GENEVA GRADUATE INSTITUTE

INSTITUT DE HAUTES ÉTUDES INTERNATIONALES ET DU DÉVELOPPEMENT

GRADUATE INSTITUTE OF INTERNATIONAL AND DEVELOPMENT STUDIES

Directive sur l'utilisation d'un transformateur génératif pré-entrainé dans le cadre d'activités administratives

INTRODUCTION

La présente directive fait partie intégrante de la Politique de sécurité de l'information (PSI) de l'Institut de hautes études internationales et du développement (ci-après l'Institut).

Elle régit l'utilisation d'un transformateur génératif pré-entrainé dans le cadre d'activités administratives.

Pour faciliter la compréhension de lecture dans la suite de ce document, la dénomination « L'institut » est utilisée pour définir l'ensemble des entités qui peuvent être concernées par la présente directive.

La forme masculine est utilisée uniquement dans le but d'alléger le texte et désigne aussi bien les femmes que les hommes.

DESCRIPTION

Un **transformateur génératif pré-entraîné** ("Generative Pre-trained Transformer" ou GPT) est un système d'intelligence artificielle capable de générer du contenu à partir d'une requête ou d'un contexte. Il peut être utilisé pour diverses applications, telles que la rédaction de documents, la création de contenus, la synthèse d'informations, etc..

Bien qu'innovant et utile dans de nombreux domaines, un tel outil présente certains risques liés à son utilisation. En effet, si un outil de type GPT n'est pas correctement supervisé ou si les données sur lesquelles il a été entraîné contiennent des biais, il peut générer du contenu inapproprié ou offensant. De plus, l'utilisation d'un outil de type GPT gratuit ou non conforme présente des risques de sécurité des données, de non-conformité réglementaire, de cybersécurité et de diffusion de fausses informations. Par ailleurs, l'utilisation de ce type d'outil doit se faire dans le respect de la transparence et des droits de propriété intellectuelle. Il est donc essentiel de respecter des mesures de sécurité et des protocoles d'éthique stricts pour encadrer son utilisation et s'assurer qu'il soit utilisé de manière responsable et éthique.

Par ailleurs, il est utile de rappeler que les versions gratuites d'outils de type GPT n'apportent pas les mêmes garanties en matière de protection de données personnelles ou confidentielles (p.ex : informations techniques, stratégiques, économiques, financières, de sécurité ou d'affaires) que les versions payantes. Ces données peuvent, en effet, être réutilisées par les outils en question pour alimenter leurs bases de données et/ou générer des réponses et perdre ainsi leur nature confidentielle. Il convient dès lors d'être particulièrement vigilants quant aux données communiquées à ces systèmes.

ART. 1 CADRE GÉNÉRAL ET CHAMP D'APPLICATION

1.1 La présente Directive a pour objet de **donner des consignes claires** et d'**encadrer l'utilisation d'un outil de type GPT** pour les tâches administratives exécutées dans un cadre professionnel par les collaborateurs de l'Institut afin de **contrôler et de minimiser les risques** associés à son utilisation.

ART. 2 UTILISATION D'OUTILS DE TYPE GPT APPROUVÉS

2.1 Dans le cadre de leurs activités administratives et afin de répondre aux exigences réglementaires en matière de protection de la vie privée et de sécurité, les collaborateurs ne peuvent utiliser que les outils qui ont été préalablement approuvés par l'Institut dont la liste est annexée et fait partie intégrante de la présente Directive. Cette liste peut être mise à jour en tout temps. Les outils y figurant ont fait l'objet d'un examen de conformité afin de s'assurer que les informations qui leur sont transmises sont suffisamment protégées.

2.2 Les fonctionnalités de type GPT disponibles sur les outils utilisés par l'Institut ne peuvent être utilisés que s'ils apparaissent dans la liste annexée.

ART. 3 CONDITIONS ET MODALITÉS D'UTILISATION

3.1 L'utilisation d'outils de type GPT approuvés par l'Institut doit respecter les conditions et modalités suivantes :

- Les utilisateurs doivent, le cas échéant, demander un accès autorisé et sécurisé à l'outil de type GPT auprès du Service Desk de l'Institut (servicedesk@graduateinstitute.ch) ;
- Les utilisateurs doivent utiliser l'outil de type GPT uniquement pour les besoins et dans le cadre de leurs activités professionnelles au sein de l'Institut, à l'exclusion de toute utilisation à des fins personnelles ou commerciales ;

ART 4. CONSIGNES GENERALES

4.1 De manière générale, l'utilisateur d'un outil de type GPT doit en faire usage de manière responsable, éthique, sécurisée et durable.

4.1.1 Utilisation responsable de l'outil

- L'utilisation d'un outil de type GPT doit être faite de manière accessoire et ne doit pas remplacer la prise de décision humaine ni négliger l'expertise humaine et le raisonnement associé. Pour se conformer aux principes de responsabilité et de transparence et aux règles applicables en matière de propriété intellectuelle, l'utilisation d'un contenu ou d'une partie de contenu produit à l'aide d'un outil de type GPT doit être clairement mentionnée dans le document réalisé (p.ex : « Ce travail a été généré (en partie) au moyen de l'IA, [nom de la compagnie], [année], [Titre de l'outil], [Version X], [Type d'application], [URL de l'outil d'IA] »).

- Il est recommandé à l'utilisateur de prendre connaissance des conditions d'utilisation de l'outil concerné de manière à avoir une connaissance plus précise des fonctionnalités de l'outil et de pouvoir l'utiliser correctement.

4.1.2 Gestion des risques liés à l'exactitude des informations

- Les utilisateurs doivent être conscients que les réponses générées par un outil de type GPT peuvent être sujettes à des erreurs ou être incomplètes et doivent être évaluées avec soin.
- Afin d'avoir une réponse correspondant le plus possible à la question posée, l'utilisateur doit veiller à fournir des indications précises et détaillées.
- Lors de l'utilisation d'un outil de type GPT, l'utilisateur doit tenir compte du fait que les outils ne sont pas forcément actualisés régulièrement et que cela peut dès lors influencer l'exactitude et la mise à jour des réponses générées. Les réponses et les sources générées par un outil de type GPT doivent être relues, vérifiées et validées avant d'être diffusées.

4.1.3 Prévention des biais et de la discrimination

- Les utilisateurs doivent être conscients qu'un outil de type GPT peut reproduire les biais présents dans les données d'entraînement.
- Il est de la responsabilité des utilisateurs de surveiller et de corriger les réponses générées par un outil de type GPT pour éviter tout contenu biaisé, discriminatoire ou offensant.

4.1.4 Empreinte environnementale

- Un outil de type GPT a un impact environnemental élevé en termes d'émissions de gaz à effet de serre. Son utilisation doit être raisonnée afin de limiter son impact environnemental.

ART 5. SÉCURITÉ ET CONFIDENTIALITÉ DES DONNÉES PERSONNELLES ET CONFIDENTIELLES

5.1 L'utilisation d'outils de type GPT approuvés par l'Institut au moyen de leur compte professionnel garantit que l'outil bénéficie des standards en conformité avec les règlementations en vigueur.

5.2 Les utilisateurs doivent respecter la législation en vigueur, en particulier, la loi sur la protection des données (LPD) et, lorsqu'il est applicable, le règlement général sur la protection des données (RGPD) lorsqu'ils traitent des données personnelles au moyen de l'outil de type GPT, en particulier lorsqu'ils transmettent des données personnelles ou des documents en comprenant.

5.3 Les utilisateurs doivent respecter la confidentialité des données ou informations de nature confidentielle (p.ex : informations techniques, stratégiques, économiques, financières, de sécurité ou d'affaires) lorsqu'ils utilisent des outils de type GPT.

5.4 Dans la mesure du possible, aucune donnée personnelle ou confidentielle ne doit être partagée dans les interactions avec un outil de type GPT, même si l'outil en question a été approuvé par l'Institut. Il incombe à l'utilisateur d'utiliser cet outil de manière à préserver ces données et informations.

5.5 Au surplus, les utilisateurs doivent respecter la politique de sécurité de l'information de l'Institut.

ART 6. PROPRIÉTÉ INTELLECTUELLE

6.1 Un outil de type GPT se base sur des documents qui peuvent être protégés par le droit d'auteur, comme des articles, des livres, du code, des peintures, de la musique, etc. Les résultats produits par ces modèles peuvent donc également violer de tels droits, s'ils contiennent des informations qui sont identiques ou très similaires à des œuvres soumises au droit d'auteur ou si les résultats produits ne citent pas suffisamment leurs sources. Les utilisateurs doivent donc veiller à respecter les droits de propriété intellectuelle, notamment le droit d'auteur, tant quant au contenu utilisé par l'outil de type GPT qu'à son résultat, en citant les sources d'information utilisées, en évitant le plagiat et en mentionnant explicitement que le contenu ou une partie de celui-ci a été généré par le biais d'un outil de type GPT.

ART 7. SURVEILLANCE

7.1 En cas de doute quant à l'utilisation adéquate d'un outil de type GPT et sur la base de la directive 1.3 relative aux moyens de contrôles d'Internet, l'Institut a la possibilité de surveiller les échanges de manière rétroactive avec tout outil basé sur GPT pour s'assurer que les utilisateurs respectent la présente Directive et les règles applicables.

ART 8. DROIT APPLICABLE ET FOR

8.1 La présente Directive est soumise au droit suisse.

8.2 Le for exclusif est à Genève, Suisse.

ART 9. ENTRÉE EN VIGUEUR

9.1 La présente Directive est adoptée par le Comité exécutif le 23 octobre 2024. Elle entre en vigueur et est applicable immédiatement.

9.2 Elle peut être modifiée en tout temps.

GENEVA GRADUATE INSTITUTE

INSTITUT DE HAUTES ÉTUDES INTERNATIONALES ET DU DÉVELOPPEMENT

GRADUATE INSTITUTE OF INTERNATIONAL AND DEVELOPMENT STUDIES

Annexe 1

à la Directive sur l'utilisation d'un transformateur génératif pré-entrainé dans le cadre d'activités administratives

Liste des transformateurs génératif pré-entrainé (type GPT) approuvés par l'Institut pour une utilisation dans le cadre d'activités administratives au sein de l'Institut :

Nom	Editeur	Conformité
Copilot	Microsoft	Oui
DeepL	DeepL	Oui
FireFlies	FireFlies	Oui
Firefly	Adobe	Oui

Les utilisateurs peuvent demander un accès autorisé et sécurisé à un outil de type GPT auprès du Service Desk de l'Institut (servicedesk@graduateinstitute.ch)

Etat au 23.10.2024