

Directives AI Swissfundraising

Etat : Juin 2025

Le groupe de travail l'AI et le fundraising est constitué de :

Christine Bill, Janine Braun, Claudia Forcellini, Olivier Graz, Susi Kammergruber, Steff Kerkhof, Arlette Pond, Christian Schipp, Simeon Stiefel, Roger Tinner

1 Introduction: But et objectifs des directives	2
2 Definition de la terminologie	2
3 Zones/Domaines de responsabilité	3
A) Ethique	Fehler! Textmarke nicht definiert.
B) Biais et équité dans l'utilisation de l'AI	3
C) Garantie de la qualité et développement des compétences	3
D) Urheberrecht/Kennzeichnung	4
E) Protection des données	4
F) Engagement durable / Ressources	5

1 Introduction: But et objectifs des directives

Selon les directives éthiques de Swissfundraising, cinq principes de base s'appliquent à l'action dans le domaine de la collecte de fonds : Honnêteté, respect, intégrité, professionnalisme et transparence. Pour les membres de l'association, ces principes s'appliquent également à l'utilisation de l'intelligence artificielle (IA) dans la collecte de fonds.

Ces lignes directrices offrent une orientation pour l'utilisation responsable de l'IA par les collecteurs de fonds, en particulier au sein d'une organisation d'utilité publique. Pour les membres de Swissfundraising, elles font partie intégrante de l'adhésion. Elles visent à utiliser efficacement le potentiel de l'IA pour la collecte de fonds, sans négliger les aspects éthiques, juridiques ou écologiques. La législation et la jurisprudence actuelles s'appliquent toujours de manière supérieure, par exemple en ce qui concerne la protection des données ou les droits d'auteur.

Les lignes directrices définissent un cadre clair pour la protection des données, l'équité, l'assurance qualité et la durabilité. Elles favorisent ainsi la transparence et la confiance auprès des collaborateurs de l'organisation et des parties prenantes externes. Les lignes directrices aident les collecteurs de fonds et les organisations à concilier les avantages de l'IA avec leurs propres valeurs et principes et à assumer leur responsabilité éthique. Les lignes directrices créent donc une base qui permet à la fois l'innovation et la responsabilité.

Ce document fait la distinction entre les exigences qui doivent absolument être remplies (« doit »), celles dont la mise en œuvre est souhaitable du point de vue de Swissfundraising (« doit ») et celles qui doivent être considérées comme des recommandations (« peut »). Les organisations sont libres de décider de règles plus étendues pour l'utilisation de l'IA.

2 Définition de la terminologie

L'intelligence artificielle (IA) est une technologie qui apprend aux ordinateurs à apprendre à partir de données et à effectuer des tâches qui nécessitent normalement une réflexion humaine, comme comprendre le langage ou reconnaître des modèles.

L'IA générative crée de nouveaux contenus tels que des textes, des images ou des vidéos à partir de données d'apprentissage. Des outils connus à cet effet sont par exemple ChatGPT, Midjourney ou Runway ML.

Le Machine Learning ou apprentissage automatique (ML) est un domaine de l'IA dans lequel des algorithmes sont utilisés pour reconnaître des modèles et faire des prédictions au sein d'un ensemble de données - sans être explicitement programmés à cet effet. Les données peuvent être des chiffres, du texte ou des photos.

LLM (Large Language Model) est un modèle d'IA entraîné avec de grandes quantités de données (texte, images, audio, vidéo, etc.) afin de générer des sorties qui pourraient également provenir d'un être humain ou de comprendre des entrées vocales et d'y réagir.

Les biais sont des distorsions ou des préjugés dans les données ou les modèles d'IA. Ils peuvent conduire à des résultats injustes ou discriminatoires s'ils ne sont pas pris en compte.

Par **hallucinations**, on entend des résultats d'une IA formulés de manière convaincante, mais qui sont objectivement faux ou non justifiés par des données d'entraînement.

3 Zones/Domaines de responsabilité

A) Éthique

L'utilisation de l'IA dans la collecte de fonds nécessite une approche éthique qui garantit que la technologie est utilisée en accord avec les valeurs et les principes de l'organisation. **L'IA doit servir d'outil de soutien aux personnes, tandis que les décisions importantes restent toujours de la responsabilité humaine.** La protection de la dignité humaine, le respect des droits individuels et la prévention de la discrimination doivent être garantis.

Il est essentiel que les dispositifs d'IA ne favorisent pas les préjugés ou les hypothèses injustifiées. **Les algorithmes ne peuvent être aussi neutres que les données avec lesquelles ils ont été entraînés.** C'est pourquoi tant les modèles utilisés que leurs résultats doivent être régulièrement contrôlés afin d'éviter des déformations contraires à l'éthique. Il est essentiel que les collaborateurs des organisations soient sensibilisés aux questions éthiques et encouragés à remettre en question les éventuels risques ou les applications problématiques et à vérifier les résultats d'une IA.

Un autre aspect important est la transparence : **les destinataires devraient être informés de façon appropriée sur la raison et la façon dont les systèmes d'IA ont pris des décisions ou généré des contenus.** Enfin, l'utilisation de l'IA ne doit pas seulement respecter les dispositions légales, mais aussi les normes les plus élevées de respect, d'intégrité, de professionnalisme et de transparence. Cela garantit non seulement la qualité des résultats, mais renforce également la confiance dans l'organisation et son usage de l'IA.

B) Biais et équité dans l'utilisation de l'AI

Une utilisation responsable de l'IA nécessite une réflexion systématique sur les biais et l'équité, car des déformations dans les données ou les algorithmes peuvent conduire à des résultats discriminatoires. Pour minimiser ces risques, une analyse critique des données d'entraînement et des modèles sous-jacents est nécessaire. Les ensembles de données biaisés ou incomplets peuvent fausser les résultats et doivent donc être identifiés et corrigés à un stade précoce.

L'organisation doit établir des lignes directrices claires pour s'assurer que les modèles d'IA fonctionnent de manière aussi neutre et inclusive que possible. Cela peut être réalisé en testant des ensembles de données diversifiés, en vérifiant régulièrement que les résultats ne contiennent pas de stéréotypes et en évaluant continuellement les algorithmes. Les collaborateurs devraient être habilités à reconnaître les hallucinations et les déformations potentielles et à développer des stratégies pour les réduire (voir également le chapitre C).

La transparence et la responsabilité dans les processus de développement et d'utilisation de l'IA sont essentielles pour obtenir des résultats équitables et non discriminatoires. La mise en place d'un mécanisme de retour d'information et une sensibilisation continue aux biais contribuent à identifier les points faibles des modèles et à apporter des améliorations continues. Cela permet de garantir que l'IA n'est pas seulement utilisée de manière efficace, mais aussi de manière éthique et socialement acceptable.

C) Garantie de la qualité et développement des compétences

Afin d'éviter la diffusion d'informations erronées ou de distorsions, il est essentiel de procéder à un contrôle systématique de la qualité des contenus générés par l'IA. **Avant toute utilisation interne ou externe, l'exactitude du contenu et le respect des principes d'équité et de diversité doivent être vérifiés pour tous les contenus créés.** Le principe

du double contrôle est particulièrement recommandé pour les contenus sensibles ou d'importance publique. Un processus de vérification clairement défini devrait garantir la cohérence et la fiabilité des résultats.

Parallèlement, le développement des compétences des collaborateurs est un facteur central pour une utilisation responsable de l'IA. Outre les bases techniques, les formations devraient également porter sur des sujets tels que les risques éthiques, la prévention des biais (voir chapitre B) et les méthodes efficaces comme l'ingénierie d'invite. **L'objectif est que les collaborateurs évaluent de manière critique les contenus générés par l'IA, reconnaissent leurs forces et leurs faiblesses et puissent les utiliser de manière ciblée pour leurs activités de collecte de fonds.** La formation continue et l'intégration de processus d'assurance qualité permettent non seulement d'améliorer la qualité du travail et l'efficacité de la collecte de fonds, mais aussi de renforcer la confiance dans l'utilisation de l'IA au sein de l'organisation.

D) Droits d'auteur/Identification

Lors de l'utilisation de l'IA, il est essentiel de respecter les droits d'auteur, car de nombreux systèmes d'IA ont été et sont formés avec des données qui peuvent être protégées par des droits d'auteur. **Les collaborateurs ne devraient pas télécharger de contenus protégés dans les outils d'IA, à moins qu'une autorisation explicite (licence) n'ait été accordée.** De même, il est indispensable de vérifier soigneusement les contenus générés afin de s'assurer qu'aucun droit d'auteur n'est violé ou que des œuvres protégées ne sont pas reproduites illégalement.

Les contenus générés par l'IA, tels que les textes, les images ou les vidéos, doivent être identifiés de manière claire et précise afin de garantir la transparence. Une formulation possible est la suivante : « Ce contenu a été créé avec l'aide de l'IA ». Pour les médias visuels ou audiovisuels, il est nécessaire d'indiquer précisément les outils ou les plateformes utilisés, par exemple « Image IA : DALL-E/nom de l'organisation ». Les textes qui ont simplement été adaptés, mais qui n'ont pas été modifiés de manière substantielle, ne nécessitent généralement pas d'être référencés.¹

Des règles claires concernant les citations et le respect des droits d'auteur permettent de renforcer la confiance dans l'organisation et de minimiser les risques juridiques. Il est donc essentiel que les collaborateurs soient informés de ces règles et aidés à utiliser l'IA de manière sûre et conforme à la loi. Des vérifications régulières et des formations sur ces sujets permettent de garantir une pratique uniforme et responsable.

E) Protection des données

La protection des données à caractère personnel est un aspect central de l'utilisation de l'IA et doit être une priorité absolue. **Afin de garantir la protection de la vie privée, seules des données anonymes peuvent être traitées dans des systèmes ouverts, et l'exploitation d'un système fermé doit être rendue transparente en conséquence dans la déclaration de protection des données.** Il est indispensable que les lois sur la protection des données et les directives internes à l'organisation soient systématiquement respectées.

Lors de l'utilisation d'outils d'IA, il faut veiller à ce que les données saisies ne soient pas enregistrées ou réutilisées pour l'entraînement des modèles. De nombreuses plateformes proposent des paramètres correspondants qui doivent garantir que les données ne parviennent pas dans les bases de données des fournisseurs - même si, ici comme

¹ L'obligation référencer s'applique à tous les contenus concernés par l'article 50 de l'EU AI Act : <https://artificialintelligenceact.eu/de/article/50/>

partout ailleurs, cela ne constitue pas une protection à 100 % contre l'utilisation abusive des données par les fournisseurs. L'utilisation de versions gratuites ou non sécurisées d'outils d'IA est à éviter, car elles n'offrent souvent pas la protection nécessaire.

En outre, les collaborateurs doivent être sensibilisés et formés pour identifier et minimiser les risques liés à l'utilisation de l'IA. Le respect du principe de réduction des données - n'utiliser que les données nécessaires - est essentiel. **Grâce à des directives claires et à un contrôle continu des outils et processus utilisés, l'organisation peut garantir que les données personnelles sont entièrement protégées et que l'IA est utilisée en toute confiance.**

F) Engagement durable / Ressources

L'utilisation de l'IA devrait toujours se faire en tenant compte des aspects écologiques, car les systèmes tels que les LLM (Large Language Models) consomment une quantité importante d'énergie (principalement lors de l'entraînement). Il est recommandé de privilégier, dans la mesure du possible, les outils peu consommateurs d'énergie ou ceux qui fonctionnent avec des énergies renouvelables, comme les moteurs de recherche conventionnels. Dans ce contexte, les bénéfices de l'utilisation de l'IA devraient toujours être mis en balance avec la consommation de ressources.