NOTE

Annexe à la charte informatique relative à l'utilisation de services d'intelligence artificielle externes (SIAE) non fournis par l'AP-HP

Création	Version	Modification
27 Mai 2025	v4.3	1 Septembre 2025

Cette charte encadre l'usage des Services d'Intelligence Artificielle Externes (SIAE) par les personnels de l'AP-HP. Elle vise à garantir un usage responsable, sécurisé et conforme aux réglementations en vigueur, notamment le RGPD, le Code de la santé publique et la doctrine de la CNIL, en cohérence avec la Politique Générale de Sécurité de l'Information (PGSI)

Elle s'applique à tout agent, stagiaire, contractuel, prestataire ou collaborateur de l'AP-HP utilisant un outil SIAE dans le cadre de ses activités professionnelles, que ce soit sur un système AP-HP ou personnel.

1. Qu'est-ce qu'un service d'intelligence artificielle externe ?

Les Services d'Intelligence Artificielle Externes (SIAE) désignent les outils en ligne (agents conversationnels (chatbots), traducteurs, services de transcription ou de génération d'image) capables de générer du texte ou d'autres types de contenus en réponse à des requêtes ("prompts"). Bien qu'ils ne fassent pas partie des ressources informatiques internes de l'AP-HP, ces services sont facilement accessibles, qu'ils soient gratuits ou payants.

Distinction entre un modèle, un système et un service d'IA

Un **modèle** d'IA est un algorithme construit à partir de données pour réaliser une ou plusieurs tâches, comme générer du texte. Un **système** d'IA désigne quant à lui l'ensemble constitué par ce modèle, les moyens techniques et interfaces nécessaires à son utilisation opérationnelle (par exemple un client web). Enfin, les **services d'IA**, qui font l'objet de cette charte, correspondent à l'offre concrète encadrant l'utilisation d'un système, accessible en ligne par les utilisateurs, éventuellement fournie par un acteur externe. Chaque service dispose de ses propres conditions d'utilisation, modalités d'hébergement et adresses web.

C1 - Interne

Exemple de distinction

Un même modèle tel que GPT-4o™ peut être intégré dans plusieurs systèmes différents selon les usages ciblés, comme l'assistant conversationnel ChatGPT™ ou l'interface de chat de Microsoft Copilot™.

De même, un système d'IA peut être déployé par différents services. Ainsi, un système conçu par une entreprise, par exemple, *Le Chat*™ (Mistral), initialement accessible au public via *chat.mistral.fr*, pourrait être déployé sur les environnements de l'AP-HP pour devenir un service interne, accessible par exemple sous l'adresse *mistral-chat.aphp.fr* (il s'agit d'un exemple, un tel service interne n'existe pas aujourd'hui).

Services d'IA externes populaires

Parmi les SIAE les plus couramment utilisés figurent notamment ChatGPT™ (OpenAI), Le Chat™ (Mistral), Claude® (Anthropic), Gemini™ (Google), Copilot™ (Microsoft), Grok™ (xAI) pour la génération de texte, ainsi que DALL-E® d'OpenAI, Stable Diffusion™ ou Midjourney® pour la création d'images. Pour la transcription et la synthèse vocale, on retrouve notamment ElevenLabs™ et Whisper™ (OpenAI). Enfin, DeepL® et Google Translate™ comptent parmi les outils les plus répandus en traduction automatique. Cette liste n'est pas exhaustive, mais illustre les services concernés par la présente charte.

2. Dans quelles situations ces SIAE peuvent-ils servir ?

À ce jour, même si certains services sont embarqués dans des logiciels métier tels que l'analyse d'images, ou la reconnaissance vocale, l'AP-HP ne propose pas directement de service d'intelligence artificielle générale à ses agents. Cependant, **des services externes existent et peuvent être utilisés sous réserve du respect rigoureux des règles de confidentialité rappelées ci-dessous, notamment en veillant à ne jamais y introduire de données personnelles ni de données stratégiques**. Ces services d'intelligence artificielle externes (SIAE), tels que les assistants conversationnels, constituent ainsi des outils précieux pour faciliter ou accélérer certaines tâches répétitives de lecture ou de rédaction nécessitant une mise en forme claire de l'information.

Ils sont particulièrement utiles pour la correction orthographique et grammaticale, la reformulation de textes, la structuration d'un compte rendu de réunion, la traduction, la rédaction d'un courriel ou de tout autre document non sensible.

Ils peuvent également être utiles pour la génération ou l'amélioration de code informatique ou de scripts à des fins internes (toujours à valider par un expert technique), ou encore l'aide à l'élaboration d'un support de cours, d'une présentation ou d'un article scientifique dès lors qu'aucune information sensible n'est partagée avec l'outil.

Enfin, l'usage de SIAE à des fins de prototypage ou d'expérimentation, qu'il s'agisse d'un contexte clinique ou de recherche, n'est possible que dans un cadre strictement défini par le pôle innovation, la DSN, le RSSI et le DPO.

3. Quels services privilégier ?

Hébergement dans l'UE

Privilégiez les services hébergés au sein de l'Union européenne ou opérés par des prestataires européens ou français et donc soumis directement au RGPD. Ce choix garantit non seulement un encadrement strict des éventuels transferts de données hors UE (qui ne peuvent se faire sans mesures spécifiques), mais aussi le respect de principes essentiels issus du cadre juridique protecteur européen, tels que la minimisation des données

collectées, la transparence des traitements, ainsi que l'obligation renforcée de sécurité et de confidentialité.

De plus, les acteurs européens sont soumis au règlement européen sur l'intelligence artificielle (AI Act), offrant des protections supplémentaires aux utilisateurs, notamment en matière de transparence, d'explicabilité et de fiabilité des systèmes d'IA.

À l'inverse, recourir à des services situés hors UE, par exemple aux États-Unis, expose vos données et requêtes à des cadres réglementaires moins stricts, où ces garanties ne sont pas nécessairement assurées (par exemple, risque de réutilisation de vos données par des tiers sans contrôle adéquat ni protection conforme au RGPD).

Exemple de fuite de données

Entre juillet et août 2025, près de 100 000 conversations "partagées" via ChatGPT, et plus de 300 000 via Grok, ont été indexées par Google. Les liens générés n'étaient pas protégés par des balises empêchant le référencement, ce qui a rendu ces échanges accessibles au public. Bien que le partage ait nécessité une action explicite des utilisateurs, l'ambiguïté de l'interface a conduit beaucoup d'entre eux à penser que leurs échanges resteraient privés. Résultat : des contenus parfois sensibles (santé, données professionnelles, confidences) ont été rendus publics et restent, pour partie, accessibles via des caches ou des archives.

Mode d'accès aux services

De principe, pour un usage professionnel les outils nécessitant un compte nominatif doivent être évités. Des exceptions sont possibles, mais doivent être réduites au maximum. Un accès dit "anonyme" permet d'éviter d'associer vos requêtes à une identité personnelle ou professionnelle, limitant ainsi la quantité d'informations collectées. Certains services proposent un mode d'utilisation directe sans enregistrement, utile pour tester une IA sans divulguer votre identité ou votre appartenance à l'AP-HP.

Si la création d'un compte est nécessaire, par exemple pour associer des documents aux requêtes, ou pour bénéficier d'autres fonctionnalités avancées, la plateforme doit proposer un système d'authentification forte (par exemple, authentification à deux facteurs).

Pour des raisons de sécurité, il est interdit d'installer des extensions de navigateurs ou d'applications de bureau non autorisées par la DSN sur les postes de travail AP-HP.

Confidentialité et conservation des données

Lisez les conditions d'utilisation ou la documentation du service pour savoir s'il conserve l'historique des requêtes. Certains outils permettent de désactiver la réutilisation de vos conversations, afin que les contenus saisis ne soient pas utilisés pour entraîner le modèle. Il est vivement conseillé d'activer cette option ou équivalent si disponible. Notez toutefois que même dans ce mode, les données peuvent parfois être conservées temporairement par le fournisseur d'où l'importance de rester très vigilant sur la nature des informations envoyées.

Exemple de SIAE conforme

✅ En France, l'assistant **Le Chat** *de* l'entreprise Mistral ne nécessite pas de connexion préalable et offre la possibilité de désactiver la réutilisation des conversations, y compris lors de sessions sans authentification.

Signalement

En cas de doute sur la sécurité d'un service ou de fuite d'informations, le canal de signalement sécurité (signalement-cyber@aphp.fr) doit être utilisé immédiatement.

4. Données personnelles, médicales, confidentielles ou stratégiques : interdiction absolue

Protection des données personnelles et médicales

Conformément à la charte informatique de l'AP-HP et au RGPD, **il est strictement interdit d'entrer des données personnelles[1] d'un tiers dans un SIAE (ex. : nom, prénom, téléphone, mail, image), a fortiori des données de santé concernant des patients ou du personnel.** Cette interdiction s'applique même aux données prétendument "anonymisées" ou pseudonymisées. En effet, remplacer un nom par des initiales ou un identifiant ne suffit pas : le contexte ou d'autres éléments pourraient permettre de reconnaître la personne, et ces données restent des données personnelles protégées par la loi. Enfin, transmettre des informations à un service d'IA privé revient à communiquer ces informations à un tiers. Ainsi pour les données de santé, toute divulgation d'une information sur un patient à un tiers non autorisé constitue une violation du secret médical, condamnable pénalement.

Protection des données stratégiques ou relevant d'un droit de Propriété intellectuelle

Conformément aux règles de classification interne de l'information et aux obligations de confidentialité issues de la charte informatique, **il est strictement interdit de charger dans un SIAE tout document ou toute information présentant un caractère confidentiel, secret, stratégique ou relevant d'un droit de propriété intellectuelle concernant l'AP-HP.** Sont notamment concernés, sans que cette liste soit exhaustive : les documents relatifs au budget de l'AP-HP, les documents relatifs aux appels d'offre et contrats avec les fournisseurs ou partenaires industriels, les protocoles de recherche clinique, les dossiers liés à des demandes de brevets, de description de savoir-faire, à la mise au point de programmes informatiques, les éléments de stratégie institutionnelle, ainsi que toute donnée pouvant contenir des informations sensibles relatives à l'organisation des soins, à la sécurité des systèmes d'information ou aux travaux de recherche menés au sein de l'AP-HP. Toute utilisation contraire est susceptible d'engager la responsabilité de l'utilisateur.

Suivi et contrôle

La DSN en lien avec la DPO peut, conformément à la charte informatique, réaliser un audit ou une enquête sur les usages. Les règles énoncées dans cette charte s'appliquent aux utilisations effectuées depuis un poste professionnel **comme depuis un terminal personnel (téléphone, ordinateur, tablette)** utilisé par un agent dans le cadre de ses fonctions. Des sanctions disciplinaires peuvent être engagées en cas de non-respect de la charte informatique et de la présente annexe relative aux SIAE. L'Utilisateur est exclusivement et seul responsable des dommages et préjudices de toute nature, matériels ou immatériels causés à l'AP-HP ou à l'un de ses agents du fait de l'utilisation non conforme du SIAE.

L'AP-HP peut être amenée à interrompre l'accès à un SIAE à tout moment, à tout ou partie de ses agents, sans préavis, à sa discrétion et notamment pour des raisons de maintenance ou en cas de force majeure. Cette interruption ne peut en aucun cas engager la responsabilité de l'AP-HP et n'ouvre droit à aucune indemnité.

[1] https://www.cnil.fr/fr/definition/donnee-personnelle

5. Quelles sont les limites de l'utilisation des SIAE ?

Hallucinations et erreurs

L'IA peut fournir, avec une forte confiance apparente, des informations inexactes ou inventées, simplement parce qu'elles semblent statistiquement probables. Un agent conversationnel peut donc inventer une référence, ou attribuer une citation à la mauvaise personne. Ces erreurs, appelées *hallucinations*, surviennent notamment lorsque ces systèmes sont interrogés sur des sujets pointus ou récents qu'ils ne maitrisent pas, ou lorsque les prémisses de la question sont erronées.

Exemple d'hallucination

À la question « *Quelle station de métro de Paris a été renommée en 2022 pour remercier les soignants de leurs efforts durant la crise du COVID ?* », la majorité des modèles de pointe[2] hallucinent, chacun donnant une station différente. Certains affirment par exemple : « *En 2022, la station de métro Saint-Michel à Paris a été renommée temporairement en Saint-Michel – Merci aux soignants* ». Aucune station de métro n'a pourtant été renommée de la sorte en 2022.

Il faut également garder à l'esprit que ces outils sont généralistes et rarement spécialisés sur le domaine hospitalier. En particulier, ils n'ont pas été entrainés sur les données de l'AP-HP et sont donc susceptibles d'halluciner plus fréquemment sur des sujets propres à notre institution (procédures, domaines médicaux de pointe, etc.). De manière générale, aucun contenu généré ne doit être cru aveuglément, ni se substituer à l'avis d'un·e expert·e.

Biais et contenus inappropriés

Les modèles d'IA apprennent sur de larges ensembles de données pouvant inclure des biais culturels ou des informations obsolètes. Ils peuvent reproduire des stéréotypes biaisés (par exemple dans la formulation de réponses sur tel ou tel groupe de population) ou des propos inadaptés si on les pousse dans leurs retranchements. Bien que des filtres existent pour éviter les dérives, aucune IA n'est infaillible. L'utilisateur doit donc relire avec un regard critique les propositions de l'IA, notamment pour détecter d'éventuels biais ou éléments choquants, et les corriger le cas échéant.

Exemple de biais

Un modèle d'IA de détection de mélanome appris seulement sur des peaux blanches aura un comportement incertain et possiblement erroné sur des peaux non-blanches.

Boites noires

L'IA ne cite pas toujours ses sources, et son raisonnement interne est une boîte noire difficile à expliquer[3]. Elle peut délivrer une réponse plausible mais erronée, et ne pas signaler son incertitude[4]. Il incombe à l'utilisateur de vérifier chaque fait important dans la sortie de l'IA à l'aide de sources fiables. Les explications que les assistants IA peuvent apporter pour compléter leurs réponses ("explique ton raisonnement") ne sont pas fiables non plus car elles peuvent également relever d'hallucinations[5]. Enfin, le cas échéant, il est nécessaire de vérifier les sources qui accompagnent la réponse d'une IA. Celles-ci peuvent être obsolètes

[2] Test en août 2025 sur GPT 5, Gemini 2.5, Claude 4.1, Mistral et Grok 4, *tous privés d'accès à Internet*.

[3] https://www.cnil.fr/fr/les-questions-reponses-de-la-cnil-sur-lutilisation-dun-systeme-dia-generative

[4] https://aclanthology.org/2024.acl-long.198.pdf

[5] https://transformer-circuits.pub/2025/attribution-graphs/biology.html - dives-cot

ou avoir été mal comprises par le modèle. Accorder une confiance excessive sans vérification peut donc conduire à de mauvaises décisions ou conclusions.

Transparence, responsabilité et maîtrise des contenus produits

Même lorsque les résultats fournis par une IA semblent justes, appropriés et exempts de biais, **l'utilisateur doit impérativement conserver une maîtrise totale du contenu final utilisé**. Par exemple, un informaticien ne doit pas partager ou déployer du code produit par une IA sans l'avoir préalablement compris, vérifié et validé. De même, dans le domaine administratif, aucun document, courrier officiel ou compte-rendu ne doit être diffusé sans avoir été soigneusement relu et validé et compris par l'utilisateur.

Si un contenu issu d'une IA, comme un compte-rendu de réunion, devait être partagé sans vérification préalable approfondie, il est impératif de signaler explicitement qu'une IA a été utilisée pour produire ce résultat. La transparence sur l'usage de l'IA est de rigueur : toute utilisation doit pouvoir être tracée, expliquée, et justifiée, notamment dans le cas des services externes.

Impact environnemental

L'usage intensif des intelligences artificielles génératives entraîne des impacts environnementaux importants, notamment une consommation énergétique élevée, une utilisation massive d'eau nécessaire au refroidissement des centres de données, ainsi qu'une forte consommation de ressources naturelles rares (comme les métaux critiques). Par exemple, il est estimé qu'une requête vers ChatGPT consomme dix fois plus d'énergie qu'une requête à un moteur de recherche classique[6]. Dans ce contexte et conformément à l'engagement de l'AP-HP à limiter son empreinte carbone, il convient d'adopter une approche raisonnée, en privilégiant l'usage des SIAE uniquement lorsque des alternatives plus sobres ne suffisent pas, et en formulant des requêtes précises pour éviter des échanges prolongés ou superflus.

6. Comment me former à leur utilisation ?

L'AP-HP encourage ses agents à se former à l'utilisation responsable et efficace des SIAE. Un module e-learning, « *Comprendre les bases de l'intelligence artificielle* », est disponible sur la plateforme Formaphp pour initier les professionnels aux fondamentaux de l'IA. D'autres ressources pédagogiques sont également accessibles en libre accès.

La plateforme Mentor (campus numérique de la fonction publique) propose notamment des modules e-learning adaptés à tous les niveaux d'expertise, comme Découvrir les IA Génératives[7] (en partenariat avec l'INRIA) ou Objectif IA : Initiez-vous à l'intelligence artificielle[8]. La Direction Interministérielle du Numérique offre également des ressources complémentaires, telles que le MOOC d'introduction Elements of AI[9] en collaboration avec l'Université d'Helsinki, une formation spécifique pour les formateurs intitulée Intelligence artificielle pour et par les enseignants[10]. L'EHESP propose quant à elle une formation axée principalement sur l'IA Générative[11]. Enfin, le Health Data Hub met à disposition l'e-learning DATA CITY – Formation citoyenne[12] à l'intelligence artificielle, particulièrement utile pour comprendre les enjeux liés aux données de santé et à l'IA dans le contexte médical.

---

[6] https://www.lecese.fr/sites/default/files/pdf/Avis/2024/2024_14_IA_Environnement.pdf
[7] https://campus.numerique.gouv.fr/catalogue/d%C3%A9couvrir-les-ia-g%C3%A9n%C3%A9ratives/
[8] https://campus.numerique.gouv.fr/catalogue/objectif-ia-initiez-vous-%C3%A0-lintelligence-artificielle/
[9] https://course.elementsofai.com/fr/
[10] https://campus.numerique.gouv.fr/catalogue/intelligence-artificielle-pour-et-par-les-enseignants/
[11] https://campus.ehesp.fr/course/view.php?id=7694
[12] https://www.health-data-hub.fr/formations/data-city-005-formation-citoyenne-lintelligence-artificielle

7. Règles d'utilisation des SIAE

Règle 1 : Ne saisissez que des informations qui peuvent être rendues publiques. En particulier, ne divulguez aucune donnée personnelle d'un tiers, médicale, confidentielle ou stratégique lors de l'utilisation d'un SIAE. Même pseudonymisées, ces données sont interdites sur des services externes non sécurisés. Il est donc interdit d'utiliser des IA connectées à Internet à des fins d'analyse de base de données internes (médicales, financières, stratégiques, etc) sans l'approbation explicite du DPO et du RSSI. Il est également interdit d'utiliser des applications de bureau ou des extensions de navigateur non autorisées par la DSN sur des postes AP-HP. Ces règles s'appliquent aux utilisations professionnelles effectuées depuis un poste fourni par l'AP-HP comme depuis un terminal personnel (téléphone, ordinateur, tablette).

Règle 2 : Privilégiez des SIAE respectant la confidentialité : hébergement en UE, pas d'inscription nominative requise, évitez les applications douteuses ou situées hors UE et paramétrez les SIAE pour maximiser la protection (ex. désactivez l'historique des conversations). N'utilisez pas un SIAE si vous ne connaissez pas, ne comprenez pas ou n'acceptez pas ses conditions d'utilisation.

Règle 3 : Vérifiez l'exactitude des informations et les sources avant de les réutiliser ou diffuser. Les productions des IA ne doivent pas être copiées aveuglément, ni être utilisées pour prendre automatiquement de décision pour un patient. Vous restez responsable du contenu final utilisé et devez en garder la maitrise. Il convient donc de relire, de corriger les erreurs factuelles, et de s'assurer que le ton et le contenu conviennent. Si un contenu généré par une IA devait être partagé sans vérification approfondie, il est impératif de le signaler explicitement comme tel.

Règle 4 : Recourez de manière raisonnée et responsable à l'IA générative en ayant conscience des impacts environnementaux. Formulez des requêtes précises, et évitez les conversations longues. Quand cela est pertinent, privilégiez des alternatives plus sobres comme des moteurs de recherche ou des services spécialisés.

Conformité du service → Sensibilité de la requête ↓	Service européen conforme	Service non conforme[13]
Requête qui ne pourrait pas être rendue publique (en lien avec un patient, des collègues ou des sujets confidentiels)	❌ Interdiction absolue	❌ Interdiction absolue
Requête sans lien avec des patients ou des collègues et sans élément confidentiel	✅ Utilisation responsable	⚠️ Seulement s'il n'existe pas d'alternative viable

[13] Service hébergé dans l'UE, conforme au RGPD et à l'AI Act, accessible sans compte nominatif, et offrant la possibilité de désactiver la réutilisation des données.