sources de connaissances approuvées, définir les droits d'accès, former les utilisateurs et établir des règles de gouvernance autour de l'utilisation de l'IA.

ÉTAPE 7: MAINTENIR ET AMÉLIORER EN CONTINU

Examiner les usages, améliorer les sources de connaissances, mettre à jour les droits d'accès, accompagner les collaborateurs et adapter l'espace de travail à l'évolution des besoins métier et de sécurité.

CONCLUSION

L'IA est déjà présente dans l'entreprise. La question n'est plus de savoir si les collaborateurs l'utiliseront, mais si l'entreprise peut leur offrir un environnement sûr et maîtrisé pour le faire. L'utilisation non maîtrisée de l'IA engendre des risques en matière de cybersécurité, de protection des données et de conformité. Se contenter d'interdire l'IA n'est pas une solution réaliste à long terme.

Les entreprises doivent adopter l'IA en toute sécurité. cogify les aide à déployer, exploiter et maintenir des espaces de travail IA privés. Ces espaces accompagnent le travail quotidien réel tout en réduisant la dépendance envers les outils d'IA publics non maîtrisés. Cet espace peut être alimenté par IVY, la couche technologique d'IA privée de cogify. La solution proposée aux clients est simple : un espace de travail IA privé où les données n'ont pas à quitter l'environnement contrôlé de l'entreprise.

Les informations de l'entreprise peuvent être réparties dans des espaces de travail cloisonnés. Les accès peuvent être définis par rôle, service, projet ou niveau de sensibilité. Les documents RH restent au service RH. Les documents financiers restent au service financier. Les données clients demeurent dans les espaces dédiés aux clients. La documentation technique reste réservée aux développeurs et aux équipes IT autorisés. Les documents de la direction et du conseil d'administration restent strictement réservés aux dirigeants. C'est ainsi que les entreprises peuvent mettre l'IA à disposition de leurs collaborateurs sans perdre le contrôle de leur savoir.

Conjugué à du conseil en cybersécurité, un accompagnement en gouvernance et une maintenance à long terme, un espace de travail IA privé offre aux organisations un moyen concret de réduire les risques liés à l'IA fantôme, de protéger les informations sensibles et de bâtir des fondations solides pour un usage responsable de l'IA.

PRÊT À ADOPTER L'IA SANS PERDRE LE CONTRÔLE DE VOS DONNÉES ?

Contactez cogify pour déployer et sécuriser votre espace de travail IA privé. Nous évaluons vos usages actuels, protégeons vos données sensibles et vous offrons une alternative maîtrisée aux outils d'IA publics.

Demander une consultation en sécurité de l'IA

AVERTISSEMENT IMPORTANT

Les cas d'usage suisses présentés dans ce livre blanc sont des scénarios métier concrets, fondés sur les pratiques courantes des PME suisses, des fiduciaires, des associations, des entreprises technologiques, des organisations du secteur de la santé et des environnements liés au secteur public.

Ils ne sont pas présentés comme des incidents de piratage avérés en Suisse, causés par l'IA. Les références suisses confirmées servent à illustrer le paysage global de la cybersécurité en Suisse, notamment les risques liés à l'exposition des données, aux rançongiciels, au phishing, à la chaîne d'approvisionnement et à la gouvernance. L'IA fantôme doit être considérée comme une nouvelle couche de risque susceptible d'accroître l'exposition des données si les entreprises ne mettent pas en place des workflows IA sécurisés et approuvés.

Le modèle de déploiement d'un espace de travail IA privé doit être défini en fonction des exigences de chaque client en matière de sécurité, de protection des données, de conformité et d'opérations. Pour certains clients, l'infrastructure suisse opérée par cogify peut convenir. Pour d'autres, une installation sur les serveurs du client ou dans un environnement privé sera préférée.

RÉFÉRENCES ET LECTURES COMPLÉMENTAIRES

1. CENTRE NATIONAL DE CYBERSÉCURITÉ SUISSE : RAPPORT SEMESTRIEL 2025/1

Le NCSC a recensé 35 727 signalements d'incidents cyber au premier semestre 2025 et a mis en évidence les rançongiciels, l'extorsion de données, le phishing, les attaques DDoS, les vulnérabilités logicielles et les attaques de la chaîne d'approvisionnement comme des cybermenaces significatives pour la Suisse.

Lien : https://www.ncsc.admin.ch/ncsc/fr/home/dokumentation/berichte/lageberichte/halbjahresbericht-2025-1.html

2. CENTRE NATIONAL DE CYBERSÉCURITÉ SUISSE : RAPPORT D'ANALYSE DES DONNÉES DE LA CYBERATTAQUE CONTRE XPLAIN

Le rapport du NCSC détaille son analyse des données publiées sur le darknet à la suite de la cyberattaque visant Xplain, un grand fournisseur de services informatiques pour les autorités nationales et cantonales. Bien qu'il ne s'agisse pas d'un incident lié à l'IA, il constitue un exemple suisse pertinent pour illustrer pourquoi l'exposition des données sensibles, les risques fournisseurs et les environnements technologiques tiers sont des enjeux cruciaux.

Lien : https://www.ncsc.admin.ch/ncsc/fr/home/dokumentation/berichte/fachberichte/bericht-datenanalyse-xplain.html

3. TOP 10 DE L'OWASP POUR LES APPLICATIONS BASÉES SUR LES GRANDS MODÈLES DE LANGAGE

Le projet GenAI Security de l'OWASP identifie et documente les risques de sécurité et de sûreté associés aux technologies d'IA générative, notamment les grands modèles de langage et les applications pilotées par l'IA. Les risques concernés incluent l'injection de prompts, la divulgation d'informations sensibles, le traitement non sécurisé des sorties et l'autonomie excessive.

Lien : https://owasp.org/www-project-top-10-for-large-language-model-applications/

4. AFFAIRE DU TÉLÉVERSEMENT DE DOCUMENTS CONFIDENTIELS SUR LA VERSION PUBLIQUE DE CHATGPT PAR LA CISA

ITPro a rapporté que le directeur par intérim de l'Agence américaine pour la cybersécurité et la sécurité des infrastructures (CISA) a téléchargé des documents sensibles classés « À usage officiel uniquement » dans une version publique de ChatGPT, suscitant des inquiétudes quant à l'IA fantôme et au contrôle des données.

Lien : https://www.itpro.com/security/data-protection/cisas-interim-chief-uploaded-sensitive-documents-to-a-public-version-of-chatgpt-security-experts-explain-why-you-should-never-do-that

5. CONSEIL FÉDÉRAL SUISSE : ORIENTATIONS EN MATIÈRE DE RÉGULATION DE L'INTELLIGENCE ARTIFICIELLE EN SUISSE

Le Conseil fédéral a annoncé l'intention de la Suisse de ratifier la Convention sur l'intelligence artificielle du Conseil de l'Europe et de la mettre en œuvre via une approche sectorielle, appuyée le cas échéant par des mesures non contraignantes.

Lien : https://www.admin.ch/fr/nsb?id=104110