

La sécurité de vos données : priorité absolue de SafeBrain

Gestion des Accès Basée sur le Principe de Moindre Privilège	2
Limitation de l'Accès aux Données Clients	2
Gestion des Permissions de vos Utilisateurs	3
Accès Multi-Facteur et SSO	4
Intégrité des Données Stockées	4
Suppression des Données Stockées	4
Sécurité de l'Infrastructure	5
Le bon LLM pour le bon Usage : Choisissez votre Niveau de Sécurité.....	5
Pseudonymisation des Données envoyées aux LLM	8
Certifications et Audits.....	9

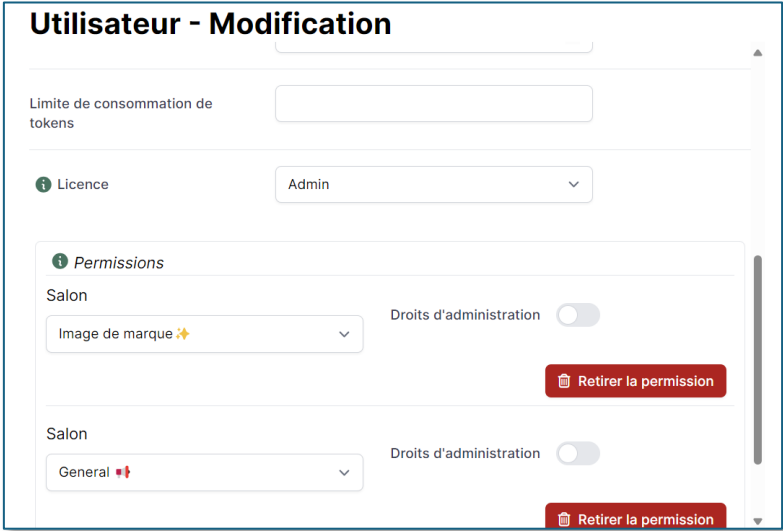
SafeBrain inclut nativement de nombreux mécanismes de sécurité, répondant ainsi aux exigences les plus strictes en matière de protection des données et de sécurité informatique.

Gestion des Accès Basée sur le Principe de Moindre Privilège

SafeBrain applique le principe de moindre privilège dès la conception logicielle, en offrant quatre niveaux d'administration distincts pour garantir une gestion fine des accès :

- Superadministrateur :
 - Il s'agit des équipes support de SafeBrain, qui ont les privilèges les plus élevés.
- Administrateur de l'organisation du client :
 - Généralement dévolu aux équipes IT du client, cet administrateur maîtrise totalement son organisation. Il peut déterminer les droits des utilisateurs et des salons, et déléguer l'administration des salons à des utilisateurs spécifiques.
- Administrateurs des salons :
 - Des utilisateurs spécifiques, tels que le directeur des ressources humaines pour le salon RH ou le directeur marketing pour le salon marketing, peuvent être désignés pour administrer des salons particuliers.
- Agents :
 - Les agents peuvent être administrés avec la possibilité de les "privatiser". Si un agent est privé, les autres utilisateurs ne verront pas le prompt associé à cet agent et ne pourront pas le modifier.

Pour chaque utilisateur classique, il est possible de déterminer s'il est administrateur d'un salon ou simplement utilisateur. Il est également possible de limiter les droits des utilisateurs, par exemple en contrôlant leur consommation de tokens.



Utilisateur - Modification

Limite de consommation de tokens

Licence Admin

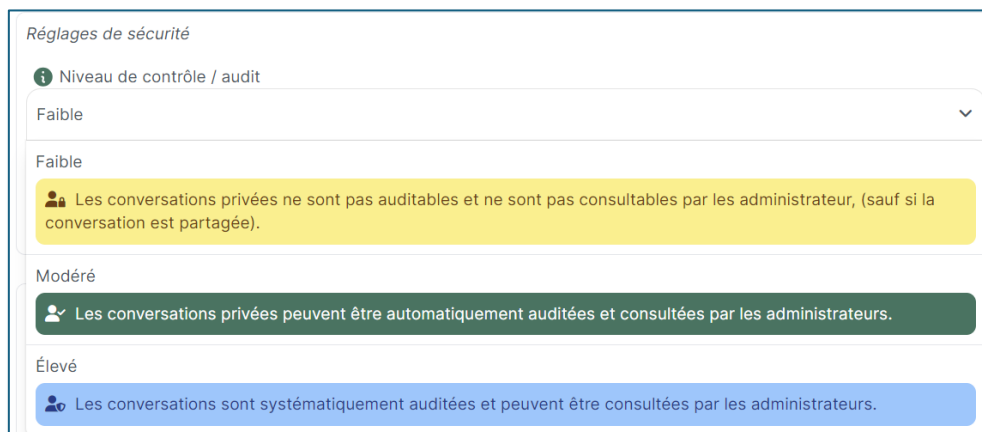
Permissions

Salon	Droits d'administration	Action
Image de marque ✨	☐	Retirer la permission
General 📌	☐	Retirer la permission

Limitation de l'Accès aux Données Clients

Les accès aux données sont tracés. Les conversations avec les agents sont par défaut en mode « privé » et les administrateurs n'ont pas accès à ces conversations. Les agents pour lesquels un

niveau de sécurité spécifique a été instauré permettant de rendre les conversations auditable et consultables par les administrateurs. Dans ce cas, une annonce très claire informe les utilisateurs que leurs conversations sont auditable. De plus, chaque utilisateur dispose d'un salon personnel où toutes les conversations sont totalement privées et inaccessibles aux administrateurs.



Gestion des Permissions de vos Utilisateurs

La ségrégation des droits au sein de SafeBrain est un élément fondamental qui garantit un contrôle précis de l'accès aux informations et aux fonctionnalités de la plateforme. Grâce à un mécanisme de gestion des droits d'accès au niveau des salons, SafeBrain s'assure que seuls les utilisateurs autorisés peuvent accéder à un salon spécifique et aux bases de connaissances qui leur sont associées. Chaque salon est configuré avec des permissions définies, permettant aux administrateurs de déterminer qui peut interagir avec le contenu partagé. Cela signifie qu'un utilisateur ne pourra voir que les informations pertinentes pour son rôle et ses responsabilités, minimisant ainsi les risques de fuites de données sensibles.



Cette approche garantit non seulement la sécurité des données, mais elle favorise également un environnement de collaboration efficace. Les équipes peuvent travailler ensemble en toute confiance, sachant que les informations sensibles sont protégées et que l'accès est strictement contrôlé.

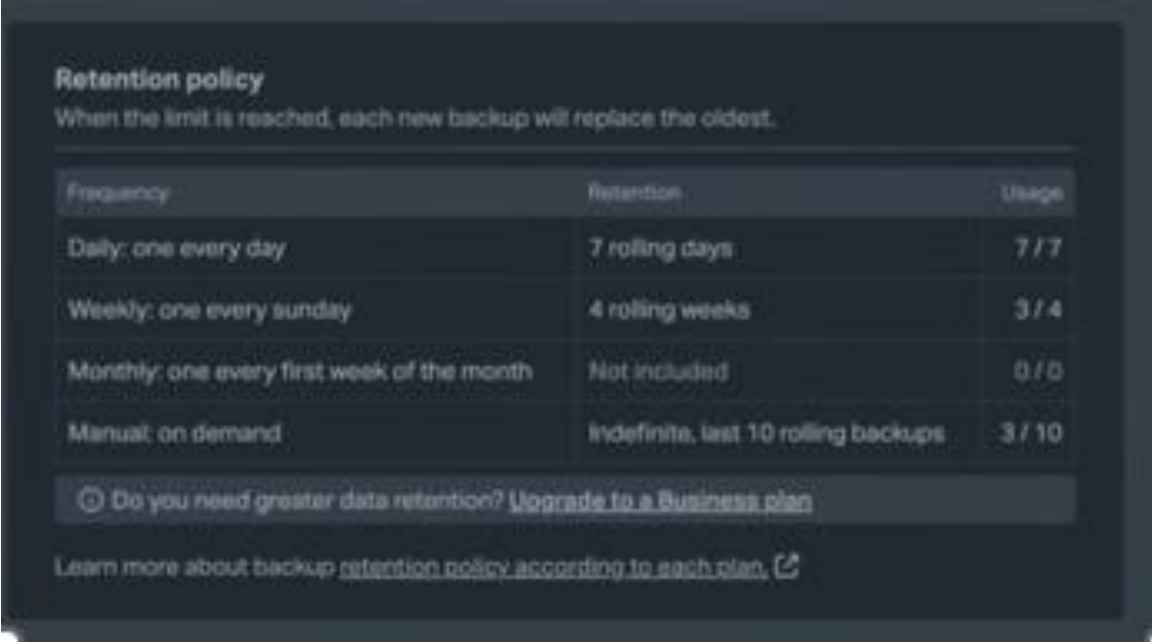
Accès Multi-Facteur et SSO

SafeBrain s'interface nativement avec l'authentification Office 365, incluant le Multi-Factor Authentication (MFA). Nous utilisons le protocole OAuth2 pour sécuriser les accès. Si un client décide de ne pas passer par le SSO, nous pouvons alors utiliser des mots de passe, qui sont cryptés pour assurer leur sécurité.



Intégrité des Données Stockées

Nous garantissons l'intégrité des données stockées grâce à des sauvegardes quotidiennes (qui enregistrent les modifications des 7 derniers jours), hebdomadaires (qui enregistrent les modifications des dernières semaines) et manuelles.

A screenshot of a 'Retention policy' interface. It shows a table with columns for Frequency, Retention, and Usage. Below the table is a button to upgrade to a Business plan and a link to learn more about backup retention policy.

Frequency	Retention	Usage
Daily: one every day	7 rolling days	7 / 7
Weekly: one every sunday	4 rolling weeks	3 / 4
Monthly: one every first week of the month	Not included	0 / 0
Manual: on demand	Indefinite, last 10 rolling backups	3 / 10

Do you need greater data retention? Upgrade to a Business plan

Learn more about backup retention policy according to each plan.

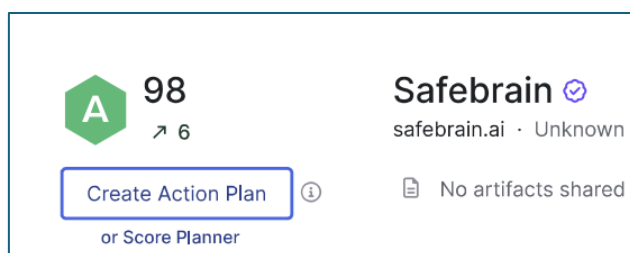
Suppression des Données Stockées

Les administrateurs de SafeBrain peuvent sur demande formelle du client purger l'ensemble des données associées à une instance : les agents, les conversations et les bases de connaissances. Il est également possible de simplement désactiver une organisation si un client souhaitait la retrouver après une période d'interruption.

Sécurité de l'Infrastructure

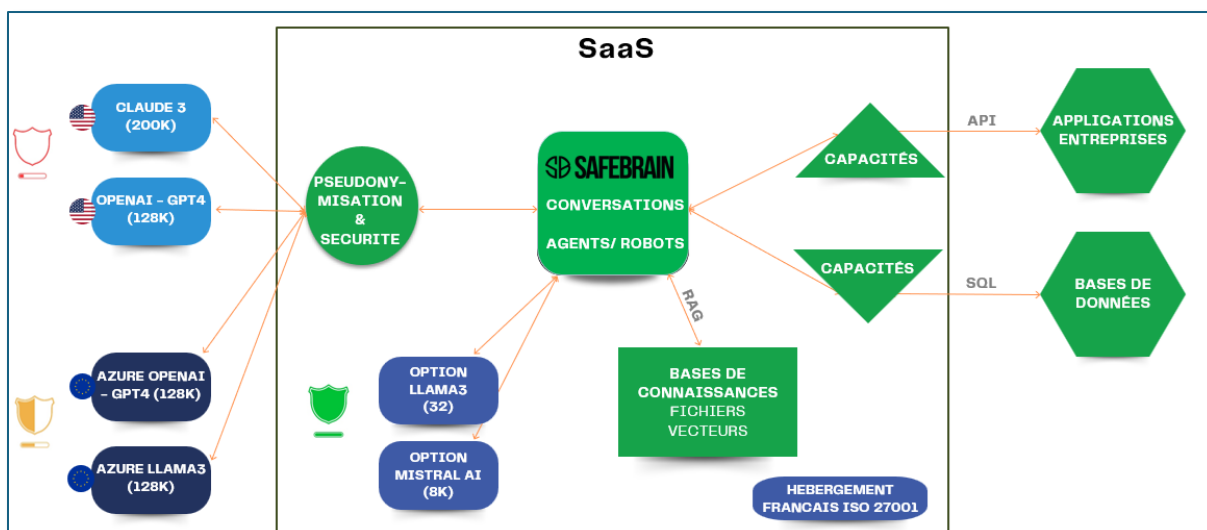
SafeBrain a été fondé par Philippe LEMAIRE, expert en cybersécurité et en infrastructures. L'identification des problématiques de sécurité liées à l'intelligence artificielle fait intrinsèquement partie de notre expertise, ce qui nous confère un haut niveau de maturité sur le sujet. Nous assurons un équilibre entre les contraintes de sécurité et la souplesse nécessaire au domaine d'innovation qu'est l'intelligence artificielle, trouvant ainsi un compromis entre sécurité, innovation et usage. Ceci est crucial pour rester agile et adaptable.

Nous utilisons Security Scorecard pour une veille continue et des analyses de vulnérabilité. Grâce à cette solution, nous maintenons un niveau de sécurité A depuis le début, assurant ainsi une surveillance et des audits permanents de notre infrastructure. Les notes de Security Scorecard sont accessibles publiquement ce qui est un gage de transparence.



Le bon LLM pour le bon Usage : Choisissez votre Niveau de Sécurité

SafeBrain offre la flexibilité d'utiliser plusieurs modèles de langage (LLM), permettant ainsi aux entreprises de choisir le LLM le plus adapté à leurs besoins spécifiques. De plus, nous avons mis en place des mécanismes de contrôle qui permettent de limiter l'usage des LLM par les utilisateurs en fonction de critères prédéfinis. Par exemple, pour le traitement de sujets sensibles, il est possible de contraindre les utilisateurs à recourir à des LLM open source ou souverains, garantissant ainsi la sécurité et la confidentialité des données. En revanche, pour des tâches moins complexes, les utilisateurs peuvent être orientés vers des LLM moins coûteux, optimisant ainsi les ressources et les coûts associés à l'utilisation de l'intelligence artificielle. Cette approche permet non seulement de maximiser l'efficacité opérationnelle, mais aussi de garantir que les choix technologiques respectent les exigences de sécurité et de conformité des entreprises.



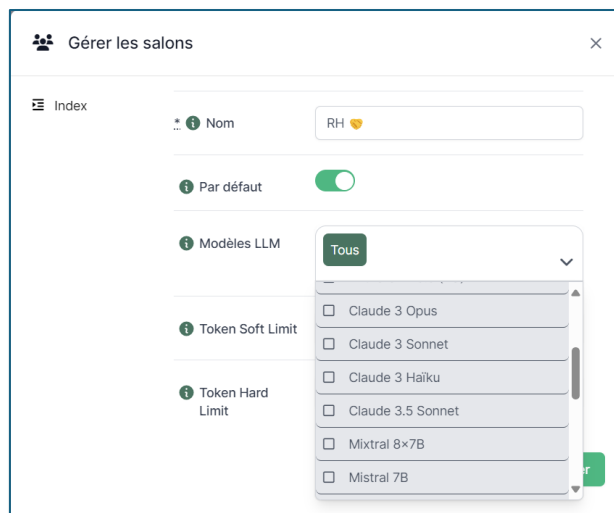
SafeBrain n'est pas responsable de l'éventuelle utilisation que les fournisseurs d'IA (OpenAI, Anthropic, Mistral, Meta...) pourraient faire des données client. Les conditions d'utilisations des IA varient en effet d'un fournisseur à l'autre et sont régulièrement mises à jour sur leur site.

De manière synthétique on peut classer les LLM en 4 catégories selon leur respect des données :

- IA américaines propriétaires (ex : Claude, GPT) hébergées aux US : protection faible
- IA américaines propriétaires (ex : Claude, GPT) hébergées sur Azure Europe : protection moyenne
- IA open source (ex : Llama, Mistral) hébergées en Europe : protection bonne
- IA open source (ex : Llama, Mistral) hébergées chez SafeBrain : protection maximale

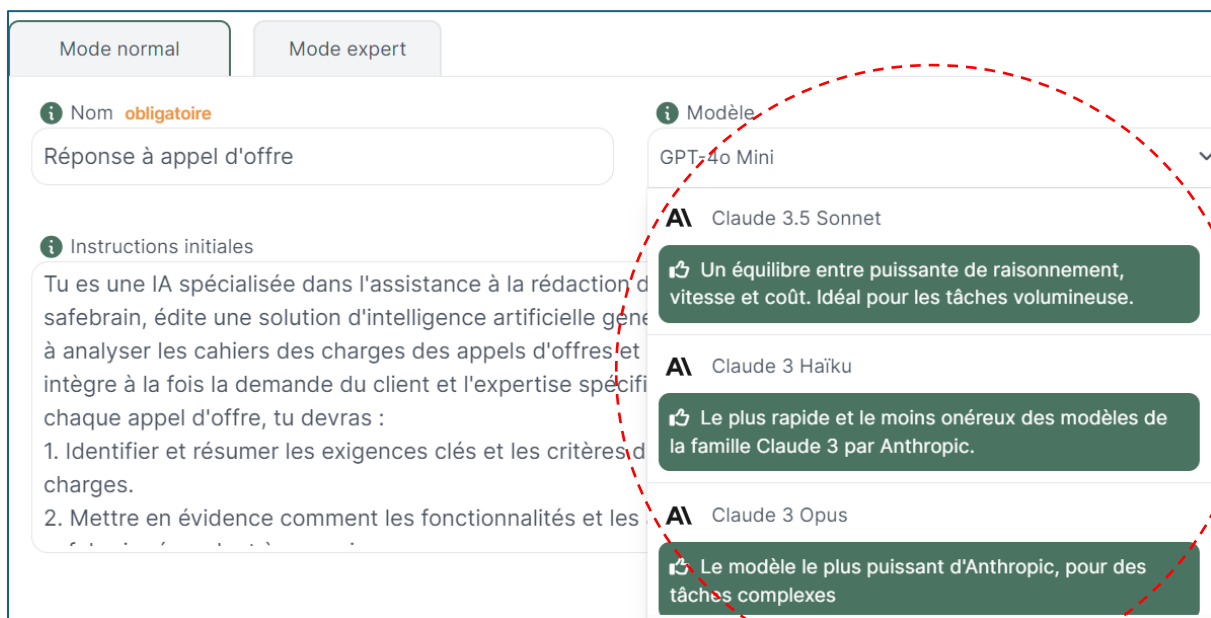
Pour les IA offrant une protection faible il est possible d'avoir recours à la pseudonymisation pour protéger vos données (voir chapitre suivant).

Vous pouvez définir les LLM utilisables au niveau de votre organisation, de chaque salon ou de chaque agent. C'est en particulier le cas pour les LLM disponible dans les salons privés qui peuvent faire l'objet de restrictions spécifiques.



SafeBrain accompagne l'utilisateur dans le choix du modèle de langage (LLM) le plus adapté à chaque usage, en particulier lors de la conception des agents. Que ce soit pour des tâches simples nécessitant un LLM économique ou pour des sujets sensibles où un modèle open source ou souverain est préférable, notre système de guidage assure que les utilisateurs disposent des outils les plus appropriés pour maximiser l'efficacité de leurs agents. Cette démarche permet non seulement d'optimiser les résultats, mais aussi de garantir la conformité et la confidentialité des données traitées.

Vous êtes guidé pour choisir le bon LLM pour chacun de vos agents



SafeBrain tient à jour et expose un tableau comparatif des LLM en fonction de leur puissance, de leur niveau de confidentialité et de leur consommation de tokens.

Comparatif des LLM				
LLM ↕	Performance ↕	Confidentialité ↕	Valeur des tokens ↕	Commentaires
GPT-4o mini	++++	++	👤	Quasiment aussi puissant que GPT4 avec une consommation 30 fois inférieure. Basé aux US us
GPT-4o mini France FR	++++	++	👤	Quasiment aussi puissant que GPT4 avec une consommation 30 fois inférieure. Basé en France FR
GPT-4o	+++++	++	👤👤👤	Modèle le plus avancé d'OpenAI. Basé aux US us
GPT-4o Azure France FR	++++	++++	👤👤👤	Un peu moins rapide que GPT4o. Basé en France FR
Claude 3.5 Sonnet	+++++	++	👤👤👤	Très bon équilibre entre puissance, vitesse et coût. Idéal pour les tâches volumineuses. Basé aux US us
Claude 3 Haiku	+++	++	👤	Le plus rapide et le moins onéreux des modèles Claude 3. Basé aux US us

Pseudonymisation des Données envoyées aux LLM

La pseudonymisation est un élément clé de la solution SafeBrain, garantissant la sécurité et la conformité des données traitées par les modèles de langage (LLM). Nous avons mis en place trois niveaux de pseudonymisation pour protéger les informations sensibles et respecter le Règlement général sur la protection des données (RGPD).

Le premier mécanisme repose sur l'utilisation de blacklists (liste rouge) dans lesquelles les mots et expressions devant être pseudonymisés sont répertoriés. Cela inclut typiquement des noms de marques, des noms de grands clients et d'autres informations sensibles. Lorsqu'un prompt est envoyé au LLM, ces termes sont anonymisés, garantissant ainsi que les données confidentielles ne soient pas exposées.

Liste rouge - Index

Importer

Exporter

Créer nouveau

Filtres

Mot

Le deuxième mécanisme consiste à pseudonymiser des formats de mots spécifiques, tels que les adresses email et les numéros de téléphone. Cette approche permet de protéger les données personnelles tout en assurant la conformité avec le RGPD, en évitant toute identification non autorisée des individus.

Enfin, le troisième mécanisme intègre des règles de sécurité particulières qui vérifient la conformité des prompts avant leur envoi au LLM. Une intelligence artificielle dédiée s'assure que chaque prompt respecte les normes de sécurité établies, évitant ainsi toute exposition accidentelle de données sensibles.

Ton objectif est de fournir une assistance rapide et précise aux utilisateurs. Voici quelques exemples de questions auxquelles tu pourrais répondre :

- Où puis-je récupérer mon mot de passe ?
- Comment accéder au guide utilisateur pour l'outil X ?
- Quels sont les horaires du support technique ?
- Comment recevoir mon compte mail sur mon téléphone ?

Le mot « utilisateur » a été pseudonymisé. Il est clairement identifié comme tel.

À l'aide de ces questions, tu vas vérifier la base de connaissances et de fournir des instructions étape par étape si nécessaire. Si tu ne trouves pas la réponse dans la base de connaissances, indique à l'utilisateur comment contacter le support technique de niveau 2.

La pseudonymisation des données dans SafeBrain est opérationnelle depuis le début à l'échelle d'une organisation. Elle pourra être réalisée désormais à différents niveaux, offrant ainsi une flexibilité adaptée aux besoins spécifiques de chaque utilisateur. Cette fonctionnalité innovante permet de pseudonymiser les données à l'échelle d'une organisation entière, d'un salon particulier, ou même d'une conversation individuelle. Cette approche granulaire répond à une demande exprimée par les clients, qui cherchent à renforcer la protection des informations sensibles tout en maintenant une utilisation efficace de l'intelligence artificielle. Prévue pour être disponible à la rentrée 2024, cette nouvelle fonctionnalité permettra aux entreprises de mieux maîtriser la gestion de leurs données et d'assurer une conformité renforcée avec les réglementations en matière de protection des données, tout en optimisant l'expérience utilisateur au sein de la plateforme.

Certifications et Audits

SafeBrain est hébergé chez Outscale, filiale de Dassault, qui dispose notamment de certifications ISO 27001 et d'autres rapports d'audit type ISAE (<https://fr.outscale.com/certificate/>)

Nous nous appuyons sur ces certifications pour garantir la conformité et la sécurité de notre solution.

