IA et données juridiques : la grille d'arbitrage

Avant de brancher une IA sur tes dossiers : une matrice de décision par sensibilité de donnée + 10 questions de diligence sur tout connecteur MCP.

BLYX-AGENCY.COM · 22 JUILLET 2026

⚖️ **IA ET DONNÉES JURIDIQUES : LA GRILLE D'ARBITRAGE** — par Pierre Cordelier.

Un cadre pour décider, donnée par donnée, ce que tu autorises, ce que tu passes en lecture seule, et ce que tu interdis. Daté et sourcé, à jour au 22 juillet 2026.

Je ne suis pas juriste. Je conçois et j'installe des connexions IA en entreprise. Cette grille traite la plomberie et la gouvernance — le périmètre, pas le droit. Chaque décision se valide avec ton DPO, ton RSSI et ton référent déontologie.

1. Le problème n'est pas l'IA. C'est le périmètre.

La question « faut-il autoriser l'IA ? » est déjà dépassée. Tes équipes l'utilisent.

- **61 %** des utilisateurs d'IA en entreprise y accèdent via un **compte personnel** au moins une fois par semaine, 38 % chaque jour

(Microsoft France × YouGov, janvier 2026, 657 cadres et dirigeants).

- **46 %** des professionnels du droit le font **malgré une directive interne contraire** (CSA Research × Lefebvre Dalloz, 627 répondants, terrain janvier-mars 2026).

Le secret professionnel ne tient pas dans un historique de chat grand public. Le vrai sujet n'est donc pas d'ouvrir ou de fermer la porte — c'est de savoir par où sortent les données, et de reprendre la main sur ce périmètre.

Jusqu'en 2026, il n'y avait pas de bonne porte : aucune IA n'avait d'accès légitime au système documentaire du cabinet. Alors les documents sont sortis par la fenêtre — copier-coller dans une interface publique, zéro trace.

2. Ce qui a changé en 2026

Les systèmes documentaires du droit ont ouvert leurs propres connecteurs.

- **iManage** a rendu son serveur MCP disponible en général le **14 mai 2026**.
- **NetDocuments** a livré sa connectivité MCP le **1er avril 2026**, en collaboration avec Anthropic sur un MCP dédié au secteur juridique.

Le MCP (*Model Context Protocol*) est la prise standard entre une IA et un système. Depuis décembre 2025, il n'appartient plus à un éditeur : il a été donné à une fondation sous la **Linux Foundation** (membres

fondateurs et soutiens : Anthropic, OpenAI, Google, Microsoft, AWS, Bloomberg). Ce n'est plus un pari sur un fournisseur, c'est un standard.

Le principe de ces connecteurs juridiques change tout par rapport au copier-coller :

- le **document ne sort pas** du système documentaire ;
- l'IA **hérite des droits exacts de l'utilisateur** (OAuth délégué), murailles de Chine (*ethical walls*) comprises ;
- chaque accès est **journalisé** ;
- **aucun export en masse** n'est possible.

C'est l'exact inverse d'un prompt collé dans une fenêtre publique. Mais « c'est possible » ne veut pas dire « c'est sans condition ». D'où la grille.

3. La matrice de décision

Croise la **sensibilité de la donnée** (en lignes) avec le **type de branchement** (en colonnes). La cellule te donne la décision par défaut.

Sensibilité de la donnée	IA grand public (compte perso)	Connecteur MCP — données hors UE	Connecteur MCP — UE, OAuth, journalisé, lecture seule
1. Données publiques (jurisprudence publiée, textes, doctrine publique)	✅ Autoriser	✅ Autoriser	✅ Autoriser

Sensibilité de la donnée	IA grand public (compte perso)	Connecteur MCP — données hors UE	Connecteur MCP — UE, OAuth, journalisé, lecture seule
2. Interne non confidentiel (clausiers anonymisés, notes non nominatives)	⚠️ Éviter	🟠 Lecture seule	✅ Autoriser
3. Données personnelles RGPD (RH, parties identifiables, contacts)	⛔ Interdire	⛔ Interdire	🟠 Lecture seule + base légale + DPA + minimisation
4. Secret professionnel / dossiers clients	⛔ Interdire	⛔ Interdire	🟠 Lecture seule + murailles de Chine + journalisation + validation humaine + information client
5. Ultra-sensible (M&A en cours, pénal, données art. 9 RGPD)	⛔ Interdire	⛔ Interdire	🔴 Cas par cas — décision DJ + DPO documentée, hors automatisation par défaut

4. Comment lire les trois colonnes

Colonne 1 — IA grand public via compte personnel. C'est le *shadow AI* d'aujourd'hui. Aucun contrôle, aucune trace, données potentiellement réutilisées. Acceptable uniquement pour de la donnée déjà publique. Dès qu'il y a de la donnée personnelle ou couverte par secret, c'est non.

Colonne 2 — Connecteur MCP, mais données hors UE. Le branchement est propre (OAuth, permissions), mais si les données sont traitées ou stockées hors de l'Union, tu ajoutes un transfert international à gérer. Pour de la donnée personnelle ou confidentielle, cette colonne reste rouge tant que la résidence n'est pas réglée.

Colonne 3 — Connecteur MCP souverain, journalisé, en lecture seule. C'est le seul setup qui ouvre la porte aux données sensibles — et encore, sous conditions cumulatives (les mentions dans les cellules ne sont pas optionnelles). Plus la donnée est sensible, plus la liste de conditions s'allonge, jusqu'au cas par cas documenté pour le niveau 5.

La logique de la grille : on ne décide pas « IA oui/non ». On décide, pour *chaque* type de donnée, le couple le plus prudent entre le branchement et le niveau d'accès.

5. Les 10 questions de diligence

À poser au fournisseur ou à ta DSI **avant** de brancher un connecteur MCP. Exige une réponse écrite. Chaque question vise un point où un vrai produit du marché coince aujourd'hui.

1. **Identité & permissions** — Le connecteur agit-il avec les droits *exacts* de l'utilisateur (OAuth délégué), ou avec un accès de service élargi ? L'IA peut-elle voir plus que la personne ne voit déjà ?
2. **Murailles de Chine** — Les cloisonnements déontologiques existants sont-ils respectés, dossier par dossier ?
3. **Lecture vs écriture** — Peut-on verrouiller le connecteur en lecture seule et bloquer écriture, envoi et suppression ? Qui peut lever ce verrou ?

4. **Résidence des données** — Où sont-elles hébergées et traitées ? Transitent-elles ou sont-elles stockées hors UE ?
5. **Exposition réseau** — Le serveur est-il joignable *uniquement* depuis ton réseau, ou doit-il être exposé sur l'internet public pour fonctionner ?
6. **Journalisation & audit** — Chaque accès est-il tracé, horodaté, exportable vers ton SIEM et conservé ? Existe-t-il un angle mort (surface non journalisée) ?
7. **Validation humaine** — Une action à conséquence (envoi, écriture, suppression) exige-t-elle une approbation humaine explicite, ou peut-elle s'exécuter seule ?
8. **Rétention & entraînement** — Tes données servent-elles à entraîner un modèle ? Quelle durée de conservation ? Un accord Zero Data Retention existe-t-il, et couvre-t-il *ce* produit précis ?
9. **Contenu non fiable** — Dans une même session, l'agent peut-il lire des données confidentielles **ET** du contenu externe non vérifié (email entrant, pièce jointe) **ET** communiquer vers l'extérieur ? Ces trois réunies ouvrent la porte à l'exfiltration.
10. **Conformité** — Un DPA au titre de l'article 28 du RGPD est-il signé ? Une analyse de transfert (post-Schrems II) est-elle documentée ? L'obligation de transparence de l'article 50 de l'AI Act (applicable le **2 août 2026**) est-elle couverte ?

6. Les 3 pièges que les vendeurs passent sous silence

Ces points sont documentés sur les pages officielles des éditeurs, mais rarement mis en avant. Vérifie-les toi-même.

- **La résidence des données.** Certaines surfaces grand public stockent les données aux États-Unis, sans option de résidence UE native. Pour un cabinet soumis au secret, c'est éliminatoire tant que ce n'est pas contractuellement réglé (question 4).
- **Le connecteur exposé sur l'internet public.** Sur plusieurs clients IA, un connecteur MCP « custom » est joint depuis le cloud du fournisseur, pas depuis ton poste : le serveur doit donc être accessible publiquement. Un MCP derrière ton VPN peut ne pas fonctionner — et un serveur exposé sans authentification solide est une porte d'entrée (question 5).
- **L'angle mort de journalisation.** Certaines surfaces agentiques ne sont pas encore capturées par les API de conformité et d'audit. Tu crois tout tracer, il te manque une pièce (question 6).

Deux principes de fond, valables quel que soit le produit :

- **La validation humaine est trop souvent facultative.** Dans la spécification MCP, l'approbation humaine avant une action est un *should*, pas un *must*. En janvier 2026, une preuve de concept publique a exfiltré des documents réels via un fichier piégé, sans qu'aucune approbation humaine ne soit requise. Impose le *human-in-the-loop* toi-même sur toute action à conséquence.
- **La triade létale.** Si un même agent cumule accès à des données privées, exposition à du contenu non fiable, et capacité de sortie vers l'extérieur, l'exfiltration devient possible. Casse au moins un des trois (question 9).

7. La règle d'or

Pour toute donnée de niveau 4 ou 5, si la réponse aux questions **4** (rédaction) **6** (journalisation) ou **7** (validation humaine) t'inquiète : la décision par défaut est **lecture seule — ou rien**.

Tu n'as pas besoin de tout brancher tout de suite. Commence par la donnée publique et interne (niveaux 1-2), là où le risque est faible et le gain immédiat — la recherche dans ton propre fonds documentaire. Le reste vient après, une fois le périmètre cadré.

> ***Rappel*** *: cette grille est un outil de cadrage technique et de gouvernance, pas un avis juridique. Elle t'aide à poser les bonnes questions et à documenter tes décisions. Elle ne remplace ni ton DPO, ni ton RSSI, ni ta déontologie.*

Sources (consultées au 22/07/2026) : iManage — serveur MCP GA (communiqué du 14/05/2026) · NetDocuments × Anthropic — MCP juridique (01/04/2026) · Anthropic, centres support & confidentialité (résidence des données, restriction des actions par connecteur, périmètre du Zero Data Retention, capture des surfaces agentiques) · Spécification Model Context Protocol (bonnes pratiques de sécurité ; validation humaine) · Linux Foundation / Agentic AI Foundation (gouvernance MCP, décembre 2025) · Microsoft France × YouGov, janvier 2026 (usage via comptes personnels) · CSA Research × Lefebvre Dalloz, 2026 (usage malgré directive) · CNB, guide déontologie & IA, mars 2026 · RGPD, article 28 · CJUE, Schrems II · Règlement européen sur l'IA (AI Act), article 50, applicable le 02/08/2026.