Révisée le 04/12/2024

Politique d'utilisation de l'IA générative par Teach Up

Historique des révisions	Date	Détail
1.0	01/04/2024	Première version
1.1	10/07/2024	Mise à jour
1.2	22/11/2024	Mise à jour
1.3	03/12/2024	Mise à jour

1. Introduction

Teach Up utilise des modèles de génération de texte basés sur l'IA déployés et hébergés par Microsoft Azure. Ce document technique vise à fournir une vue d'ensemble détaillée de la politique d'utilisation de l'IA générative par Teach Up, en mettant l'accent sur la collecte de données, les mesures de sécurité, l'éthique et les bonnes pratiques.

Teach Up s'engage à utiliser l'IA de manière éthique, sécurisée et responsable, en respectant les régulations de protection des données et en s'efforçant d'éviter les biais et les discriminations dans ses solutions d'IA. Les mesures de sécurité, les processus d'audit et les engagements environnementaux sont au cœur de la politique d'utilisation de l'IA de Teach Up.

2. Collecte et traitement des données

Base Légale
Teach Up n'utilise ni ne traite de données personnelles pour la conception, le développement ou la fourniture de ses solutions d'IA (Intelligence Artificielle) et de ML (Machine Learining). Toute donnée tierce utilisée est collectée légalement et est compatible avec les objectifs d'utilisation.

Utilisation des données :

- Données personnelles : aucune donnée personnelle n'est intégrée dans les solutions d'IA/ML.
- Données anonymisées : les processus de ML n'ont pas le potentiel de réidentifier des individus.
- Sources de données : les modèles sont utilisés, sans accès aux sources spécifiques de données.

3. Mesures de sécurité et de conformité

Protection des données
Pour protéger les résultats des processus de ML, Teach Up met en place des mesures techniques et organisationnelles strictes, notamment :

- Accès restreint aux personnels autorisées
- Audits de sécurité réguliers (internes et externes)
- Procédures de sauvegarde et de récupération des données robustes

Sécurité et confidentialité des données
La sécurité des données constitue une priorité absolue pour Teach Up. Dans le cadre de la fourniture de l'Expérience IA de Teach Up, un contrat avec Microsoft Azure a été établi. Ce partenariat permet l'accès à tous les modèles d'IA générative d'Open AI et à une partie des modèles de Mistral.

Les données transmises dans le cadre de l'utilisation de l'Expérience IA de Tech Up ne sont jamais utilisées pour former, pour réentraîner ou pour améliorer tous les modèles d'IA utilisés par Teach Up.

Les données fournies sont utilisées exclusivement pour fournir le service demandé, et ce, uniquement pour la durée du contrat. Il est essentiel de s'assurer que l'utilisation des services d'Azure est conforme à la politique de confidentialité de votre entreprise avant de déployer l'Expérience IA de Teach Up. L'option d'activation ou de désactivation est disponible sur tous les modules et reste à la main de l'utilisateur.

Sur Azure OpenAI Service qui embarquent tous les modèles d'IA Générative utilisés par la version en cours de Teach Up, les inputs (entrées), les outputs (sorties) et les embeddings :

- Ne sont pas accessibles à d'autres clients.
- Ne sont pas accessibles à OpenAI.
- Ne sont pas utilisés pour améliorer les modèles OpenAI.
- Ne sont pas utilisés pour améliorer les produits ou services de Microsoft ou de tiers.
- Ne sont pas utilisés pour améliorer les modèles Azure OpenAI.

Le service Azure OpenAI est entièrement contrôlé par Microsoft.

Microsoft héberge les modèles OpenAI dans l'environnement Azure de Microsoft.

Le service n'interagit PAS avec les services opérés par OpenAI (par exemple, ChatGPT ou l'API OpenAI).

Azure OpenAI Service est configuré de manière que les modèles soient stateless (Modèles sans état).

Certifications

Azure est conforme à des normes internationales de confidentialité et de sécurité, notamment la RGPD (Règlement Général sur la Protection des Données) pour les utilisateurs en Europe.

Microsoft publie des rapports d'audit et garantit que vos données ne sont pas utilisées pour ses propres besoins commerciaux.

Hébergements des modèles

Les modèles sont hébergés et déployés exclusivement sur des serveurs européens, notamment en Suède et à Paris.

Microsoft héberge les modèles OpenAI dans ses propres centres de données Azure, qui sont conformes aux normes de sécurité les plus strictes (ISO 27001, SOC 2, etc.). Azure utilise des protocoles avancés pour le chiffrement des données en transit (TLS/SSL) et au repos (AES-256).

Précision et intégrité des données

La qualité des données et la performance des modèles sont continuellement surveillées et améliorées grâce aux retours des utilisateurs et des mises à jour régulières des modèles. Teach Up s'appuie sur des techniques de génération augmentée de récupération qui permet générer des contenus à partir d'informations provenant de sources de données privées ou propriétaires.

Durée de conservation des données

Teach Up ne stocke pas de données personnelles dans ce contexte. La conservation des données est régie par le Règlement Général sur la Protection des Données (RGPD).

Évaluation des risques et réponses

Identification et gestion proactive des risques associés à l'utilisation des API Azure :

- Analyse des risques : évaluation périodique des risques de sécurité liés à l'utilisation des API.
- Plan de réponse aux incidents : protocoles en place pour réagir rapidement en cas de faille de sécurité ou de violation de données.
- Mise à jour des politiques : révision continue des politiques de sécurité pour s'adapter à l'évolution du paysage technologique.

4. Contrôle des données générées et souveraineté

Lorsqu'une décision automatisée produit un résultat négatif, une intervention humaine est possible pour générer une réponse alternative basée sur des critères différents.

À tout moment et partout dans Teach Up, un utilisateur peut soit choisir de désactiver l'IA Générative, soit modifier le contenu proposé par l'IA Générative, soit enfin consulter les informations sources qui ont servi à la création des contenus.

5. Performance et stabilité des modèles

Stabilité et performance
La performance des modèles et la santé des systèmes sont surveillées en temps réel pour assurer leur stabilité et éviter les pannes.

Audits et revues des algorithmes
Les algorithmes sont examinés et audités chaque fois qu'ils sont déployés, et en continu, pour s'assurer qu'ils restent conformes aux attentes industrielles et sociétales.

Modèles Open AI Utilisés

Teach Up utilise les modèles GPT-4, GPT-4 Turbo et GPT-4o, les modèles d'IA générative les plus avancés développés par OpenAI. Ces modèles se distinguent par leur capacité à générer du texte avec une précision et une fluidité exceptionnelle. Leurs améliorations par rapport aux versions précédentes, telles que GPT-3, incluent une meilleure gestion des contextes complexes, une réduction des biais dans les réponses, et une capacité accrue à produire des contenus pertinents et cohérents.

En intégrant ces modèles, en développant ses propres RAG et en développant des centaines de prompts uniques à haute valeur pédagogique, Teach Up est en mesure de fournir des contenus de haute qualité, adaptés aux besoins spécifiques des concepteurs et des apprenants.

6. Propriété intellectuelle des contenus générés par l'IA

Propriété des contenus générés

Dans la majorité des cas, les utilisateurs d'IA générative détiennent la propriété intellectuelle des contenus qu'ils produisent via ces outils. Cela s'applique aux textes pédagogiques, aux modules créés ou aux suggestions générées dans nos solutions.

- Garantie utilisateur : Teach Up garantit que les contenus créés à partir de nos outils appartiennent pleinement à leurs auteurs, à l'exception des cas où ces contenus reproduiraient des œuvres existantes ou enfreindraient des droits tiers (cela peut arriver par exemple si un auteur utilise dans Teach Up des contenus dont il n'est pas propriétaire et qui seraient utilisés en l'état).

- Clause explicite : aucune appropriation des contenus générés par l'IA n'est effectuée par Teach Up ou ses partenaires techniques.

Conformité des données d'entraînement : respect des droits d'auteur

Les modèles d'IA utilisés dans Teach Up sont entraînés sur des ensembles de données publics ou sous licence. Ces données sont collectées dans le strict respect des lois sur la propriété intellectuelle.

- Sources validées : OpenAI, le fournisseur de nos modèles, s'assure que les bases d'entraînement excluent les données protégées sans autorisation légale.

- Absence de rétention des données utilisateurs : les contenus fournis ou générés dans Teach Up ne sont pas utilisés pour entraîner les modèles, protégeant ainsi la confidentialité et la propriété des données des utilisateurs.

Prévention des litiges et cadre juridique évolutif

Teach Up s'engage à surveiller les évolutions réglementaires concernant la propriété intellectuelle liée à l'IA. Ce cadre évolue rapidement, notamment au niveau européen avec des initiatives comme l'AI Act.

- Audits juridiques réguliers : Teach Up s'associe à des experts pour évaluer la conformité de ses pratiques en matière de droits d'auteur et de protection des données.

- Clause de non-responsabilité : nos conditions générales précisent les responsabilités de chaque partie, limitant les risques juridiques tout en garantissant une utilisation responsable de l'IA.

Une utilisation éthique et responsable de l'IA

En limitant, en contrôlant et en optimisant l'utilisation des modèles d'IA générative dans sa solution, Teach Up offre une technologie de pointe qui respectent les principes éthiques et le cadre juridique. Nous encourageons nos utilisateurs à :

- Consulter nos politiques pour comprendre les implications des contenus générés.
- Adopter une utilisation réfléchie des outils IA pour maximiser leur potentiel tout en minimisant les risques.

7. Responsabilité environnementale

Teach Up reconnaît les impacts environnementaux des systèmes d'IA/ML et s'efforce de minimiser les émissions de CO2 générées par les besoins en serveurs et en énergie de refroidissement des centres de données, conformément aux principes de Gouvernance Environnementale, Sociale et d'Entreprise (ESG).

Pour ce faire, et pour plus de 96 % de ses requêtes, Teach Up sélectionne dans ses RAG, parmi toutes les données disponibles et pertinentes de l'utilisateur, les quatre blocs de 1000 caractères les plus pertinents pour a) limiter les flux de données envoyées et reçues dans le cadre de l'utilisation des IA Génératives, afin de limiter l'impact environnemental de l'utilisation de l'IA et pour b) améliorer la pertinence et la précision des contenus générés.

Glossaire

Accès restreint : limitation de l'accès aux données et systèmes aux seuls personnels autorisés pour protéger la sécurité des informations.

Analyse des risques : évaluation périodique des risques de sécurité liés à l'utilisation des technologies d'IA pour prévenir et gérer proactivement les menaces potentielles.

Base légale : ensemble de lois et règlements sur lesquels s'appuie une organisation pour garantir que toutes les données utilisées sont collectées et traitées légalement.

Collecte et traitement des données : processus par lequel une organisation obtient et utilise des données pour développer et fournir des solutions d'IA, tout en respectant les lois sur la protection des données.

Données anonymisées : données traitées de manière à ne pas permettre l'identification d'individus, utilisées dans les processus de Machine Learning (ML) pour garantir la confidentialité.

Données personnelles : informations relatives à une personne identifiable.

Décisions et prédictions : capacités des algorithmes à générer des contenus personnalisés et à évaluer les réponses des utilisateurs pour améliorer les expériences de formation ou d'utilisation.

Durée de conservation des données : période pendant laquelle Teach Up conserve les données, conformément à des réglementations comme le Règlement Général sur la Protection des Données (RGPD).

Évaluation des risques et réponses : processus d'identification et de gestion proactive des risques associés à l'utilisation des technologies d'IA, incluant l'analyse des risques et les plans de réponse aux incidents.

Interventions humaines : possibilité d'intervenir manuellement dans les décisions automatisées pour générer des réponses alternatives basées sur des critères différents.

Mesures de sécurité : actions mises en place pour protéger les données, incluant l'accès restreint, les audits de sécurité, et les procédures de sauvegarde et de récupération des données.