

Liste de contrôle pour une intelligence artificielle responsable

Cette liste de contrôle vise à aider les développeurs et les responsables du déploiement de systèmes d'intelligence artificielle (IA), y compris ceux qui déploient des systèmes d'IA acquis auprès d'un tiers, à examiner leurs activités liées à l'IA afin de déterminer si elles respectent les bonnes pratiques applicables.

Remarque : les déployeurs de systèmes d'IA tiers restent responsables des systèmes d'IA qu'ils déploient. Les bonnes pratiques décrites dans ce document s'appliquent également aux systèmes d'IA tiers que votre organisation acquiert et déploie, et doivent être prises en compte dans la mesure où votre organisation a la capacité de s'y conformer.

Table des matières

Gouvernance de l'IA	2
Leadership en matière IA.....	3
Éthique des données	3
Éthique dès la conception.....	4
Gouvernance des données.....	5
Entrées et sorties de données.....	6
Achats.....	6
Transparence et explicabilité	7
Interprétabilité	9
Évaluation d'impact.....	9
Supervision et intervention humaines	10
Formation et sensibilisation	11
Suivi et contrôle.....	12
Tests	13
Détection des incidents et réponse	14
Ressources	14

Gouvernance de l'IA



Intégrez des pratiques responsables en matière d'IA dans la stratégie de responsabilité numérique de l'organisation, en mettant l'accent sur la prise de responsabilité vis-à-vis des systèmes d'intelligence artificielle (IA) développés ou utilisés par l'entreprise



Déterminez qui assume la responsabilité finale et rend compte des décisions de l'entreprise en matière d'IA



Définissez et mettez en œuvre des politiques et des lignes directrices internes concernant l'utilisation et la gouvernance des systèmes d'IA au sein de l'organisation :

- Inspirez-vous des Principes de l'OCDE relatifs à l'IA¹ (ainsi que d'autres principes et normes faisant autorité en matière d'IA) :
 - Croissance inclusive, développement durable et bien-être
 - Valeurs centrées sur l'humain et équité
 - Transparence et explicabilité
 - Robustesse, sécurité et sûreté



Élaborez et appliquez des processus et des règles standardisés pour développer, tester, déployer et exploiter des systèmes d'IA (par exemple, codage, documentation, tests, directives architecturales) :

- Examinez les recommandations des autorités de régulation compétentes, des défenseurs des droits et des leaders du secteur



Réalisez des audits internes pour évaluer la conformité des pratiques de l'organisation en matière d'IA aux processus standardisés



Désignez des personnes chargées de tenir l'entreprise informée des évolutions réglementaires et techniques :

- Réexaminez les politiques et procédures internes afin de s'assurer qu'elles tiennent compte de toute évolution du cadre réglementaire en matière d'IA, notamment :
 - les politiques de confidentialité
 - les politiques d'utilisation des données
 - Politiques de classification et de gestion de l'information
 - Conditions d'utilisation
- Veillez à ce que les employés sachent à qui s'adresser en cas de questions, de demandes de renseignements et/ou de préoccupations concernant les politiques de l'entreprise en matière de confidentialité et de protection des données relatives à l'IA.

¹ Principes et recommandations de l'OCDE en matière d'IA à l'intention des décideurs politiques <https://oecd.ai/en/ai-principles>

Direction en matière IA

- ☐ Désignez un membre de la haute direction chargé de la gouvernance du développement et de l'utilisation des systèmes d'IA au sein de l'entreprise, y compris de l'application des politiques et procédures internes
- ☐ Confiez au responsable de l'IA la mission de collaborer avec le comité d'éthique des données de l'entreprise (ou son équivalent)
- ☐ Définissez clairement tous les rôles et responsabilités organisationnels liés à la conception, au développement, à l'utilisation et au déploiement des systèmes d'IA afin de garantir que ces rôles et responsabilités s'inscrivent dans le périmètre de responsabilité du responsable de l'IA

Éthique des données

- ☐ Distinguez les responsabilités éthiques de l'entreprise de ses responsabilités juridiques :
 - Une pratique peut permettre à l'entreprise de se conformer légalement à une obligation réglementaire (par exemple, l'obtention du consentement), mais cela ne rend pas automatiquement cette pratique, ni la manière dont elle est mise en œuvre, éthique
- ☐ Élaborez un ensemble de principes éthiques en matière d'IA ainsi qu'un code de conduite interne ou un code de conduite de l'entreprise qui rende les développeurs et les utilisateurs de systèmes d'IA responsables du respect des principes éthiques en matière d'IA :
 - Inspirez-vous des principes de l'OCDE en matière d'IA² (et d'autres principes faisant autorité)
 - Intégrez les principes relatifs à l'IA dans le code de conduite/code de conduite d'entreprise et inversement
 - Dans un souci de transparence, élaborez des lignes directrices éthiques internes qui s'alignent sur les principes éthiques de l'entreprise en matière d'IA, de manière à ce qu'elles puissent être facilement rendues publiques en cas de besoin
- ☐ Mettez en place un comité d'éthique des données (ou équivalent) chargé de conseiller l'entreprise sur le développement et l'utilisation éthiques des systèmes d'IA :
 - Ce comité doit exercer un rôle purement consultatif, mais ses avis ne doivent pas pouvoir être facilement écartés par les services opérationnels ou la direction
- ☐ Composez le Comité d'éthique des données à partir de personnes et de fonctions issues de tous les niveaux de l'entreprise :
 - Le comité doit être composé de cadres supérieurs, de membres de la direction générale ou du conseil d'administration
- ☐ Intégrez au comité d'éthique des données des personnes issues de milieux divers et ayant des opinions divergentes

² Principes et recommandations de l'OCDE en matière d'IA à l'intention des décideurs politiques <https://oecd.ai/en/ai-principles>

- ☐ Offrez un « espace sûr » aux membres du comité pour qu'ils puissent remettre en question les pratiques de l'entreprise et formuler des recommandations à son égard
- ☐ Veillez à ce que les questions et les sujets abordés accessibles à tous les membres du comité (par exemple, ceux qui n'ont pas de formation juridique ou opérationnelle)
- ☐ Distinguez la responsabilité éthique de la responsabilité juridique :
 - Une pratique peut permettre à l'entreprise de se conformer légalement à une obligation réglementaire (par exemple, l'obtention d'un consentement), mais cela ne rend pas automatiquement cette pratique, ou la manière dont elle est mise en œuvre, éthique

L'éthique au cœur de la conception

- ☐ Identifiez les principales parties prenantes qui doivent être impliquées dans la conception, le développement et la mise en œuvre de solutions d'IA responsables
- ☐ Définissez ce que signifie la « justice » pour le groupe de parties prenantes. Cela peut se faire avec l'aide du comité d'éthique
- ☐ Définissez et articulez clairement la valeur commerciale ou le contexte d'utilisation commerciale d'un système d'IA :
 - Commencez par formuler le problème (c'est-à-dire le problème que l'entreprise cherche à résoudre) :
 - Évitez de partir d'une solution ou d'une capacité dont dispose déjà l'entreprise, car cela risquerait d'imposer une solution d'IA alors qu'elle n'est pas justifiée
 - Consignez les hypothèses sous-jacentes concernant le fonctionnement du système d'IA afin que ces hypothèses puissent être validées par un suivi et des tests
- ☐ Définissez les termes, les processus et les ensembles de données nécessaires à la mise au point de la solution :
 - Si des technologies de renforcement de la confidentialité (PET) sont utilisées, intégrez-les tout au long du cycle de vie du développement plutôt que d'essayer de les intégrer à la fin.
- ☐ Testez le système d'IA dans le cadre d'un projet pilote avant de le déployer en production afin de vous assurer qu'il fonctionne comme prévu
- ☐ Déterminez si l'entreprise applique les bonnes pratiques et quelles sont les conséquences de certaines pratiques
- ☐ Collaborez avec le comité d'éthique de l'entreprise pour vous assurer que les parties prenantes et le comité d'éthique utilisent un langage et une terminologie cohérents, et qu'ils ne travaillent pas à contre-courant les uns des autres pour développer et produire des algorithmes éthiques
- ☐ Mettez en place des politiques et des pratiques organisationnelles visant à favoriser la pensée critique et une approche privilégiant la sécurité dès la conception



Tenez compte des risques liés à la saisie d'informations sensibles dans des modèles d'IA :

- Les modèles d'IA générative ne doivent pas être alimentés ni incités à produire des informations sensibles ou confidentielles, à moins que les données ne soient traitées localement et/ou soumises à des contrôles d'accès et d'utilisation appropriés
- Une attention particulière doit être accordée aux points suivants :
 - Données relatives aux enfants
 - Données relatives aux employés
 - Données relatives à l'éducation
 - Données relatives à la santé des patients
 - Données relatives au recrutement ou au lieu de travail susceptibles de donner lieu à des plaintes pour discrimination ou harcèlement
 - Autres types de données réglementées



Mettez en place des mécanismes et des processus permettant de démontrer que le système d'IA déployé est valide et fiable, par exemple en ce qui concerne :

- La collecte des données
- La licéité du traitement
- La limitation du traitement à la finalité spécifiée



Mettez en place des mécanismes permettant au personnel chargé de la conception ou du déploiement des systèmes d'IA d'intégrer régulièrement les retours d'expérience du comité d'éthique et des acteurs concernés par l'IA (par exemple, les développeurs de logiciels, les utilisateurs finaux, les opérateurs, les data scientists, les services juridiques et de conformité, etc.) dans la conception et la mise en œuvre du système



Mettez en place des procédures d'urgence pour gérer les défaillances ou les incidents liés à des données tierces ou à des systèmes d'IA considérés comme présentant un risque élevé :

- Veillez à ce que le système d'IA puisse présenter des défaillances en toute sécurité, par exemple lorsqu'il est amené à fonctionner au-delà de ses limites de connaissance, et à ce que les risques négatifs résiduels ne dépassent pas le seuil de tolérance au risque de l'entreprise

Gouvernance des données



Mettez en œuvre des politiques et procédures de gouvernance des données qui définissent les normes régissant la manière dont l'organisation collectera, utilisera, stockera, conservera et diffusera les données à caractère personnel liées à l'utilisation des systèmes d'IA³



Documentez et tenez à jour des registres de provenance des données qui précisent les éléments suivants :

- Les catégories de données collectées (à caractère personnel et non personnel)
- L'origine/la source des données
- La finalité initiale de la collecte
- Des explications claires sur les données utilisées, la manière dont elles sont collectées et les raisons de cette collecte
- Qui a étiqueté les données et si des tests de biais ont été effectués pour évaluer si les données étiquetées présentaient un biais
- Comment les données sont transformées au fil du temps
- Date de la dernière mise à jour ou modification (cela peut inclure la balise de date dans les métadonnées)

³ Voir l'exemple de politique de gouvernance des données dans la catégorie « Gestion de la vie privée » 1 - Intégrer la protection de la vie privée dans le programme d'éthique et de gestion des données



Mettez en place des mécanismes permettant aux personnes de contester l'exactitude des données à caractère personnel utilisées par les systèmes d'IA déployés par l'organisation :

- Rectifiez toute donnée à caractère personnel inexacte dès réception d'informations suffisantes confirmant l'identité de la personne concernée, conformément aux procédures de l'organisation relatives au traitement des demandes liées à la protection de la vie privée et de toutes les lois applicables
- Veillez à ce que ces mécanismes soient simples et faciles à utiliser, et les présenter de manière claire et bien visible

Entrée et sortie des données



Limitez la saisie des données à caractère personnel utilisées dans le système d'IA :

- Révissez et mettez à jour les politiques et procédures existantes en matière de minimisation des données afin de limiter la saisie des données utilisées dans les systèmes d'IA
- Déterminez et documentez les techniques utilisées pour garantir que les données utilisées comme entrées du système d'IA soient limitées, dans la mesure du possible, à ce qui est nécessaire et pertinent pour atteindre l'objectif pour lequel le système d'IA a été déployé. Ces techniques peuvent inclure :
 - Des examens périodiques de la quantité de données utilisées et de la nature des données d'entrée
 - La suppression des données qui ne sont plus nécessaires ni pertinentes
 - La limitation des types de données pouvant être utilisées ou soumises (par exemple, via les paramètres et configurations disponibles dans le système d'IA)



Limitez l'utilisation des résultats du système d'IA aux finalités pour lesquelles ce dernier est destiné, si ces résultats sont susceptibles d'aboutir à des décisions entraînant des effets juridiques défavorables ou des conséquences tout aussi importantes, y compris tout risque pour les droits et libertés des personnes

- Révissez et mettez à jour les politiques existantes relatives à la limitation de la finalité afin de restreindre l'utilisation des résultats du système d'IA
- Élaborez et documentez des politiques décrivant les utilisations prévues du système d'IA

Achats

Système d'IA tiers acquis en vue de son déploiement au sein de votre organisation :



Élaborez des critères (en collaboration avec les équipes chargées des achats et de la gestion des risques d'entreprise) pour évaluer et approuver les logiciels et services tiers, nouveaux ou mis à jour, qui s'intègrent à des API d'IA ou offrent des fonctionnalités d'IA :

- Demandez à des évaluateurs internes d'examiner les ensembles de données utilisés pour créer les résultats (par exemple, pour les modèles d'IA générative) afin de déterminer les risques spécifiques associés à chaque logiciel et service
- Examinez les résultats en termes d'explicabilité, d'équité et de partialité : les résultats sont-ils cohérents par rapport aux données d'entrée, sont-ils équitables et non discriminatoires ?
- Évaluez les conditions d'utilisation et la documentation du système du fournisseur tiers ; sont-elles suffisamment transparentes concernant les fonctionnalités d'IA ?



Efforcez-vous de comprendre comment le système d'IA tiers est conçu, notamment :

- Les données utilisées pour entraîner le système d'IA
- La manière dont l'efficacité et l'explicabilité du système d'IA sont évaluées par le développeur
- Dans quelles circonstances le système d'IA présente-t-il des performances insuffisantes ?



Veillez à ce que les contrats conclus avec des fournisseurs tiers de systèmes d'IA abordent les points suivants :

- Une description et/ou des instructions concernant le système d'IA prévu
- Toutes les données obtenues en lien avec le modèle acquis
- Les mécanismes permettant à chaque partie de signaler à l'autre les vulnérabilités, risques ou biais potentiels qui apparaissent dans le système d'IA pendant la durée du contrat d'acquisition

Recours à un prestataire de services tiers utilisant un système d'IA pour fournir des services :



Définissez des exigences relatives au partage de données avec les fournisseurs afin de garantir le respect des lois, réglementations et bonnes pratiques applicables en matière de partage ou de vente de données à des tiers :

- Réduisez au minimum le partage de données à caractère personnel dans le cadre des ensembles de données utilisés pour entraîner les systèmes d'IA de tiers



Réviser les clauses contractuelles afin de s'assurer que toute utilisation des données par les fournisseurs respecte les dispositions contractuelles réglementaires obligatoires ou relève d'exceptions approuvées :

- Conformément à la législation locale, veillez à ce que les notifications et consentements appropriés soient obtenus avant tout partage de données à caractère personnel avec les fournisseurs
- Assurez-vous que les fournisseurs sont en mesure de se conformer aux demandes individuelles concernant l'utilisation de leurs données à caractère personnel dans les systèmes utilisant l'IA (par exemple, les demandes « Ne Pas Vendre » en vertu des lois des États Unies, les demandes d'accès et de suppression)

Transparence et explicabilité



Examinez et documentez les risques liés à la transparence et à l'explicabilité :

- Déterminez comment l'entreprise doit garantir la transparence vis-à-vis du public ou des personnes concernées quant à son utilisation de l'IA
- Déterminez quand et comment les employés sont tenus de divulguer si un produit de travail interne et/ou externe a été créé en tout ou en partie par des outils d'IA, en particulier des outils d'IA générative



Informez les personnes concernées du moment et de la manière dont elles interagissent avec l'IA :

- Informez les personnes qu'elles interagissent avec un système d'IA au moment même de l'interaction, par exemple au moyen d'une déclaration de confidentialité ou d'autres mécanismes (par exemple, des étiquettes, des mentions légales)
- Fournissez des informations de manière claire et concise
- Rendez ces informations facilement accessibles
- Indiquez clairement aux personnes lorsqu'elles interagissent avec une machine (bot) ou avec du contenu généré automatiquement



Informez les personnes qui interagissent avec le système d'IA de leurs droits concernant les résultats et les conséquences de ce système



Documentez et mettez à la disposition des personnes/communautés concernées et du public, le cas échéant, des explications sur le fonctionnement du modèle d'IA, notamment :

- Les mécanismes sous-jacents au fonctionnement du système d'IA
- Les limites des connaissances du système d'IA

- Les données collectées et/ou utilisées par le système
- Justification de tous les résultats générés par le système
- Comment les résultats du système peuvent être utilisés ; toutes les décisions prises et les paramètres permettant de parvenir à ces décisions
- Comment le système produit-il différents résultats ?
- L'implication d'intervenants humains (par exemple, comment et à quel stade du cycle de traitement du système l'intervention humaine a-t-elle lieu, et dans quelle mesure)

☐ Fournissez des explications adaptées à l'étape du cycle de vie de l'IA et au niveau de risque du système d'IA, et adaptez ces explications aux différences individuelles au sein du public cible, telles que son rôle, ses connaissances et son niveau de compétence

☐ Utilisez diverses méthodes pour expliquer les systèmes d'IA, telles que

- Des visualisations
- L'extraction de modèles
- Importance des caractéristiques

☐ Veillez à ce que les explications fournies sur le système d'IA correspondent aux processus utilisés par celui-ci pour générer le résultat :

- Rendez les explications pertinentes :
 - Utilisez un langage, des concepts et des termes qui soient les plus compréhensibles possible pour le public visé par l'explication (c'est-à-dire qu'une même explication peut être fournie avec différents niveaux de complexité en fonction du public)
- Vérifiez que le système est utilisé uniquement pour l'usage auquel il est destiné

☐ Envisagez de rendre le code source de l'algorithme accessible au public, ou à tout le moins disponible sur demande

☐ Mettez en place des procédures permettant aux utilisateurs finaux et aux personnes/communautés concernées par les décisions de l'IA de signaler des problèmes et de contester les résultats du système ;

☐ Communiquez les incidents et les erreurs aux acteurs concernés par l'IA ainsi qu'aux personnes et communautés concernées

☐ Faites preuve de transparence quant à la conception des systèmes d'IA – divulguer des informations sur :

- Les objectifs de conception
- Les données d'entrée
- La construction et le fonctionnement du système
- Les résultats du système et leurs impacts sur les personnes/communautés ciblées et/ou la société dans son ensemble

Interprétabilité



Déterminez quelles informations doivent être interprétables afin que la conception et les résultats du système puissent être facilement expliqués :

- Les informations nécessaires à l'interprétabilité dépendront des parties prenantes et du contexte des exigences en matière de responsabilité



Réexaminez régulièrement la documentation expliquant le fonctionnement du système d'IA, la manière dont il prend des décisions et génère des résultats, et mettez-la à jour si nécessaire. La documentation doit aborder les points suivants :

- Le suivi du temps consacré aux examens
- Le personnel impliqué dans la révision
- Les décisions/résultats qui font l'objet d'un examen
- Les mises à jour effectuées, le cas échéant



Rendez les systèmes d'IA vérifiables :

- Permettez à des tiers d'examiner, de comprendre et d'évaluer les systèmes d'IA avant leur déploiement et pendant leur fonctionnement.

Évaluation de l'impact



Mettez en œuvre des processus permettant d'évaluer, d'identifier, de documenter et de traiter les impacts et les risques associés à la prise de décision automatisée et aux résultats des systèmes d'IA :

- L'évaluation doit inclure les éléments suivants :
 - Une description de l'utilisation prévue du système d'IA
 - Une description des processus de l'organisation permettant de déterminer si le système d'IA sera utilisé conformément à l'usage prévu
 - Une description de la durée et de la fréquence d'utilisation prévues du système d'IA
 - Les types de données utilisées par le système en entrée et les types de données générées par ses résultats
 - Les catégories ou caractéristiques démographiques des personnes et des groupes susceptibles d'être affectés par l'utilisation du système
 - L'impact positif du système d'IA sur les individus, la société ou l'environnement
 - Les risques spécifiques de préjudice susceptibles d'affecter les catégories d'individus ou de groupes identifiées
 - Une description des mesures de contrôle humain
 - Les mesures à prendre au cas où ces risques se concrétiseraient, y compris les dispositions relatives à la gouvernance interne et aux mécanismes de réclamation.



Afin d'analyser efficacement les risques pour l'organisation, l'évaluation doit :

- Identifier les conséquences potentielles pour l'organisation et les personnes concernées par les résultats du système d'IA si un risque identifié venait à se concrétiser
- Déterminer les niveaux de risque en fonction de la sensibilité des données au sein du système d'IA et de ses utilisations prévues
- Déterminer les contrôles nécessaires pour atténuer les risques identifiés

- Mettez l'évaluation à la disposition des parties intéressées et des parties prenantes concernées, telles que déterminées par l'organisation ou la législation applicable

☐

Hiérarchisez les risques liés à l'IA identifiés en fonction de leur impact, de leur probabilité et des ressources ou méthodes disponibles.

☐

Élaborez, planifiez et documentez les réponses aux risques liés à l'IA jugés hautement prioritaires :

- Les options de réponse aux risques comprennent l'atténuation, le transfert, la prévention ou l'acceptation du risque

Supervision et intervention humaines

☐

Documentez les limites des connaissances du système d'IA et la manière dont les résultats du système peuvent être utilisés et supervisés par des humains (c'est-à-dire l'intervention humaine)

☐

Mettez en œuvre des mécanismes visant à alerter les opérateurs humains en cas de résultats ou d'impacts indésirables du système d'IA

☐

Mettez en place des mécanismes de contrôle et d'intervention humains lorsque la finalité du traitement est susceptible d'aboutir à des décisions entraînant des effets juridiques néfastes ou des effets tout aussi significatifs, pouvant inclure, sans s'y limiter, tout risque pour les droits et libertés des personnes, en particulier lorsque le système d'IA :

- ne peut pas détecter ou corriger ses propres erreurs
- nécessite une supervision humaine pour que le système d'IA fonctionne correctement conformément à ses instructions ou à toute autre documentation relative à son utilisation prévue

☐

Définissez, évaluez et documentez les mécanismes de contrôle et d'intervention humains conformément aux politiques de l'organisation

☐

Déterminez le niveau approprié de supervision et d'intervention humaines nécessaire :

- Les opérateurs humains doivent avoir le pouvoir de passer outre les erreurs et les décisions prises par le système d'IA
- Les intervenants humains doivent disposer de processus clairs et documentés et être soutenus par des mécanismes d'interprétabilité

☐

Documentez des procédures détaillées à l'intention des intervenants humains concernant la prise de décision et la mise en œuvre des mesures nécessaires qui en découlent, notamment :

- La manière dont un opérateur humain est informé ou prend connaissance des situations nécessitant une intervention (par exemple, contrôles proactifs, surveillance régulière des résultats à la recherche d'inexactitudes, audits aléatoires)
- Les procédures permettant à un opérateur humain de corriger les erreurs du système d'IA

☐

Mettez en place des mécanismes et des moyens permettant aux personnes concernées par les résultats du système d'IA de signaler des effets indésirables (par exemple, une injustice) et/ou de demander une intervention humaine

Formation et sensibilisation

☐

Informez et formez vos employés sur leurs responsabilités lorsqu'ils travaillent avec des systèmes d'IA dès qu'ils commencent à exercer des fonctions professionnelles qui s'appuient sur les résultats fournis par ces systèmes, notamment :

- Les implications et les conséquences de l'utilisation d'outils d'IA sur le lieu de travail (par exemple, les résultats de certains outils peuvent être inexacts) :
 - Gardez à l'esprit que l'utilisation de l'IA et la notion d'éthique peuvent être des concepts nouveaux pour certains employés
- Leurs responsabilités et obligations concernant l'utilisation des données et du système, l'interprétation des résultats du système, la sécurité et la confidentialité
- Les implications du non-respect des exigences et de l'usage prévu du système d'IA, ainsi que les répercussions des décisions prises sur la base des résultats fournis par le système

☐

Indiquez clairement aux employés s'il existe des outils spécifiquement recommandés ou interdits :

- Fournissez des directives claires sur la manière dont les employés peuvent ou ne peuvent pas utiliser les outils d'IA pour exercer leurs fonctions essentielles
- Sensibilisez les employés aux politiques pertinentes, telles que la politique de l'organisation en matière d'IA ou la politique d'utilisation acceptable

☐

Rappelez aux employés que les obligations légales applicables continuent de s'appliquer à l'utilisation des nouveaux outils ;

☐

Mettez en garde les employés contre la saisie de données à caractère personnel, d'informations commerciales confidentielles, de secrets d'affaires ou d'autres données sensibles dans les systèmes d'IA

☐

Envisagez de mettre en place des formations spécifiques pour aider les employés à comprendre et à limiter la responsabilité juridique liée à l'utilisation de certains outils d'IA, en particulier pour les entreprises évoluant dans un secteur réglementé

- Mettez à jour les ressources destinées aux employés, notamment les guides du personnel, afin d'y refléter les politiques relatives à l'utilisation de l'IA

☐

Établissez et imposez des directives de développement afin de responsabiliser les employés utilisant des systèmes d'IA

☐

Mettez en place et communiquez des protocoles à l'intention des employés utilisant des applications d'IA sur des appareils fournis par l'entreprise :

- Indiquez aux employés quels paramètres ou autorisations sont acceptables

☐

Fournissez aux employés des ressources traitant de l'utilisation responsable de tous les types d'IA. Organisez des sessions de formation et des ateliers sur différents aspects de l'IA, tels que :

- L'éthique et les biais
- La minimisation des données
- Exactitude/inexactitude des données
- La transparence et l'explicabilité
- Sécurité des données

- ☐ Dispensez une formation spécialisée aux employés chargés des processus d'intervention humaine : élaborer, documenter et former les équipes à tous les processus d'intervention humaine avant le lancement ou l'utilisation d'un système d'IA, afin que les clients puissent recourir à cette alternative dès le premier jour
- ☐ Renforcez la formation en mettant en place des sessions de remise à niveau à intervalles réguliers afin de tenir les employés informés des restrictions et autorisations légales en vigueur ainsi que des risques émergents associés aux systèmes d'IA (par exemple, une fois par an)

Suivi et contrôle

- ☐ Identifiez et évaluez la manière dont l'entreprise mesurera les risques liés à l'IA (par exemple, du plus important au moins important) ; les risques qui ne peuvent être mesurés doivent être correctement documentés :
 - Mettez en place des indicateurs de sécurité qui reflètent :
 - La fiabilité et la robustesse du système
 - La surveillance en temps réel
 - Le temps de réponse en cas de défaillance du système
 - Reliez les approches de mesure permettant d'identifier les risques liés à l'IA au(x) contexte(s) de déploiement
- ☐ Mettez en œuvre des plans post-déploiement pour surveiller en continu le fonctionnement et le comportement des systèmes d'IA et de leurs composants, afin de s'assurer qu'ils fonctionnent comme prévu pendant toute la durée du déploiement :
 - Pour les systèmes tiers, déterminez si le système d'IA fonctionne conformément à l'usage prévu tel que décrit par le développeur de l'IA
 - Vérifiez que les résultats sont conformes aux attentes de l'organisation
 - Intégrez les retours d'expérience et les problèmes signalés par les utilisateurs finaux et les personnes/communautés concernées par les décisions de l'IA dans les indicateurs d'évaluation
 - Mettez en œuvre des politiques et des procédures qui traitent des points suivants :
 - Qui participera au suivi et à l'analyse
 - Ce qui doit faire l'objet d'un suivi
 - Les méthodes nécessaires au suivi, à l'analyse et à l'évaluation
 - La fréquence du suivi
 - Les performances attendues en fonction de l'usage prévu
 - Quand le suivi et l'analyse seront-ils effectués ?
 - Les indicateurs définis pertinents pour le suivi
- ☐ Déterminez les mesures à prendre si les systèmes d'IA ne fonctionnent pas conformément à leur utilisation prévue, par exemple :
 - Mettez à jour et réentraînez les entrées et/ou les sorties du système d'IA afin que celui-ci fonctionne conformément à son usage prévu (dans la mesure où l'organisation a le contrôle sur les entrées des systèmes tiers)
 - Cessez d'utiliser le système d'IA lorsque ses sorties sont susceptibles d'entraîner des décisions ayant des conséquences juridiques défavorables ou des effets similaires d'une importance comparable, et, dans le cas d'un système tiers, informez le développeur que le système ne fonctionne pas conformément à l'usage prévu



Évaluez régulièrement l'efficacité des indicateurs utilisés pour mesurer les risques liés à l'IA et mettez-les à jour si nécessaire :

- Les évaluations doivent être menées par des évaluateurs indépendants ou des experts internes n'ayant pas participé à la conception et au développement du système.



Mettez en place des mécanismes et des moyens permettant aux tiers de signaler des effets indésirables (par exemple, des injustices) :

- Mettez en place des procédures permettant d'examiner, de corriger ou d'annuler les résultats préjudiciables, en particulier ceux qui aboutissent à des décisions entraînant des effets juridiques préjudiciables ou des effets similaires d'une importance comparable, y compris tout risque pour les droits et libertés des personnes – tels que :
 - Des procédures d'examen des réclamations
 - Des procédures permettant d'évaluer et de déterminer si une non-conformité ayant entraîné des conséquences négatives pour les personnes s'est produite
 - Une procédure permettant d'évaluer s'il existe des non-conformités similaires
 - Un processus permettant de réexaminer les procédures susmentionnées et d'y apporter des modifications si nécessaire



Obtenez et consignez les retours d'information des experts du domaine et des acteurs concernés de l'IA sur les résultats des mesures relatives à la fiabilité des systèmes d'IA dans le(s) contexte(s) de déploiement et tout au long du cycle de vie de l'IA, afin de vérifier si le système fonctionne comme prévu



Mettez en place des procédures de détection, de remontée et de gestion des incidents liés aux systèmes d'IA, ainsi qu'un plan d'intervention, ou mettez à jour les procédures et le plan d'intervention existants afin de les adapter aux systèmes d'IA

Tests

Avant le déploiement :



Mettez en œuvre des processus et des procédures pour tester le système d'IA avant son déploiement. Les politiques et procédures doivent inclure les éléments suivants :

- Un plan de test décrivant les objectifs des tests (par exemple, les critères de mise en production) et les indicateurs de performance à atteindre
- Des tests à la fois automatisés et manuels, menés par des opérateurs humains, tenant compte de la technologie utilisée et du rôle des opérateurs humains sur les résultats et l'efficacité du système d'IA
- Des cas d'utilisation concrets dans lesquels le système d'IA sera déployé – reproduire ces cas d'utilisation aussi fidèlement que possible
- Une comparaison des performances du système d'IA par rapport aux processus actuels menés par des humains

Après le déploiement : ⁴



Examinez régulièrement les données d'entrée et testez le système d'IA afin d'identifier les biais du système :

Mettez en place des processus garantissant l'intégration de mesures aux différentes étapes (par exemple, l'obligation d'utiliser un outil ou une méthode de test spécifique pour remédier aux injustices ou aux biais indésirables)



Mettez en place des mécanismes permettant de vérifier la précision des résultats du système d'IA par rapport aux performances attendues du système :

- Par exemple, mettez en œuvre des procédures de test pour mesurer la précision des résultats du système d'IA qui répondent aux critères suivants :

⁴ Voir « Procédures de test de l'IA » dans la catégorie 4 de la gestion de la vie privée – Maintenir des politiques/procédures pour la responsabilité algorithmique

- 

- Validez les hypothèses formulées au sujet du système d'IA lors de la phase de conception ou, en cas d'utilisation d'un système d'IA tiers, lors de la phase de tests préalable au déploiement
- Vérifiez que la collecte et/ou le traitement des données par le système sont conformes à l'objectif visé par celui-ci
- Assurez-vous que l'impact des résultats ou des décisions pris par le système d'IA n'est préjudiciable à aucun individu, à aucune communauté ni à la société dans son ensemble



Détection des incidents et réponse



- Mettez en place des mécanismes permettant de déterminer si un incident concerne des données à caractère personnel au sein d'un système d'IA

INSTRUMENTATION: <https://www.fishbase.org/traits/traits-portal/traits/ANISDA11dC4line20230> based on the fishTRUST team for Anisodactylus

TrustArc
Commencez votre essai gratuit