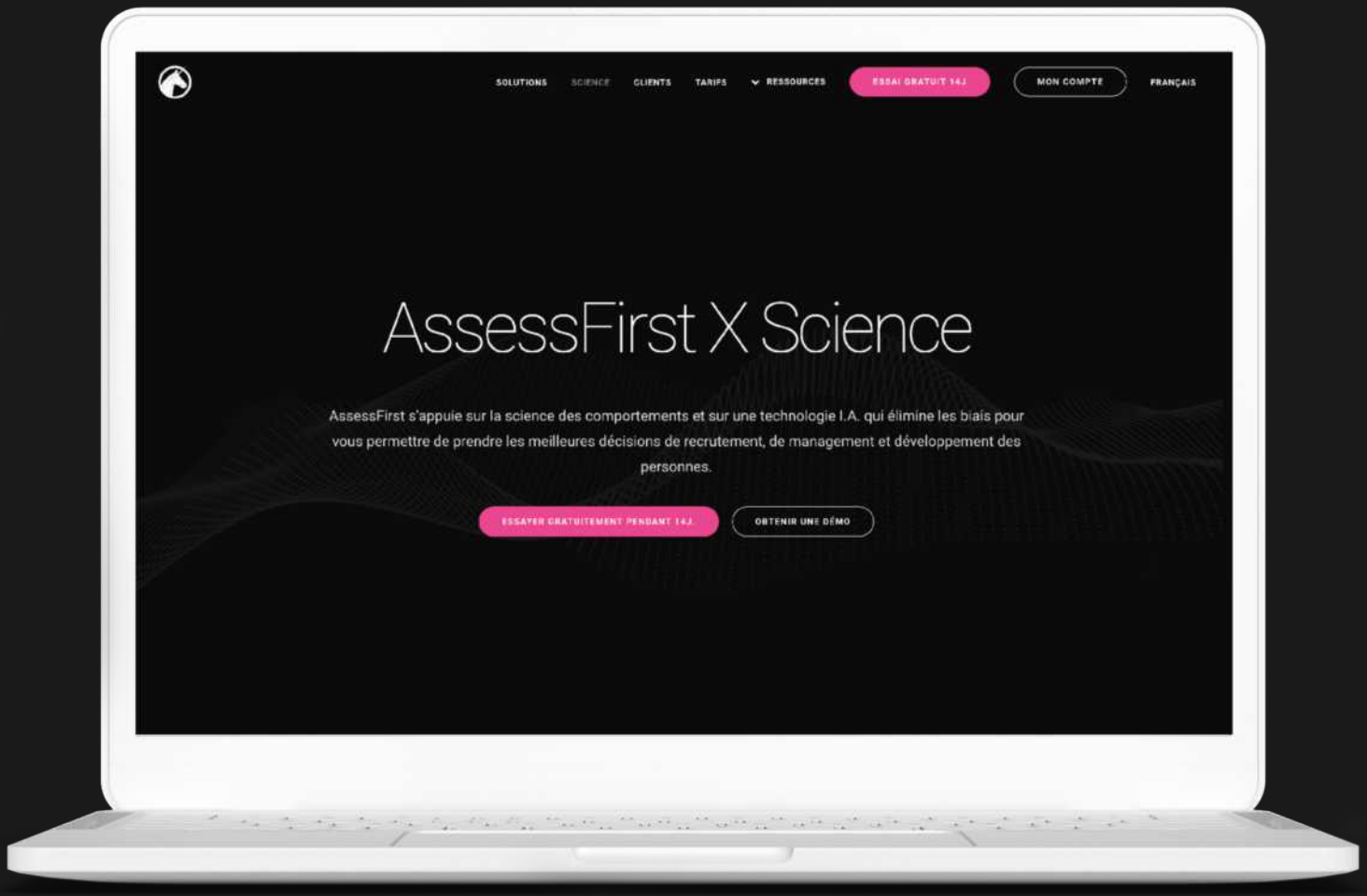


Guide d'auto-évaluation pour les systèmes d'IA

Le guide d'auto-évaluation proposé dans les pages suivantes a pour objectif d'étudier le positionnement d'AssessFirst au regard de l'ensemble des aspects pertinents en matière de données personnelles et d'éthique d'un projet de traitement dans le cadre d'un système d'intelligence artificielle. À travers plusieurs fiches d'évaluation conçues par la CNIL, ce guide justifie des bonnes pratiques éthiques, déontologiques et techniques d'AssessFirst.

SCIENCE X LEGAL

Préface

AssessFirst est une solution de recrutement et de gestion prédictive des talents, qui se concentre sur la compréhension profonde du potentiel humain afin d'aider les entreprises à prendre de meilleures décisions Talents. AssessFirst couple sciences comportementales et intelligence artificielle pour analyser les caractéristiques psychologiques, comportementales et cognitives des candidats, et de prédire leur succès et leur épanouissement dans un rôle et un contexte donnés. Aussi, la mise en conformité de ces outils avec les standards préconisés par l'Association Américaine de Psychologie (APA) et la Commission Internationale des Tests (ITC) permet aujourd'hui à AssessFirst de garantir un haut niveau de qualité dans les conception des questionnaires et d'améliorer continuellement la fiabilité de ses outils d'évaluation. Avec de nombreuses publications dans des journaux scientifiques ou congrès internationaux soumis à peer-review, la Science AssessFirst est mondialement reconnue.

Dans cette démarche, l'éthique est un pilier essentiel, qui anime et régule le développement de notre intelligence artificielle (IA). Premièrement, le respect des utilisateurs est primordial : en matière d'IA, l'éthique est étroitement liée à la manière dont les données personnelles sont traitées. Il est donc fondamental pour AssessFirst de garantir une utilisation transparente et sécurisée de ces données, dans le respect des droits et des attentes de chaque individu. Deuxièmement, l'éthique est au cœur de la fiabilité du modèle d'IA : une IA doit être équitable et non discriminatoire. Pour cela, elle doit être conçue de manière à éviter tout biais, que ce soit lors de la collecte de données ou lors de leur traitement. Une IA qui respecte l'éthique prend des décisions sur la base de critères justes et objectifs. Troisièmement, l'éthique joue un rôle déterminant dans la confiance accordée à l'IA. Les utilisateurs, ainsi que le grand public, doivent être assurés qu'AssessFirst utilise l'IA de manière responsable. Cette confiance est un élément fondamental pour le succès à long terme de l'entreprise. Enfin, l'éthique de l'IA est un élément clé pour la conformité légale. De nombreux pays ont commencé à mettre en place des réglementations pour encadrer le développement et l'utilisation de l'IA. Le non-respect de ces règles peut ou pourra entraîner des sanctions sévères. En intégrant l'éthique dès le début du développement de l'IA, AssessFirst s'assure de respecter les meilleures pratiques légales concernant l'IA.

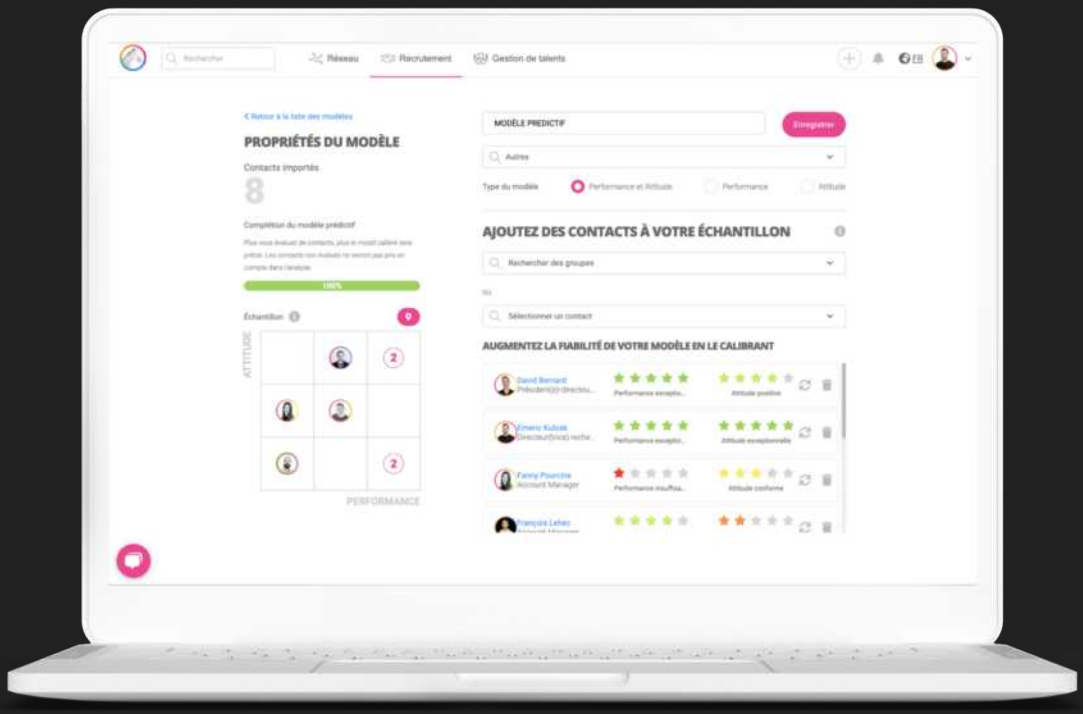
En somme, l'éthique est fondamentale pour AssessFirst dans le développement de son modèle d'IA. Elle contribue à protéger les utilisateurs, à améliorer la fiabilité de l'IA, à créer de la confiance et à assurer la conformité légale. L'éthique n'est pas seulement une question de faire ce qui est juste - elle est aussi une nécessité intellectuelle, commerciale et juridique. C'est dans ce cadre qu'a été rédigé ce document : basé sur le guide d'auto-évaluation proposé par la CNIL, il permet de faire l'état des lieux des pratiques d'AssessFirst en matière d'IA, et de démontrer l'éthique et la fiabilité de notre modèle d'IA.

Emeric Kubiak
Head of Science

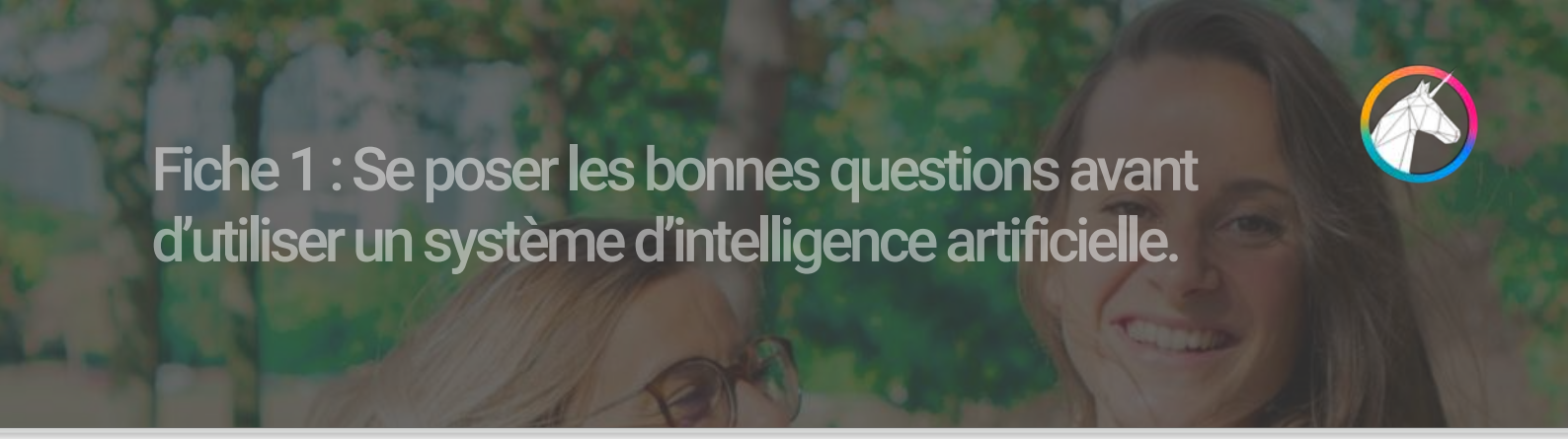
Lucile Whitbeck
Head of Legal & DPO

Note introductive

Il est important de souligner que ce guide a été soigneusement élaboré en mettant l'accent sur l'état actuel de l'art de l'intelligence artificielle, tel qu'il se reflète dans l'application AssessFirst en juin 2023. Pour l'instant, l'usage de l'intelligence artificielle chez AssessFirst est exclusivement réservé à la création de modèles prédictifs basés sur l'analyse de contacts - voir image présentée ci-dessous. Les autres typologies de modèles prédictifs, tels que les modèles de benchmarking et les modèles personnalisés, n'intègrent pas d'intelligence artificielle en phase de production. De la même façon, aucune autre fonctionnalité de l'application n'implémente, à ce jour, l'intelligence artificielle. Par conséquent, ce document a été rédigé en tenant compte du fait que le module d'analyse de contacts est le seul outil d'intelligence artificielle actuellement disponible dans l'application AssessFirst. Cependant, nous sommes pleinement conscients de la dynamique rapide de l'évolution des technologies associées à l'IA et de leur intégration croissante dans les processus de production chez AssessFirst. C'est pourquoi, afin d'assurer la pertinence et l'actualité de ce guide, nous nous engageons à le réviser semestriellement en fonction des évolutions de notre produit. Cette approche nous permet de rester constamment alignés sur les dernières avancées et pratiques adoptées chez AssessFirst .



Sommaire



Fiche 1 : Se poser les bonnes questions avant d'utiliser un système d'intelligence artificielle.

Intégrer l'IA de manière proportionnée et en définissant un objectif clair.

Objectif de proportionnalité

Des données personnelles sont-elles traitées par le système ?

Oui, des données personnelles sont traitées par les systèmes d'IA AssessFirst. Il s'agit des données suivantes : (1) résultats aux questionnaires SWIPE, DRIVE, BRAIN, (2) évaluations de performance et d'attitude (si applicable).

L'objectif du traitement reposant sur l'utilisation d'un système d'IA est-il clairement défini ?

Dans le cadre de l'utilisation du système d'IA, l'objectif du traitement est clairement défini et a trait à la création de modèles prédictifs du succès et/ou de l'épanouissement dans un poste spécifique. Ces modèles prédictifs sont utilisés par les clients d'AssessFirst afin de mieux gérer leurs actions de recrutement et de mobilité, en ayant accès à des analyses fiables et objectives.

Les phases d'apprentissage et de production du système d'IA sont-elles distinctes ? Si oui, une seconde évaluation est-elle prévue concernant la phase de production ?

Oui, les phases d'apprentissage et de production de nos systèmes d'IA sont distinctes. Elles représentent deux étapes différentes dans le cycle de vie de nos systèmes d'IA :

- Phase d'apprentissage (ou phase d'entraînement) : Pendant cette phase, le système d'IA est alimenté avec un ensemble de données d'apprentissage afin d'acquérir des connaissances et de développer sa capacité à effectuer une tâche spécifique (e.g. identifier les traits de personnalité qui expliquent la performance sur un poste spécifique). Cette phase implique généralement l'utilisation d'algorithmes d'apprentissage automatique et supervisé, tels que différents types de régressions. L'objectif est d'optimiser les performances du modèle en minimisant l'erreur entre les prédictions du modèle et les vérités terrain associées aux données d'apprentissage ;

- Phase de production (ou phase d'inférence) : Une fois que le modèle d'IA a été entraîné, il est déployé dans un environnement de production pour effectuer des prédictions ou prendre des décisions sur de nouvelles données (e.g., identifier les personnes qui ont le plus de probabilité de succès sur un poste). Pendant cette phase, le modèle utilise les connaissances qu'il a acquises lors de la phase d'apprentissage pour effectuer des tâches en fonction des entrées fournies ;

Il est à noter que bien que les phases d'apprentissage et de production soient distinctes, elles sont interdépendantes. Les performances du modèle dans la phase de production dépendent de la qualité des données d'apprentissage utilisées et de la manière dont il a été entraîné. De plus, le modèle peut nécessiter une rétroaction régulière de la phase de production pour être amélioré, en utilisant des mécanismes tels que le retour d'information des utilisateurs ou l'apprentissage par renforcement.

En ce qui concerne une seconde évaluation pour la phase de production, AssessFirst produit des évaluations supplémentaires après le déploiement du modèle dans un environnement de production, car les conditions réelles peuvent différer des conditions de test initiales. Cela peut être dû à des variations dans les données d'entrée, des contraintes de temps réel, des changements de distribution des données, etc. Il est important de noter que l'évaluation continue et le suivi des performances du modèle dans la phase de production sont essentiels pour identifier les problèmes potentiels, ajuster et améliorer le modèle si nécessaire, et assurer une utilisation efficace et fiable du système d'IA. À ce titre, AssessFirst produit de manière régulière des études de validité des modèles suite à leur mise en production, généralement de manière bi-annuelle ou annuelle. Ces études ont vocation à étudier les performances du modèles en production et à apporter les correctifs et optimisations nécessaires pour améliorer le modèle et assurer sa pérennité.

Les personnes ayant vocation à interagir avec le système d'IA ou envers lesquelles une décision automatisée sera prise, sont-elles identifiées ? Quelles sont leurs caractéristiques (âge, genre, spécificités physiques, etc.) ? Quel est leur nombre ?

Les personnes ayant vocation à interagir avec le système d'IA sont identifiées et prises en compte lors de la conception et de l'utilisation du système d'IA. Cela est particulièrement important pour garantir la transparence, l'éthique et la responsabilité dans l'utilisation des systèmes d'IA. Il est toutefois à noter que les systèmes d'IA AssessFirst n'amènent pas de prise de décision automatisée : la décision est toujours celle d'un utilisateur humain. L'identification des personnes concernées concerne deux niveaux :

- Les utilisateurs « clients » du système : Les personnes qui interagissent directement avec le système d'IA, que ce soit pour fournir des entrées, recevoir des résultats ou prendre des décisions basées sur les informations fournies par le système. Ils sont identifiés lors de la création de leur compte, et notamment via leur nom, prénom, adresse mail et entreprise cliente. ;

- Les bénéficiaires « candidats » ou « employés » : Les personnes ou groupes qui sont affectés par les décisions prises à partir des résultats des systèmes d'IA. Ils sont identifiés lors de la création de leur compte, et notamment via leur nom, prénom, et adresse mail. Sur une période allant du 01.09.2020 au 01.09.2021, en France, ces personnes présentent les caractéristiques suivantes (voir tableaux ci-dessous). À noter qu'AssessFist a mis en place une suppression automatique des comptes candidats inutilisés au bout de 2 ans.

Échantillon	Total	Hommes	Femmes
France	466 537	47 %	53 %

Tableau 1. Distribution du genre.

Échantillon	Catégorie d'âge	Total
France	15-19	4 %
	20-24	23 %
	25-29	21 %
	30-34	15 %
	35-39	12 %

France	40-44	9 %
	45-49	8 %
	50-54	5 %
	55-59	3 %
	60-64	0,5 %
	65-69	0,1 %
	70	0,1 %

Tableau 2. Distribution de l'âge.

Les résultats en termes de distribution des catégories d'âge montrent que les tranches 20-24 ans, 25-29 ans, 30-34 ans, 35-39 ans, 40-44 ans et 45-49 ans sont davantage représentées que les catégories d'âge « extrêmes ». Cette distribution reflète réalistement le cœur d'activité de la solution AssessFirst : nos questionnaires sont, en effet, principalement utilisés dans le cadre de décisions de carrière (orientation, recrutement, talent management). Cela implique que les utilisateurs de la solution sont en âge de travailler. De même, ces chiffres correspondent aux données de description de la population active française présentée par l'INSEE, entre 2014 et 2020 : la tranche d'âge 25 à 49 ans étant celle où l'on retrouve le plus de personnes actives et qui présente le taux de mobilité le plus important (étude Deloitte 2015).

Échantillon	Diplôme	Total
France	Doctorat	1 %
	Master	32 %
	Licence	36 %
	Baccalauréat	16 %
	Diplôme professionnel	10 %
	Aucun	4 %

Tableau 3. Distribution du niveau de diplôme.

Les résultats en termes de distribution des niveaux d'éducation montrent une légère sur-représentativité des utilisateurs ayant une Licence ou un Master. Cette sur-représentativité s'explique par 2 raisons principales : (1) la représentativité naturelle et croissante des personnes ayant un niveau d'enseignement tertiaire dans la population française - environ 50% selon les chiffres de l'OCDE entre 2017 et 2020, et, (2) l'utilisation légèrement plus accrue de la solution pour le recrutement de profils cadres par nos clients.

Échantillon	Niveau hiérarchique	Total
France	Associé(e)	2 %
	Vice président(e)	0,1 %
	Comité de direction	1,4 %
	Directeur(rice)	6 %
	Manager	12 %
	Expérimenté(e)	39 %

France	Jeune diplômé(e)	22 %
	Stage et alternance	9 %
	Non applicable	9 %
	Bénévoles	0,2 %

Tableau 4. Distribution du niveau hiérarchique.

Le traitement pourra-t-il viser directement ou indirectement des personnes vulnérables (enfants, patients, employés) ?

En tant que systèmes d'IA utilisés pour prendre des décisions de recrutement ou de mobilité interne, les systèmes d'IA AssessFirst visent directement des personnes qualifiées de vulnérables, à savoir des employés et des candidats à un poste (recrutement).

Le système d'IA remplace-t-il un autre type de système pour la tâche qui lui est confiée? Pour quelle raison celui-ci doit-il être remplacé?

Nos systèmes d'IA sont utilisés pour aider à prendre des décisions de recrutement et de mobilité interne, et nous reconnaissons l'importance de considérer les conséquences potentielles sur les personnes impliquées. Bien que les impacts puissent être plus limités par rapport à d'autres domaines sensibles tels que la santé ou les finances, nous prenons ces considérations très au sérieux. Nous nous engageons dès lors à garantir l'équité, la transparence et la protection des données dans nos systèmes. Nous veillons aussi à ce que les critères utilisés soient pertinents dans le cadre du recrutement ou de la mobilité, et exempts de biais discriminatoires. Il est toutefois à noter que les systèmes d'IA AssessFirst n'amènent pas de prise de décision automatisée : la décision finale est toujours celle d'un utilisateur humain. Candidats et employés ont ainsi l'opportunité de discuter des recommandations et des résultats produits par nos systèmes d'IA. Notre objectif est de maximiser les opportunités professionnelles et la carrière des individus concernés tout en minimisant les risques d'effets négatifs. Nous surveillons en permanence les performances du système, évaluons les risques et nous efforçons d'adopter les meilleures pratiques éthiques et responsables dans notre utilisation de l'IA pour le recrutement et la mobilité.

Le système d'IA présente-t-il un avantage notable (en termes d'efficacité technique, de coût, de protection de la vie privée, etc.) en comparaison avec les autres solutions disponibles? Cet avantage notable compense-t-il les risques supplémentaires potentiellement liés à l'utilisation d'un système d'IA?

Nos systèmes d'IA présentent plusieurs avantages notables par rapport aux autres solutions disponibles, notamment en termes d'amélioration de la qualité des décisions, de gain de temps pour les utilisateurs et les équipes RH. Ces avantages compensent les risques potentiels supplémentaires liés à l'utilisation d'un système d'IA. AssessFirst mène régulièrement des études de validité prédictive de ses modèles auprès de clients cibles, permettant ainsi de conclure que, en moyenne, l'utilisation de la solution AssessFirst permet de :

- Recruter des candidats qui sont 33% plus performants ;
- De diviser le turniver par 2 ;
- Un gain de diversité de 30% ;
- Un time to hire qui diminue de 33% ;
- Un nombre d'entretiens à réaliser diminué de 50%.

Au vu de vos réponses aux questions précédentes, l'utilisation d'un système d'IA pour atteindre la finalité identifiée semble-t-elle proportionnée et nécessaire?

Oui, au vu des réponses précédentes, l'utilisation d'un système d'IA pour atteindre la finalité identifiée semble proportionnée et nécessaire. L'introduction d'un système d'IA dans le processus de recrutement et de mobilité interne vise à améliorer l'efficacité, l'objectivité et la qualité des décisions prises, tout en prenant en compte les préoccupations éthiques et les risques potentiels. Il est important de souligner que le recrutement et la mobilité interne sont des domaines où la prise de décision humaine peut être influencée par des biais cognitifs, des jugements subjectifs et des limitations individuelles. L'utilisation d'un système d'IA peut contribuer à atténuer ces biais et à améliorer la justesse des décisions prises.

Intégrer l'IA de manière proportionnée et en définissant un objectif clair.

Fournisseurs, utilisateurs de systèmes d'IA et individus

Si des données personnelles sont collectées et/ou utilisées, un responsable du traitement a-t-il été identifié?

AssessFirst considère agir comme responsable de traitement pour les traitements de données personnelles liées à son système d'IA.

Les personnes morales en charge du développement du système d'IA, de son déploiement et de son contrôle sont-elles clairement définies ?

Seule la société AssessFirst est en charge du déploiement et du contrôle de ses systèmes IA.

Les personnes physiques en charge du développement du système ont-elles la formation adéquate ? Ont-elles été sensibilisées aux enjeux juridiques, techniques, éthiques et moraux de l'IA ?

Les personnes responsables du développement des systèmes d'IA chez AssessFirst sont clairement identifiées et bénéficient d'une formation complète et nécessaire pour assurer un développement efficace et éthique de nos systèmes. Leur formation couvre un large éventail de domaines, notamment les sciences comportementales, la science des données, l'ingénierie des données, l'intelligence artificielle et l'apprentissage automatique. Cette expertise multidisciplinaire est essentielle pour garantir la qualité et la fiabilité de nos systèmes d'IA. En plus de leur formation initiale, ces professionnels ont également été formés et sensibilisés aux enjeux juridiques, techniques, éthiques et moraux de l'IA. Cela a été réalisé à travers diverses initiatives, telles que des interventions régulières menées par notre délégué à la protection des données (DPO).

Existe-t-il une charte ou politique interne encadrant la conception et le déploiement d'IA ?

À date d'écrite de ce document, il n'existe pas de charte ou politique interne qui encadre la conception et le déploiement de systèmes d'IA. Toutefois, AssessFirst a mis en place une politique de protection des données personnelles, qui englobe des sujets essentiels tels que la confidentialité des données, la transparence, la responsabilité et la protection de la vie privée. Cette politique de confidentialité sert de référence pour toutes nos pratiques en matière de gestion des données, garantissant ainsi un engagement fort envers la protection des informations de nos utilisateurs. En plus de notre politique de protection des données, nous nous engageons également à promouvoir la transparence et la confiance envers nos systèmes d'IA. Dans cet esprit, nous produisons régulièrement des études approfondies, telles que des livres blancs et des articles de recherche

publiés dans des revues scientifiques à comité de lecture, ainsi que des présentations lors de conférences scientifiques internationales. Ces études mettent en lumière la conception de nos systèmes d'IA, leurs performances et démontrent l'absence de biais.

Les personnes en charge du maintien du système d'IA, de la correction des problèmes et de l'intervention sur son fonctionnement sont-elles clairement identifiées et connues de tous ?

Oui, les personnes responsables de la maintenance du système d'IA, de la résolution des problèmes et de l'intervention sur son fonctionnement sont clairement identifiées et connues de tous. Il s'agit des membres de l'équipe Science d'AssessFirst, ainsi que de certaines personnes de l'équipe de développement backend.

Un suivi est-il effectué afin d'assurer une permanence de ce service ?

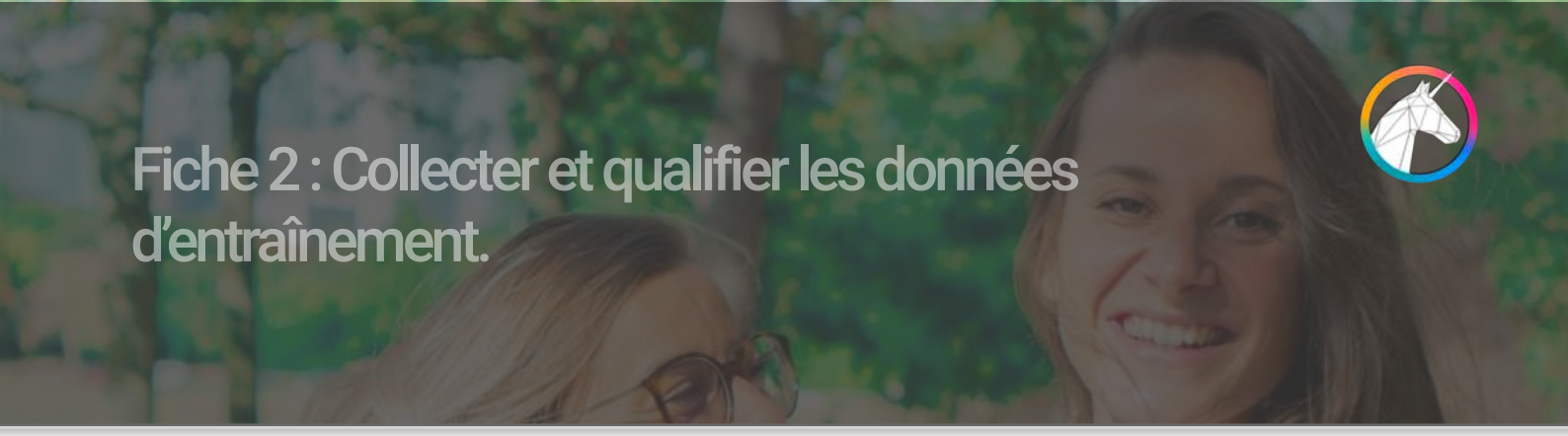
Oui, nous accordons une grande importance à assurer une permanence du service et à réagir rapidement en cas d'anomalies. Pour cela, sont mises en place les mesures suivantes:

- Équipe de support réactive : Nous avons constitué une équipe dédiée au support, chargée de gérer rapidement les anomalies identifiées sur l'application. Cette équipe est disponible pour répondre aux problèmes signalés par les utilisateurs et prendre les mesures nécessaires pour les résoudre dans les meilleurs délais ;

- Suivi continu par les équipes Science et Back : Nos équipes Science et Back assurent un suivi permanent de nos systèmes, même pendant les périodes de congés. Cela garantit une surveillance constante de nos services et la possibilité de réagir rapidement en cas de dysfonctionnement ou d'anomalie. Nous comprenons l'importance de maintenir la qualité et la disponibilité de notre service à tout moment ;

- Système de logs pour la traçabilité : Nous avons mis en place un système de logs qui enregistre les activités. Cela nous permet de remonter à l'origine d'un problème en cas de dysfonctionnement. Les logs nous fournissent des informations détaillées sur les opérations effectuées, et facilitent l'identification et la résolution des problèmes.

En combinant ces mesures, nous veillons à ce que notre service reste opérationnel et réactif aux besoins de nos utilisateurs.



Fiche 2 : Collecter et qualifier les données d'entraînement.

Respecter le RGPD lors de la collecte et constituer une base de qualité.

Traiter les données d'apprentissage de manière licite en respectant la réglementation.

Les données d'entraînement font-elles l'objet de réutilisation (réutilisation d'une base interne ou publiquement accessible, acquisition, etc.) ou d'une collecte spécifique ?

Les systèmes d'IA en production sont basés sur des données d'entraînement issues d'une collecte spécifique. L'objectif de cette collecte spécifique est d'obtenir des données de haute qualité qui permettent au modèle d'apprendre de manière précise, et de créer un modèle prédictif valide et pertinent. En ce sens, les données collectées concernent :

- Des données liées à la personnalité, aux motivations et aux capacités de raisonnement des individus. Ces données sont récoltées à travers 3 questionnaires dédiés : SWIPE (personnalité), DRIVE (motivations) et BRAIN (raisonnement). La mise en conformité de ces outils avec les standards préconisés par l'Association Américaine de Psychologie (A.P.A.) et la Commission Internationale des Tests (I.T.C.) permet aujourd'hui à AssessFirst de garantir un haut niveau de qualité dans la conception des questionnaires et d'améliorer continuellement leur fiabilité ;

- Des données liées à la performance et à l'attitude des personnes qui composent l'échantillon d'entraînement. Ces données brutes sont directement fournies et annotées manuellement par l'utilisateur "client" pour fournir des informations supplémentaires aux modèles d'apprentissage automatique.

Dans le cas d'une réutilisation, la base de données a-t-elle été constituée conformément à la réglementation en matière de protection des données ?

Les systèmes d'IA en production sont alimentés par des données d'entraînement provenant d'une collecte spécifique, et non par des données réutilisées. Pour garantir la qualité et l'efficacité de nos systèmes d'IA, nous nous assurons de collecter des données spécifiquement adaptées à chaque modèle. Plutôt que de simplement réutiliser des ensembles de données existants, nous procédons à une collecte ciblée et méthodique afin de garantir que les données utilisées pour l'entraînement de nos modèles sont représentatives et pertinentes pour les tâches spécifiques à accomplir.

Si une base de données publiquement accessible a été utilisée, a-t-elle fait l'objet d'études, en particulier en ce qui concerne la présence de biais ?

Les systèmes d'IA en production sont alimentés par des données d'entraînement provenant d'une collecte spécifique, et non par des données publiquement accessibles.

Sur quelle base légale s'appuie le traitement des données d'entraînement?

Le traitement des données d'entraînement des systèmes d'IA d'AssessFirst est effectué conformément aux lois et réglementations en vigueur sur la protection des données personnelles et la confidentialité. La base légale sur laquelle ce traitement s'appuie notamment sur le Contrat : le traitement des données est nécessaire à l'exécution d'un contrat avec les personnes concernées, en l'occurrence les Conditions Générales d'Utilisation d'AssessFirst.

Si le traitement porte sur des données sensibles (santé, infraction, etc.), le traitement n'est possible qu'en mobilisant une des exceptions prévues à l'article 9 du RGPD. Sur laquelle de ces exceptions repose le traitement ?

AssessFirst ne traite pas de données sensibles au sens de l'article 9 du RGPD pour le système d'IA.

Comment est suivie la conformité du traitement des données d'entraînement (réalisation d'une AIPD, d'une analyse des risques de réidentification, etc.) ?

AssessFirst a effectué une AIPD globale, mise à jour en mars 2023. Toutefois, AssessFirst n'a pas conduit d'AIPD unique dans le cadre du système d'IA.

Les jeux de données utilisés pour l'entraînement ont-ils été constitués de façon à satisfaire le principe de minimisation ?

Lors de la constitution des ensembles de données utilisés pour l'entraînement des modèles d'IA, AssessFirst adopte une approche de minimisation des données. Cela signifie que seules les données nécessaires à l'objectif spécifique de l'entraînement sont incluses, évitant ainsi la collecte excessive de données personnelles non pertinentes. Voici quelques pratiques mises en place par AssessFirst pour respecter le principe de minimisation des données lors de la constitution des ensembles de données d'entraînement :

- Collecte limitée : AssessFirst collecte uniquement les données qui sont strictement nécessaires pour atteindre l'objectif d'entraînement spécifique. Aucune donnée excessive ou non pertinente n'est collectée ;

- Anonymisation et pseudonymisation : Lorsque cela est possible, AssessFirst réduit les informations personnelles directes dans les données d'entraînement en utilisant des techniques d'anonymisation ou de pseudonymisation. Cela permet de réduire les risques potentiels pour la vie privée des individus ;

- Échantillonnage représentatif : AssessFirst utilise des techniques d'échantillonnage représentatif pour sélectionner un sous-ensemble représentatif des données plutôt que de collecter la totalité des données disponibles. Cela peut réduire la quantité de données nécessaires tout en maintenant la qualité de l'entraînement ;

- Durée de conservation limitée : AssessFirst a mis en place une suppression automatique des comptes candidats inutilisés au bout de 2 ans.

Les données sont-elles anonymisées ? Par quel moyen ?

Pour les systèmes d'IA en production, les données sont pseudonymisées.

Sont-elles pseudonymisées ? Par quel moyen ?

Pour les systèmes d'IA en production, les données sont pseudonymisées. La pseudonymisation implique le remplacement des données personnelles identifiables par des identifiants artificiels ou des pseudonymes. Les données pseudonymisées conservent une certaine association avec la personne concernée, mais l'identifiant utilisé est dissocié des informations personnelles directes. Contrairement à l'anonymisation, la pseudonymisation permet une réidentification ultérieure en utilisant des informations supplémentaires, comme une clé de correspondance sécurisée qui lie les pseudonymes aux informations d'identification réelles. La pseudonymisation est souvent utilisée pour protéger les données personnelles lorsqu'il est nécessaire de conserver une certaine utilité des données pour des fins spécifiques tout en réduisant les risques pour la vie privée. Dans le cadre des systèmes AssessFirst, les données sont pseudonymisées grâce à l'utilisation d'un user ID.

Les risques de réidentification ont-ils été évalués ?

Dans le contexte de notre système d'IA, il est important de préciser qu'il n'existe aucun risque qu'une personne puisse être réidentifiée par l'intermédiaire de l'accès à l'algorithme. En effet, seul un accès direct à la base de données pourrait potentiellement permettre une telle réidentification. L'accès à la base de données est contrôlé de manière stricte Ainsi, l'intégrité et la confidentialité des informations personnelles sont préservées, mettant en avant l'engagement de notre entreprise pour une utilisation sécurisée des technologies de l'IA.

Le volume des données recueillies est-il justifié au vu de la difficulté de la tâche d'apprentissage ?

Le volume de données nécessaire pour l'apprentissage des modèles d'IA peut varier en fonction de plusieurs facteurs, notamment la complexité de la tâche, la diversité des données, la qualité des données et les algorithmes d'apprentissage utilisés. Dans le cas des systèmes d'IA AssessFirst, les modèles utilisés (e.g., régression linéaire, logistique) sont généralement considérés comme moins complexes par rapport à d'autres approches d'apprentissage automatique, ce qui peut réduire les besoins en volume de données. Néanmoins, si AssessFirst conseille à ses clients d'utiliser ses systèmes d'IA avec le minimum de données disponibles, il reste toutefois nécessaire de disposer d'un nombre suffisant de données pour assurer la fiabilité des analyses. Ainsi, il est préférable d'avoir un plus grand échantillon pour une meilleure fiabilité des résultats (notamment pour les modèles de régression qui sont utilisés dans nos analyses). Le volume des données recueillies est ainsi pleinement justifié au vu de la nature de la tâche d'apprentissage.

Les variables envisagées pour l'entraînement du modèle sont-elles toutes nécessaires ?

AssessFirst attache une grande importance au respect du principe de minimisation des variables utilisées dans nos systèmes d'IA. Dans cette optique, seules les variables suivantes sont intégrées à l'entraînement : personnalité, motivations, raisonnement, performance et attitude. Cependant, nous tenons à souligner que selon l'objectif d'utilisation spécifique du client, certaines de ces variables peuvent ne pas être incluses dans le processus d'entraînement. Bien que nous reconnaissons la possibilité de personnaliser les modèles en fonction des besoins de chaque client, nous recommandons généralement d'intégrer l'ensemble de ces variables afin d'optimiser la précision des modèles. En incluant toutes ces dimensions, nos modèles sont en mesure de fournir des résultats plus complets et précis, offrant ainsi une meilleure compréhension des candidats ou des métiers évalués. Chez AssessFirst, nous nous efforçons de trouver le bon équilibre entre la minimisation des variables pour garantir la confidentialité et la protection des données, tout en prenant en compte les informations nécessaires pour fournir des recommandations pertinentes et précises. Notre approche est centrée sur l'optimisation des résultats en utilisant un ensemble approprié de variables, tout en respectant les besoins et les préférences spécifiques de nos clients.

La collecte de certaines valeurs qui ne s'avèreraient pas utiles à l'apprentissage, en particulier si elles constituent des données sensibles, pourrait-elle évitée ?

Nous n'effectuons aucune collecte de données qualifiées de sensibles au regard de l'article 9 du RGPD. De plus, nous adhérons strictement au principe de minimisation des données. Ainsi, toutes les informations recueillies par AssessFirst sont impérativement requises pour les besoins de l'apprentissage. Nous nous limitons exclusivement à l'utilisation de données indispensables pour cette tâche et excluons l'intégration de toute donnée supplémentaire.

Si leur collecte ne peut être évitée, pourraient-elles être supprimées ou masquées ?

Non applicable, voir question précédente.

Respecter le RGPD lors de la collecte et constituer une base de qualité.

Des données brutes à un ensemble de données d'entraînement de qualité

La véracité des données a-t-elle été vérifiée ?

Nous nous appuyons sur des tests psychométriques dont les qualités, telles que la validité, la fidélité, la sensibilité et l'équité, ont été rigoureusement évaluées et démontrées. Ces tests sont soigneusement construits par nos équipes pour garantir des mesures fiables et valides des caractéristiques psychologiques et comportementales. Pour plus d'informations détaillées, nous vous invitons à consulter notre manuel psychométrique qui fournit des informations supplémentaires sur ces tests.

Si une méthode d'annotation a été utilisée, est-elle vérifiée ?

Les seules données "annotées" qui sont intégrées sont celles fournies par nos clients, sous forme de scores de performance et/ou d'attitude notés de 1 à 5. Nous encourageons nos clients à baser ces notations sur des éléments d'évaluation objective et vérifiable, tels que les entretiens annuels d'évaluation. Cela permet d'apporter une dimension objective aux données annotées et de renforcer la validité des évaluations. Nous comprenons l'importance de disposer de données fiables et pertinentes pour assurer la qualité de nos modèles d'IA. Par conséquent, nous travaillons en étroite collaboration avec nos clients pour garantir que les notations fournies sont basées sur des critères objectifs et représentatifs de la performance et de l'attitude des individus évalués.

Si l'annotation est réalisée par des personnes, ont-elles été formées ?

Chez AssessFirst, nous attachons une grande importance à la qualité des annotations effectuées par nos clients. C'est pourquoi nos équipes de Customer Success jouent un rôle essentiel dans la formation et l'accompagnement quotidien des clients chargés de l'annotation. Nous comprenons que l'annotation des données peut être un processus délicat et nécessite une attention particulière pour garantir la fiabilité et l'objectivité des résultats. C'est pourquoi nos équipes sont là pour fournir un soutien constant aux clients, en les formant aux meilleures pratiques d'annotation et en répondant à toutes leurs questions et préoccupations. En travaillant en étroite collaboration avec nos clients, nous nous assurons que le processus d'annotation est bien maîtrisé et que les résultats obtenus sont cohérents et de haute qualité. Nous nous engageons à fournir un accompagnement continu pour garantir que nos clients se sentent soutenus tout au long du processus d'annotation et qu'ils puissent effectuer leur travail avec précision et confiance.

La qualité de leur travail est-elle contrôlée ?

La qualité du travail d'annotation est contrôlée selon trois approches complémentaires :

- Premièrement, nous encourageons les clients à réaliser les annotations en groupe. Cette méthode collective permet de limiter les éventuelles erreurs dues à l'intuition individuelle. De plus, lorsque cela est possible, nous utilisons des extraits de données provenant des équipes comptables du client pour effectuer les annotations. Cela garantit une base solide et objective pour le processus d'annotation ;

- Deuxièmement, notre équipe Customer Success apporte un regard critique et humain sur les notations réalisées. Un membre dédié de notre équipe travaille en étroite collaboration avec le client pour valider la méthodologie d'annotation, discuter de l'objectivité des données et s'assurer de la cohérence globale du processus. Cette approche permet d'obtenir des annotations de haute qualité, en tenant compte des spécificités et des objectifs du client ;

- Enfin, nous utilisons des analyses statistiques et des distributions pour évaluer la représentativité de l'échantillon d'annotations. Cela nous permet de détecter d'éventuels biais ou incohérences dans les annotations réalisées. En identifiant ces anomalies, nous pouvons prendre des mesures correctives pour garantir la fiabilité et la validité des résultats.

Ces trois approches combinées assurent un contrôle de qualité robuste tout au long du processus d'annotation. Elles permettent de minimiser les erreurs subjectives, de valider la méthodologie utilisée et de garantir la représentativité des annotations effectuées.

La qualité de leur travail est-elle contrôlée ?

Les données utilisées sont directement collectées en environnement réel, et plusieurs actions sont mises en place pour garantir leur représentativité, notamment :

- **Plan de collecte de données** : en collaboration avec le “client”, AssessFirst élabore un plan de collecte de données détaillé qui identifie les variables et les contextes spécifiques à capturer. Nous nous assurons que les données soient pleinement ciblées sur les situations et conditions rencontrées dans l'environnement réel ;

- **Taille de l'échantillon** : AssessFirst s’assure également de faire les analyses avec des échantillons d’une taille suffisante et nécessaire pour représenter de manière adéquate la variabilité des situations réelles. Une taille d'échantillon plus grande peut être nécessaire pour des environnements réels complexes ou diversifiés ;

- **Sélection stratifiée des données** : Une méthode consiste à sélectionner les données de manière stratifiée afin de représenter différentes caractéristiques ou classes présentes dans l'environnement réel, notamment en termes de notation de performance et/ou d’attitude ;

- **Équilibrage des classes** : AssessFirst s’assure aussi de collecter suffisamment d'exemples de chaque classe pour éviter les biais d'une représentation inégale ;

- **Échantillonnage aléatoire** : L'utilisation d'un échantillonnage aléatoire permet de sélectionner des données de manière aléatoire dans l'ensemble des données disponibles. Cela contribue à éviter les biais de sélection et à garantir une représentation plus équilibrée des observations réelles ;

- **Collecte de données longitudinales** : Si les données évoluent dans le temps, la collecte de données longitudinales sur une période prolongée peut fournir une meilleure représentation des variations et des changements observés dans l'environnement réel ;
- **Validation et vérification** : AssessFirst effectue des validations croisées pour garantir la qualité et la représentativité des données collectées en les comparant à des références externes ;
- **Expertise du domaine** : AssessFirst garantit un accompagnement du client par des experts du domaine qui ont une connaissance approfondie de l'environnement réel et des données peut aider à garantir la représentativité des données utilisées ;
- **Rétroaction des utilisateurs ou des parties prenantes** : AssessFirst sollicite les commentaires et les retours des utilisateurs ou des parties prenantes impliquées dans l'environnement réel pour évaluer la représentativité des données. Leurs perspectives peuvent contribuer à identifier des aspects importants à prendre en compte lors de la collecte de données.

En combinant ces stratégies, nous maximisons ainsi les chances d'obtenir des données représentatives et de qualité qui reflètent fidèlement l'environnement réel.

Une étude formalisée de cette représentativité a-t-elle été réalisée ?

AssessFirst ne réalise pas d'étude formalisée de représentativité pour chaque modèle créé.

Si le traitement repose sur une solution d'apprentissage fédéré, le caractère indépendant et identiquement distribué des données utilisées au sein des centres (condition garantissant que les informations tirées des données reflèteront les mêmes tendances sans spécificité propre à chaque centre) a-t-il été évalué ? S'il n'est pas vérifié, quelles mesures sont prises pour y remédier ?

Le traitement effectué ne repose pas sur un apprentissage fédéré.

Dans le cas d'un système d'IA utilisant de l'apprentissage continu, quel mécanisme est mis en œuvre afin de garantir la qualité des données utilisées de façon continue ?

Les données collectées de manière continue sont soumises aux mêmes normes et procédures que les données initiales. Elles sont recueillies au moyen de tests psychométriques dont la validité, la fidélité, la sensibilité et l'équité ont été reconnues et démontrées. Pour obtenir des informations détaillées, nous vous recommandons de consulter notre manuel psychométrique qui fournit des informations supplémentaires sur ces tests et leur utilisation. Ce manuel est une ressource précieuse pour comprendre en profondeur les principes et les standards appliqués dans la collecte continue des données.

Des mécanismes réguliers d'évaluation des risques de perte en qualité ou de changement dans la distribution des données sont-ils mis en œuvre ?

AssessFirst accorde une grande importance à la qualité des outils de collecte de données et mène régulièrement des études pour évaluer et prévenir les risques de perte de qualité. Nous nous engageons à faire évoluer en permanence nos outils afin de maximiser la qualité des données recueillies. De plus, des recalibrations périodiques sont effectuées sur certains échantillons de données chaque année pour éviter toute perte de qualité ou de précision due à d'éventuels changements dans leur distribution naturelle. Cette approche garantit la fiabilité et la pertinence des données utilisées dans nos analyses.

Respecter le RGPD lors de la collecte et constituer une base de qualité.

Identifier les risques de biais et les corriger efficacement

La méthode utilisée pour la collecte des données d'entraînement est-elle suffisamment connue ?

Oui, les données sont collectées en utilisant des tests psychométriques dont les qualités (validité, fidélité, sensibilité, équité) ont été reconnues et démontrées. Ces outils sont largement reconnus et optimisés depuis de nombreuses décennies, ce qui garantit leur fiabilité et leur pertinence. Pour plus de détails, nous vous invitons à consulter notre manuel psychométrique qui fournit des informations complémentaires sur ces tests. De plus, les données relatives à la performance et/ou à l'attitude des individus inclus dans l'échantillon reposent sur des échelles de collecte traditionnelles, qui sont des méthodes éprouvées et largement utilisées dans le domaine de l'évaluation.

Des biais peuvent-ils exister du fait de la méthode utilisée ou des conditions de la collecte ?

Il peut exister des biais du fait de la méthode utilisée ou des conditions particulières de la collecte, comme par exemple des biais liées aux conditions de collecte (e.g., si le questionnaire est administré dans un environnement bruyant ou perturbant, cela peut affecter la concentration des participants et donc la précision de leurs réponses.), ou à la culture et à la linguistique (e.g., des biais peuvent survenir en raison des différences culturelles ou des nuances linguistiques qui peuvent affecter la compréhension des questions.). Néanmoins, AssessFirst met en place toutes les actions pour identifier, prévenir et neutraliser ces potentiels biais. Les actions mises en place par AssessFirst pour assurer et améliorer l'accessibilité de la solution, mais aussi l'absence de biais, incluent :

- L'orientation professionnelle du contenu : Les questionnaires et les résultats de la solution AssessFirst ont été spécifiquement développés pour être pertinents dans un but professionnel. Les dimensions évaluées ont été choisies en raison de leur pertinence pour l'efficacité professionnelle. Les conclusions tirées de l'utilisation d'AssessFirst se limitent à ce cadre précis ;

- Le niveau de langue : AssessFirst s'appuie sur une équipe « Localization » composée de psychologues et d'experts de la gestion linguistique, afin de proposer des contenus textuels compréhensibles et accessibles par tous, dans toutes les langues (15 langues disponibles à ce jour). En ce sens, nous travaillons avec des traducteurs natifs afin de construire et de valider l'ensemble de nos contenus ;

- Validation psychométrique des questionnaires : La mise en conformité de ces outils avec les standards préconisés par l'Association Américaine de Psychologie (A.P.A.) et la Commission Internationale des Tests (I.T.C.) permet aujourd'hui à AssessFirst de garantir un haut niveau de qualité dans la conception des questionnaires et d'améliorer continuellement leur fiabilité ;

- Méthodologie de scoring optimisée : En intégrant des modèles I.R.T. (Item Response Theory) à nos algorithmes de scoring, nous sommes en mesure de réduire de 50% environ le nombre de questions à poser aux individus afin de dresser un bilan fiable de leurs talents, et d'assurer une meilleure précision de la mesure ;

- Fairness by design : Nous construisons nos questionnaires à partir de contenus neutres, dans la mesure où ils ne font référence à aucun code culturel ou social. Par exemple, sur le raisonnement, BRAIN fait appel exclusivement à des logiques primaires (rotations, transformations, répétitions...) qui ne sont pas affectées par des handicaps tels que la dyslexie ou le daltonisme. Ainsi, nos tests offrent des conditions de succès maximales pour tous ;

- Text to speech : AssessFirst a développé son propre outil de text-to-speech, ou lecture automatique des questionnaires. Cette fonctionnalité permet ainsi d'accéder à un assistant vocal qui lit les questions, renforçant dès lors l'accessibilité pour les personnes en situation de handicap visuel ;

- Gestion des contrastes : AssessFirst met en place des actions permettant de personnaliser les contrastes et l'affiche des contenus web, afin de les rendre plus faciles à lire pour les utilisateurs ayant des déficiences visuelles ;

- Intégration des clients : De nombreux partenariats sont mis en place avec des clients cible qui proposent la solution à des publics qui pourraient avoir des difficultés d'accessibilité à l'outil (e.g. utilisateurs en situation de handicap, qui sont éloignés de l'emploi, publics jeunes, publics éloignés du numérique, publics sans expérience professionnelle). Des échanges réguliers avec ces partenaires et les utilisateurs nous permettent de sans cesse améliorer la solution afin de répondre au mieux à leurs besoins. Des exemples sont présentés dans la suite de ce document ;

- Étude des biais : AssessFirst produit régulièrement des études (e.g., livre blanc, papiers de recherches publiés dans des revues scientifiques à comité de lecture, conférence dans des congrès scientifiques internationaux) pour démontrer l'absence de biais dans nos modules de collecte de données.

Ces actions permettent de neutraliser les biais, comme le démontre nos dernières études, menées en Juillet et Décembre 2022 :

- Que cela soit en termes de genre, de catégories d'âge, d'origine ethnique, ou de handicap, il n'existe pas de différences majeures entre les résultats obtenus par les différents groupes aux questionnaires SWIPE, DRIVE et BRAIN d'AssessFirst. En somme, les résultats obtenus permettent de soutenir l'hypothèse selon laquelle les questionnaires proposés ne discriminent aucun public, et s'avèrent équitables. Les effets les plus marqués sont relatifs aux handicaps (notamment handicaps mentaux et cognitifs) - même si les effets mentionnés sont très faibles ou faibles, et explicables par d'autres interactions. De même, les effets les plus marqués concernent le genre et uniquement quelques facettes. Aussi, ces effets mentionnés sont très faibles ou faibles, potentiellement explicables par d'autres interactions ou échantillonnages, et trouvent tous une justification conceptuelle ;

- L'usage des algorithmes proposés par AssessFirst s'avère non-discriminant et équitable : les caractéristiques (distribution) des personnes recommandées - ou non recommandées, par l'algorithme (sortie du processus), sont similaires à celles en entrée du process, au regard du genre et de l'âge. Autrement dit, pour un poste spécifique, les profils qui seront recommandés par l'algorithme au recruteur présentent des caractéristiques démographiques similaires à l'ensemble des candidats qui ont postulé à ce poste.

Les méthodes de collecte des données, les analyses et les algorithmes proposés par AssessFirst sont donc d'une fiabilité et d'une équité reconnue. Nos équipes mettent en ce sens tout en œuvre pour veiller au caractère équitable des analyses prédictives, et pour s'assurer que l'usage de nos algorithmes dans les processus décisionnels n'amène pas de discriminations via des biais algorithmiques. C'est d'ailleurs ce que démontrent nos études : une équité dans les résultats obtenus au questionnaire (au regard du genre, de l'âge, de l'origine ethnique, et du handicap), et une équité dans les résultats et les recommandations issues de l'algorithme. Les détails concernant ces études sont présentés dans nos manuels techniques et guide d'accessibilité et d'équité, disponibles ici.

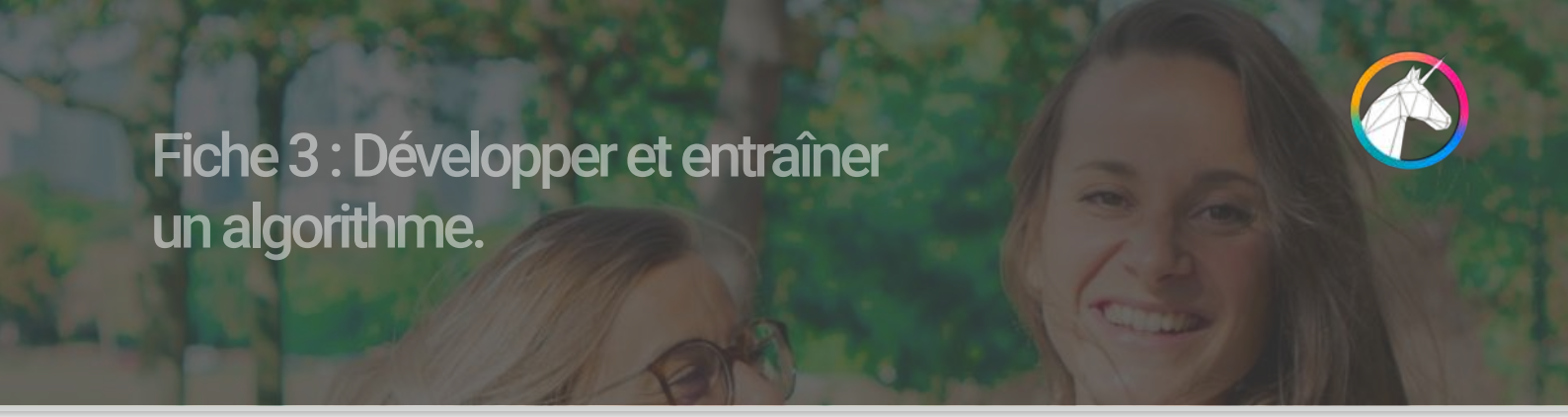
Les données d'entraînement comportent-elles des données liées aux caractéristiques particulières des personnes telles que leur sexe, leur âge, leurs caractéristiques physiques, des données sensibles, etc. ? Lesquelles ?

Les données telles que le sexe, l'âge ou les caractéristiques physiques ne sont pas, actuellement, utilisées par nos systèmes d'IA.

Une étude des biais a-t-elle été effectuée ? Selon quelle méthode ? Si un biais a été identifié, quelles mesures ont été prises pour le réduire ?

Oui, des études régulières sont menées, notamment pour valider l'absence de biais liés au genre dans les recommandations faites par le système. Ces analyses se basent par exemple sur des études d'impact ratio. Dans notre dernière étude, menée sur un échantillon de 208 modèles prédictifs construits pour 18 employeurs, qui ont été testés sur un échantillon mondial de 273 293 candidats potentiels pour chaque poste, des impact ratio pondérés moyens de 0,91 (Femmes-Hommes) et de 0,90 (Hommes-Femmes) ont été observés. Nous avons obtenu des résultats similaires lors de l'analyse d'impact ratio pour 21 catégories d'emplois différentes. Sachant que, selon les standards exigés par l'EEOC (Equal Employment Opportunity Commission), un impact ratio compris entre .8 et 1 est requis pour démontrer l'absence d'adverse impact d'une pratique de sélection, ces conclusions soutiennent l'absence de biais de genre dans nos systèmes.

Impact ratio : fait référence au taux de recommendation ou de sélection, pour une méthode de sélection, d'un groupe de personnes dans une classe protégée, divisé par le taux de sélection du groupe ayant le taux de sélection le plus élevé. Les ratios d'impact sont comparés à la règle des 80 pour cent pour déterminer l'impact défavorable.



Fiche 3 : Développer et entraîner un algorithme.

Mettre en place les bonnes pratiques lors de cette phase cruciale

Concevoir et développer un algorithme fiable

Quel type d'algorithme est utilisé et quel est son principe de fonctionnement (apprentissage supervisé, non-supervisé, continu, fédéré, par renforcement, etc.) ? Pour quelle raison cet algorithme a été choisi ?

Les systèmes d'IA d'AssessFirst actuels se basent principalement sur des modèles de régressions linéaires ou logistiques. Ces régressions font partie des algorithmes d'apprentissage supervisé. En résumé, les régressions linéaires sont utilisées pour modéliser des relations linéaires entre une variable dépendante continue et des variables indépendantes, tandis que les régressions logistiques sont utilisées pour modéliser des relations entre une variable dépendante binaire (ou discrète) et des variables indépendantes. Les deux méthodes d'apprentissage supervisé cherchent à estimer les coefficients de pondération qui minimisent l'erreur ou maximisent la vraisemblance afin de faire des prédictions sur de nouvelles données. Plus en détails :

- Les régressions linéaires sont utilisées pour modéliser une relation linéaire entre une variable dépendante continue et un ensemble de variables indépendantes. L'objectif est de trouver la meilleure ligne droite (ou hyperplan dans le cas de régressions linéaires multiples) qui minimise l'écart entre les valeurs prédites et les valeurs réelles de la variable dépendante. L'algorithme de régression linéaire cherche à ajuster les coefficients de pondération (ou poids) associés à chaque variable indépendante afin de minimiser l'erreur quadratique moyenne entre les valeurs prédites et les valeurs réelles. Cela peut être réalisé à l'aide de méthodes d'optimisation telles que la méthode des moindres carrés ou la descente de gradient ;

- Les régressions logistiques sont utilisées pour modéliser la relation entre une variable dépendante binaire (ou discrète) et un ensemble de variables indépendantes. Elles sont particulièrement adaptées lorsque la variable dépendante est une probabilité ou une catégorie. L'algorithme de régression logistique utilise la fonction logistique (ou sigmoïde) pour transformer une combinaison linéaire des variables indépendantes en une probabilité. Les coefficients de pondération sont estimés à l'aide de méthodes d'optimisation telles que la descente de gradient.

Ces algorithmes ont été choisis en raison de leur (1) simplicité : les régressions linéaires et logistiques sont relativement simples à comprendre et à mettre en œuvre. Elles fournissent une approche intuitive pour modéliser les relations entre les variables ; et leur (2) interprétabilité : les résultats des régressions sont faciles à interpréter. Les coefficients attribués à chaque variable indépendante indiquent la force et la direction de l'impact de cette variable sur la variable dépendante.

L'algorithme utilisé a-t-il été testé par des tiers indépendants ?

À ce jour, l'algorithme utilisé n'a pas encore été soumis à des tests par des tiers indépendants. Cependant, il est important de noter que les régressions sont des modèles largement répandus dans le domaine de l'IA et font partie des algorithmes les plus utilisés. Bien que des tests par des tiers indépendants puissent apporter une validation supplémentaire, l'utilisation de régressions dans notre contexte est conforme aux pratiques courantes et aux normes de l'industrie. Nous restons attentifs aux développements futurs dans le domaine de la validation des algorithmes et nous nous engageons à évaluer régulièrement la performance de notre modèle afin d'assurer sa qualité et son efficacité.

Une revue de la littérature a-t-elle été effectuée ?

Oui, une revue de la littérature a été effectuée. Cette revue a porté sur les données pertinentes à utiliser pour prédire la performance en poste, ainsi que sur les types d'algorithmes appropriés pour modéliser les relations entre les traits de personnalité, les dimensions de motivation et la performance. La revue de la littérature a permis d'identifier des études antérieures qui se sont intéressées à l'utilisation des données liées à la personnalité et aux motivations pour prédire la réussite sur un poste. Ces travaux ont fourni des informations précieuses sur les variables pertinentes à prendre en compte et les méthodes les plus adaptées pour les analyser. En ce qui concerne les algorithmes, la revue de la littérature a examiné différentes approches utilisées dans des contextes similaires. Cela a permis d'identifier les régressions linéaires et logistiques comme des méthodes couramment utilisées pour modéliser les relations entre les variables prises en compte. Cependant, la revue a également révélé que d'autres algorithmes, tels que les réseaux de neurones, les méthodes ensemblistes ou les techniques avancées d'apprentissage automatique, peuvent offrir des avantages dans la modélisation de relations non linéaires ou complexes. À ce titre, il est important de noter qu'AssessFirst s'investit de manière continue dans la recherche et le développement. Ce fort engagement dans la R&D témoigne de notre volonté de rester constamment à jour avec les avancées scientifiques et technologiques. Cela nous permet de suivre les dernières tendances en matière d'algorithmes et de les appliquer à nos données. Cette approche proactive garantit que nos modèles restent à la pointe et nous permet d'innover continuellement.

Un comparatif aux systèmes analogues a-t-il été réalisé ?

Un comparatif aux systèmes analogues est en cours de réalisation.

Des outils tiers sont-ils utilisés ?

Non.

Sont-ils réputés fiables et éprouvés ?

Aucun outil tiers n'est utilisé.

Une recherche des possibles défauts dans la documentation des outils et parmi la communauté de développeurs a-t-elle montré des points d'intérêts ?

Aucun outil tiers n'est utilisé.

Un suivi des mises à jour des outils est-il prévu ?

Aucun outil tiers n'est utilisé.

L'algorithme d'IA est-il disponible en open source ?

En raison de considérations commerciales, les algorithmes utilisés dans nos systèmes d'IA ne sont pas accessibles en open source. Cette décision a été prise pour protéger la propriété intellectuelle et préserver notre avantage concurrentiel. Cependant, nous accordons une grande importance à la transparence et à la confiance de nos clients, c'est pourquoi nous nous engageons à fournir des informations détaillées sur notre approche méthodologique et les principes sur lesquels reposent nos modèles d'IA.

Les algorithmes ont-ils été suffisamment testés par des tiers ?

Les algorithmes des systèmes d'IA d'AssessFirst sont mis à l'épreuve quotidiennement par des milliers d'utilisateurs dans le monde entier. Cette utilisation intensive de notre solution nous offre une opportunité précieuse de tester nos algorithmes dans des conditions réelles, en couvrant une large gamme de cas d'utilisation. Grâce à cette diversité de situations concrètes, nous sommes en mesure d'analyser et d'étudier attentivement les résultats générés par nos modèles d'IA, afin de garantir leur pertinence et la qualité des algorithmes utilisés. Cette approche basée sur des cas réels nous permet d'améliorer en permanence nos systèmes d'IA et d'assurer une performance optimale pour nos utilisateurs.

Les algorithmes correspondent-ils à l'état de l'art ?

L'état de l'art en matière d'algorithmes évolue constamment avec les progrès de la recherche. Les régressions linéaires et logistiques sont des algorithmes classiques et bien établis, mais ils ne représentent pas nécessairement l'état de l'art actuel dans tous les domaines. Le choix de l'algorithme dépend avant tout du problème spécifique que l'on essaie de résoudre et des performances recherchées. Au sens que les régressions linéaires et logistiques fournissent des résultats satisfaisants en termes de précision et d'efficacité dans notre cadre d'étude précise (i.e., prédire la performance ou le turnover à partir de données liées à la personnalité, aux motivations et au raisonnement), il est ainsi admis d'affirmer que nos algorithmes reposent sur des méthodes éprouvées et bien connues. Cependant, il est important de noter que l'état de l'art peut évoluer, et il existe peut-être des approches plus avancées ou des algorithmes plus performants qui pourraient être explorés pour améliorer encore les résultats. C'est en ce sens que notre équipe Science suit les développements récents dans le domaine de l'apprentissage automatique pour rester à jour sur l'état de l'art des algorithmes.

Les retours de la communauté à son sujet sont-ils encouragés et étudiés ?

Tous les retours sont soigneusement étudiés par AssessFirst. Nous attachons une grande importance à la recherche de feedback sur nos systèmes actuels. Dans cette optique, nous avons mis en place un comité scientifique en 2019, composé de chercheurs renommés et indépendants issus du monde académique et de la recherche universitaire. Ce comité joue un rôle essentiel en fournissant un regard externe et objectif sur nos systèmes d'IA. Leurs expertises nous permettent d'améliorer continuellement nos approches et de prendre en compte les dernières avancées scientifiques dans le domaine.

Mettre en place les bonnes pratiques lors de cette phase cruciale

Mener un protocole d'entraînement méticuleux

Quelles sont les stratégies d'apprentissage mises en œuvre ?

Les stratégies d'apprentissage actuellement utilisées sont principalement supervisées. Dans cette stratégie, un modèle est entraîné à partir de données étiquetées, c'est-à-dire des exemples d'entrées associés à des sorties attendues. Le modèle apprend à prédire les sorties correspondantes pour de nouvelles données non étiquetées. Les régressions linéaires et logistiques utilisées sont des exemples d'algorithmes d'apprentissage supervisé.

Quelle répartition a été utilisée entre ensembles d'entraînement, de test, de validation ?

Non applicable.

Des stratégies telles que par exemple la validation croisée sont-elles utilisées ?

Non applicable.

La qualité des sorties du système d'IA est-elle jugée suffisante ?

Oui, la qualité est jugée suffisante et mesurée à travers les métriques suivantes :

- Accuracy : Mesure la proportion de prédictions correctes par rapport à toutes les prédictions effectuées. Il s'agit donc de la capacité à prédire correctement les observations positives et négatives. Elle se calcule en divisant le nombre total de prédictions correctes par le nombre total de prédictions effectuées. L' accuracy varie de 0 à 1 ;

- Recall : Mesure la proportion de vrais exemples positifs qui sont correctement prédits parmi tous les exemples positifs. En d'autres termes, le recall mesure la capacité à trouver toutes les observations positives. Il se calcule en divisant le nombre de prédictions positives correctes par le nombre total d'exemples positifs. Le recall varie de 0 à 1 ;

- Précision : Mesure la proportion de prédictions positives qui sont correctes parmi toutes les prédictions positives. En d'autres termes, la précision mesure notre capacité à ne pas classer à tort une observation négative comme positive. Elle se calcule en divisant le nombre de prédictions positives correctes par le nombre total de prédictions positives. La précision varie de 0 à 1 ;

- F1-score : Mesure combinée de la précision et du recall. Il s'agit de la moyenne harmonique de la précision et du recall. Le F1-score peut être considéré comme l'indicateur global d'efficacité. Le F1-score varie de 0 à 1, où une valeur de 1 indique une performance optimale en termes de précision et de recall ;

- ROC AUC : La courbe ROC (Receiver Operating Characteristic) et l'aire sous la courbe ROC (ROC AUC) sont utilisées pour évaluer les performances d'un modèle de classification binaire. La courbe ROC représente le taux de vrais positifs en fonction du taux de faux positifs, en faisant varier le seuil de décision du modèle. L'aire sous la courbe ROC (ROC AUC) mesure la capacité de discrimination du modèle, c'est-à-dire sa capacité à classer correctement les exemples positifs et négatifs ;

- Pouvoir prédictif : Le pouvoir prédictif fait référence à la capacité d'un modèle à faire des prédictions précises sur de nouvelles données non vues auparavant. Il mesure la qualité globale des prédictions du modèle ;

- Impact ratio : Mesure de l'équité, qui correspond au taux de recommandation pour les femmes divisé par le taux de recommandation pour les hommes. Le taux de recommandation a été utilisé plutôt que le taux de sélection, car l'algorithme ne prend pas de décisions de manière autonome, mais est uniquement utilisé comme recommandation par les employeurs. Conformément aux directives de l'Equal Employment Opportunity Commission (ou EEOC), un modèle prédictif avec un impact ratio compris entre 0,8 et 1 serait considéré comme équitable. Les modèles prédictifs dont le taux d'impact est inférieur au seuil de 0,8 indiquent un impact discriminatoire.

Ces métriques permettent d'évaluer différentes facettes de la performance d'un modèle d'intelligence artificielle et aident à comprendre son comportement.

Permettent-elles de mesurer la performance du système de manière satisfaisante et en prenant en compte les conséquences pour les personnes étudiées ?

Oui, ces métriques permettent de mesurer la performance du système de manière satisfaisante et en prenant en compte les conséquences pour les personnes étudiées.

Les cas d'erreur ont-ils été étudiés ?

Oui, les cas d'erreurs sont constamment étudiés.

Les situations limites où les sorties du système ne sont pas suffisamment fiables sont-elles clairement identifiées ?

Oui, nous identifions clairement les situations où les sorties du système peuvent ne pas être suffisamment fiables grâce à nos métriques précédentes. Nous accordons une grande importance à l'explicabilité du modèle et de ses résultats, c'est pourquoi nous informons directement les utilisateurs "clients" des limitations du système via des informations spécifiques affichées dans l'application. Cette approche transparente permet aux utilisateurs de prendre des décisions éclairées en comprenant les niveaux de confiance associés aux résultats fournis par le système d'IA.

Une sécurité est-elle ajoutée pour prendre en charge ces cas (en rendant systématiquement la main à un opérateur humain par exemple ?)

Quelles que soient les performances du modèle, un opérateur humain (le "client") conserve toujours le contrôle pour corriger les résultats. Nous reconnaissons l'importance de l'intervention humaine dans le processus décisionnel et nous donnons aux clients la possibilité de modifier ou d'ajuster les résultats du modèle selon leur expertise. Toutefois, nous ne recommandons pas cette approche, au sens qu'elle peut amener l'introduction de biais. Dans le cas où le client souhaite modifier le modèle généré par l'IA, ce modèle deviendra ainsi un « modèle personnalisé ».

Cas de l'apprentissage continu : des mesures sont-elles prises en amont pour éviter une dégradation de la performance, une dérive du modèle ou une attaque visant à orienter les résultats de l'algorithme (par exemple des chatbots en ligne devenus "racistes") ?

Nous mettons en place plusieurs mesures pour améliorer notre approche :

- Sélection des données : Nous veillons à ce que nos clients soient formés et encadrés pour garantir que les modèles soient créés à partir de données non biaisées ;

- Vérification humaine : Les utilisateurs ont accès aux résultats du modèle et ont la possibilité de les corriger si nécessaire. Nous encourageons ainsi une collaboration entre le modèle et l'expertise humaine pour assurer des prédictions de qualité ;

- Évaluation de la performance du modèle : En accord avec nos clients, nous effectuons régulièrement des études de validité prédictive pour évaluer la performance des modèles. Cela nous permet de démontrer leur efficacité et d'apporter d'éventuelles améliorations si nécessaire.

Mettre en place les bonnes pratiques lors de cette phase cruciale

Vérifier la qualité du système en environnement contrôlé

Le traitement a-t-il été validé par une expérimentation ?

Oui, les systèmes d'IA d'AssessFirst font l'objet de tests et d'expérimentations régulières pour évaluer leur performance et leur validité prédictive dans des environnements réels. Nous effectuons des analyses régulières pour mesurer l'efficacité et la précision des modèles générés par nos systèmes, en utilisant des données réelles provenant de situations concrètes. Cette approche nous permet de valider et d'améliorer continuellement nos algorithmes, en nous assurant qu'ils fournissent des résultats fiables et pertinents.

Le contexte dans lequel évolue le système d'IA (quantité de variables à considérer, difficulté à évaluer la représentativité des données, etc.) est-il particulièrement complexe ?

Un environnement complexe dans le contexte des systèmes d'IA se caractérise par :

- **Quantité de variables** : Un environnement complexe peut impliquer un grand nombre de variables à prendre en compte dans le processus de prise de décision. Ces variables peuvent interagir de manière non linéaire et présenter des relations complexes, ce qui rend l'analyse et la modélisation plus difficiles ;

- **Interactions dynamiques** : Les environnements complexes peuvent être dynamiques, c'est-à-dire que les conditions et les relations entre les variables peuvent évoluer dans le temps. Les systèmes d'IA doivent être capables de s'adapter à ces changements et de prendre des décisions en temps réel en fonction de l'état actuel de l'environnement ;

- **Incertitude et bruit** : Dans un environnement complexe, il peut y avoir des sources d'incertitude et de bruit qui rendent la tâche de modélisation et de prédiction plus difficile. Les données peuvent être incomplètes, bruitées ou manquer de fiabilité, ce qui nécessite des techniques avancées pour gérer cette incertitude ;

- **Relations non linéaires** : Les environnements complexes peuvent présenter des relations non linéaires entre les variables, ce qui signifie que les effets d'une variable sur une autre ne sont pas simplement proportionnels. Cela nécessite l'utilisation de modèles et d'algorithmes capables de capturer ces relations non linéaires.

En ce sens, le contexte dans lequel évolue le système d'IA AssessFirst peut être qualifié de moyennement complexe. En effet, il se caractérise par moins de variables, des relations linéaires ou simples à modéliser, et une certaine stabilité temporelle des variables mesurées. Dans un tel environnement, les tâches d'analyse, de modélisation et de prise de décision peuvent être plus directes et moins complexes. Toutefois, la nature des variables prises en compte apportent davantage de complexité et d'incertitude aux modèles.

L'algorithme en question a-t-il déjà été utilisé dans un contexte similaire ?

Oui, l'algorithme est constamment utilisé dans ce type d'environnement.

Dans quel environnement a été menée l'expérimentation ?

Nos expérimentations sont réalisées avec certains clients cibles et volontaires, sur des cas d'usage réels, non-contrôlés et en environnement ouvert.

Les conditions se rapprochent-elles suffisamment des conditions réelles ?

Toutes nos expérimentations ou études sont conduites en environnement réel (ex. données collectées à partir de l'environnement réel, cas d'usage client, etc.).

Les précautions appropriées ont-elles été prises, comme un contrôle systématique par un opérateur humain qui pourrait être ensuite réduit lors de la phase de production ?

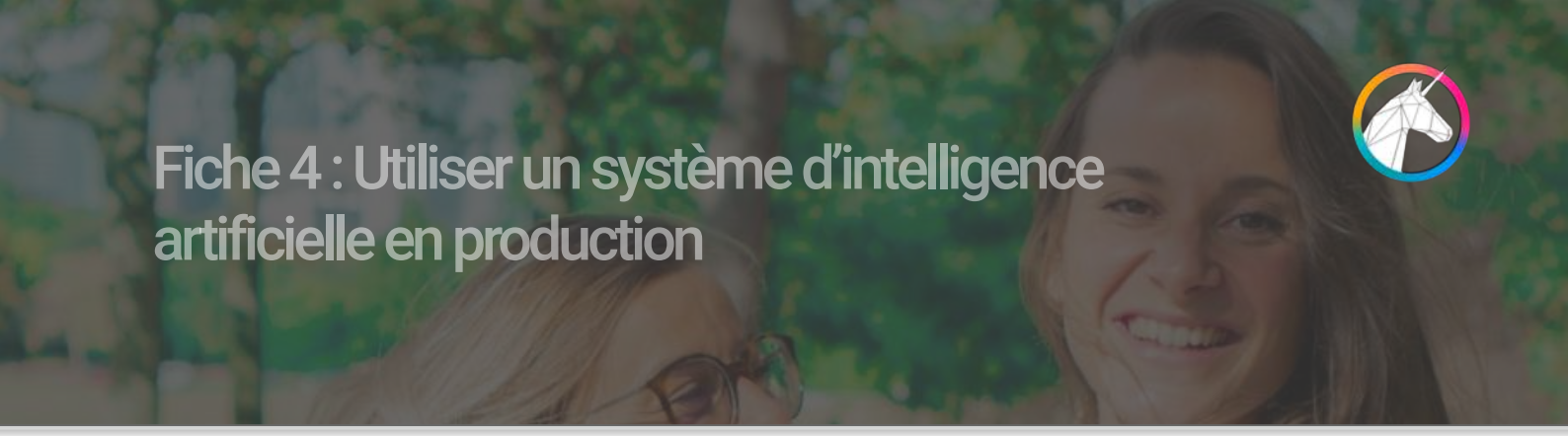
L'ensemble des résultats fournis par le système d'IA (dans le cadre de la génération de modèles prédictifs) fait systématiquement l'objet d'un contrôle par un opérateur humain. Cet opérateur peut, s'il le juge nécessaire, apporter des modifications au modèle généré par l'IA.

Les métriques utilisées pour valider l'expérimentation sont-elles adaptées et suffisantes ?

Les métriques utilisées correspondent à celles déjà présentées, et sont jugées adaptées et suffisantes pour nos cas d'études actuels.

Un bilan exhaustif et objectif de l'expérimentation a-t-il été réalisé ?

Oui, chaque étude est documentée.



Fiche 4 : Utiliser un système d'intelligence artificielle en production

Garantir la qualité et la transparence du système au cours de son utilisation

Permettre à l'humain de garder la main

Une supervision par un opérateur humain est-elle prévue ?

Tous les résultats produits par le système d'IA dans le cadre de la génération de modèles prédictifs font l'objet d'un contrôle par un opérateur humain. Cet opérateur a le pouvoir, s'il le juge nécessaire, d'apporter des modifications au modèle généré par l'IA. Chaque modèle prédictif est examiné attentivement par un opérateur qualifié, qui évalue sa performance, son adéquation aux besoins spécifiques et sa conformité aux exigences. Si l'opérateur identifie des aspects à améliorer ou des ajustements nécessaires, il est en mesure d'apporter des modifications au modèle généré par l'IA. Cette approche permet de combiner les avantages de l'IA avec l'expertise et la prise de décision humaine, assurant ainsi des résultats fiables.

Des mécanismes ont-ils été prévus pour que l'opérateur puisse modifier manuellement une décision prise par le système d'IA (par ex. : modifier le profil donné à l'utilisateur d'une plateforme de partage de vidéos) ou en arrêter le fonctionnement (par ex. couper l'accès d'une chatbot en ligne dont l'apprentissage automatique aurait conduit à une dérive ?)

Tous les systèmes d'IA d'AssessFirst, ainsi que les résultats qu'ils produisent, peuvent être modifiés manuellement par un opérateur humain. Cette possibilité de modification s'applique aussi bien (1) côté « client », qui peuvent faire des ajustements dans les modèles générés par l'IA, que (2) côté AssessFirst, qui garde la plein de maîtrise de chaque étape du process, jusqu'à interruption totale du fonctionnement du système si nécessaire. Nous accordons une grande importance à la flexibilité et au contrôle offerts aux opérateurs humains. Cette approche permet de garantir une supervision et une responsabilité humaine dans l'utilisation des systèmes d'IA. Nous reconnaissons l'importance de l'intervention humaine pour assurer la pertinence, la fiabilité et l'éthique des résultats produits par nos systèmes d'IA.

Quel type d'intervention est prévu ?

Plusieurs types d'intervention sont prévus :

- Human-in-the loop : capacité d'intervention humaine dans chaque cycle de décision du système ;
- Human-on-the loop : capacité d'intervention humaine pendant le cycle de conception du système et la surveillance du fonctionnement du système ;
- Human-in-command : capacité de superviser l'activité globale du système d'IA et de décider quand et comment utiliser le système d'IA dans une situation donnée.

Ces mécanismes font-ils l'objet de protocoles clairement formulés et connus de tous ? Sont-ils intégrés de manière naturelle dans le traitement ? Les ressources humaines et matérielles nécessaires ont-elles été prévues ?

Ces mécanismes d'intervention sont intégrés de manière transparente à l'application, offrant ainsi une expérience fluide et intuitive pour les utilisateurs. Dans cette optique, nous accordons aux utilisateurs "clients" une entière liberté pour décider quand et comment utiliser le système d'IA dans une situation donnée (Human-in-command) ou pour modifier les décisions prises par le système (Human-in-the-loop). De plus, du côté de la conception, nos équipes Science et Tech disposent de pleins pouvoirs pour apporter des modifications manuelles aux décisions prises par le système d'IA ou à son fonctionnement. Nous veillons à ce que les protocoles soient définis et, si nécessaire, nous prévoyons les ressources matérielles, humaines et techniques nécessaires pour soutenir ces modifications.

Les personnes en charge de cette tâche sont-elles clairement identifiées ?

Oui, les personnes en charge de cette tâche sont clairement identifiées, et concernent des membres cibles de l'équipe Science d'AssessFirst, ou de l'équipe de développement Back.

Sont-elles en mesure d'exercer leur contrôle sur le système d'IA facilement ?

Oui, elles sont en mesure d'exercer leur contrôle sur le système d'IA facilement.

Les personnes ont-elles reçu une formation suffisante ?

Les individus responsables du développement des systèmes d'IA chez AssessFirst sont clairement identifiés et ont tous suivi une formation adéquate et essentielle pour garantir un développement efficace et éthique de nos systèmes. Leur formation comprend des domaines tels que les sciences comportementales, la science des données, l'ingénierie des données, l'intelligence artificielle et l'apprentissage automatique. Nous accordons une grande importance à l'expertise et aux compétences de notre équipe de développement. Leur formation multidisciplinaire leur permet de comprendre les aspects comportementaux, les défis techniques et les implications éthiques liés à nos systèmes d'IA. Cette approche holistique assure un développement de qualité et une prise en compte appropriée des enjeux complexes associés à l'intelligence artificielle.

L'information qui leur est fournie lors de la supervision est-elle suffisante ?

Les informations fournies lors de la supervision sont adéquates. De plus, ces personnes ont un contrôle total sur les données à collecter et possèdent les compétences nécessaires pour effectuer des analyses complémentaires si besoin. Nous nous assurons que les superviseurs disposent de toutes les informations requises pour prendre des décisions éclairées. Leur expertise leur permet d'identifier les éléments pertinents à considérer et de mener des analyses supplémentaires, le cas échéant, afin d'approfondir la compréhension du problème. Nous reconnaissons l'importance de leur rôle dans le processus de supervision, et nous leur faisons confiance pour prendre des décisions judicieuses et garantir la qualité des résultats. Leur expertise et leur capacité à explorer et à analyser les données de manière approfondie contribuent à renforcer la fiabilité et la pertinence des modèles générés par nos systèmes d'IA.

Une identification des cas où l'intervention humaine est nécessaire a-t-elle été mise en place (via le calcul d'un indicateur de confiance par exemple) ?

Nous utilisons plusieurs métriques pour évaluer la qualité des modèles générés par nos systèmes d'IA (comme mentionné dans la fiche précédente). Cependant, il n'y a pas de seuil prédéterminé pour déclencher une intervention humaine. Nous maintenons la possibilité d'une intervention humaine indépendamment des résultats produits par le système d'IA. Nous reconnaissons l'importance de la supervision et de l'expertise humaine dans le processus décisionnel, et nous offrons la flexibilité nécessaire pour permettre une intervention humaine à tout moment. Cela garantit que les décisions prises reposent sur une combinaison d'intelligence artificielle et de discernement humain, offrant ainsi une approche équilibrée et responsable.

Une étude de la proportion de cas soumis à l'humain et de son efficacité est-elle prévue ?

Non, ce genre d'étude n'a pas été réalisé.

Comment le risque lié au biais d'automatisation (tendance d'un opérateur à accorder une confiance trop importante à un procédé automatisé) a-t-il été pris en compte ?

Pour prendre en compte le risque lié au biais d'automatisation et à la tendance d'un opérateur à accorder une confiance excessive à un processus automatisé, plusieurs mesures sont prises :

- Sensibilisation et formation : Les "clients" sont informés et sensibilisés sur les biais potentiels associés à l'automatisation. Ils sont formés pour comprendre les limites et les préjugés éventuels des systèmes d'IA, afin d'éviter une confiance aveugle et d'adopter une approche critique dans l'interprétation des résultats ;

- Surveillance et évaluation : Un suivi régulier est effectué pour évaluer la performance des systèmes d'IA et détecter tout biais ou distorsion dans les résultats. Des indicateurs de performance sont définis et surveillés pour identifier les cas où une intervention humaine supplémentaire peut être nécessaire ;

- Mécanismes de relecture et de validation : Des mécanismes sont mis en place pour permettre la relecture et la validation des résultats générés par l'automatisation. Les opérateurs ont la possibilité de vérifier les décisions prises par les systèmes d'IA et de les remettre en question si nécessaire, en s'appuyant sur leur expertise ;

- Transparence et explication des résultats : Les systèmes d'IA sont conçus de manière à fournir des explications et des justifications claires sur les décisions prises. Cela permet aux "clients" de comprendre comment les résultats ont été obtenus et de détecter tout biais éventuel. La transparence favorise une évaluation critique et une prise de décision éclairée.

En combinant ces différentes mesures, nous visons à atténuer le risque de biais d'automatisation en favorisant une approche équilibrée entre l'automatisation et la supervision humaine. Nous reconnaissons que la confiance aveugle envers les systèmes automatisés peut entraîner des limites, c'est pourquoi nous promouvons une approche critique et réflexive dans l'utilisation de nos systèmes d'IA.

Garantir la qualité et la transparence du système au cours de son utilisation

Mettre en oeuvre la transparence pour assurer la confiance

Une journalisation des éléments (données utilisées pour l'inférence, indicateurs de confiance, version du système, etc.) servant à la prise de décision par le système d'IA existe-t-elle ?

Nous effectuons une journalisation de certains éléments clés, tels que les données d'entrée et de sortie, les résultats générés par le modèle, le pouvoir prédictif du modèle généré, ainsi que la rareté des profils étudiés. Cette journalisation nous permet de suivre et d'analyser en détail le fonctionnement de nos systèmes d'IA, ainsi que leurs performances. Elle joue un rôle essentiel dans la transparence et la traçabilité de nos processus, en fournissant une documentation précise des étapes et des résultats obtenus. Cette pratique nous permet d'évaluer l'efficacité de nos modèles, d'identifier d'éventuelles améliorations à apporter et de garantir la fiabilité de nos systèmes d'IA.

Les éléments journalisés permettent-ils d'expliquer, a posteriori, une décision particulière prise par le système d'IA ?

Oui, les éléments journalisés jouent un rôle important dans l'explicabilité des décisions prises par notre système d'IA. En enregistrant les données d'entrée et de sortie, ainsi que les étapes de traitement effectuées par le modèle, il est possible de retracer le cheminement qui a conduit à une décision spécifique, et de reproduire l'ensemble des analyses et décisions. Cela permet de mieux comprendre les facteurs pris en compte par le système et d'identifier les raisons qui ont influencé la décision finale. La journalisation facilite également l'analyse rétrospective des résultats. En résumé, la journalisation des éléments pertinents fournit une documentation essentielle pour expliquer et évaluer les décisions prises par le système d'IA.

Ces informations sont-elles conservées pour une durée justifiée ?

Ces informations sont conservées tant que le modèle prédictif est actif sur le compte client.

Sont-elles limitées aux strictes catégories nécessaires à l'explication de la décision ?

Elles sont limitées aux catégories nécessaires pour pouvoir ré-effectuer, a posteriori, l'ensemble des analyses et comprendre les décisions du système.

Le fonctionnement du système est-il expliqué aux personnes amenées à interagir avec ?

Oui, AssessFirst s'engage à fournir une explication claire et compréhensible du fonctionnement de son système d'IA aux personnes amenées à interagir avec lui. Ces explications pour la plupart présentées lors de la formation des clients, en amont de l'utilisation. L'objectif est de garantir une transparence et une compréhension adéquate de la manière dont le système fonctionne, afin que les utilisateurs puissent interagir en toute confiance. Pour atteindre cet objectif, AssessFirst met à disposition des informations détaillées sur le système d'IA, notamment à travers des documents explicatifs, des guides d'utilisation et des formations appropriées. Ces ressources sont conçues pour expliquer les principes fondamentaux de l'IA, les méthodes utilisées, les données prises en compte et les résultats attendus. En fournissant ces explications, AssessFirst vise à promouvoir une utilisation éclairée de son système d'IA et à permettre aux utilisateurs de comprendre les limites et les implications du système. Il s'agit d'une démarche visant à favoriser la confiance, l'engagement et la collaboration efficace entre les utilisateurs et le système d'IA.

L'explication est-elle rendue suffisamment claire et compréhensible, grâce à des cas concrets par exemple, et en expliquant les limitations, hypothèses et cas extrêmes du système ? (ex. un robot industriel assistant un humain dans sa tâche saura perforer mais pas visser)

Oui, dans le souci de rendre l'explication du système d'IA claire et compréhensible, AssessFirst met en place diverses stratégies pour faciliter la compréhension des utilisateurs. Cela comprend l'utilisation de cas concrets, d'exemples et de scénarios pratiques pour illustrer le fonctionnement du système. L'explication du système d'IA comprend également une discussion sur les limitations, les hypothèses sous-jacentes et les cas extrêmes. AssessFirst met l'accent sur la transparence en identifiant et en communiquant clairement les limites du système, les conditions dans lesquelles il peut ne pas fournir des résultats optimaux, ainsi que les hypothèses sur lesquelles il se base. En fournissant ces informations, AssessFirst vise à permettre aux utilisateurs de comprendre les performances attendues du système dans différents contextes, d'anticiper les situations où le système peut présenter des difficultés et de prendre des décisions éclairées en fonction de ces informations. Dans l'ensemble, l'objectif est de rendre l'explication du système d'IA aussi accessible et transparente que possible, en fournissant des explications claires, des exemples concrets et en soulignant les limitations et les cas extrêmes.

Des outils techniques et méthodologiques sont-ils utilisés pour permettre l'explicabilité du système ?

Oui, pour garantir l'explicabilité du système, AssessFirst utilise des méthodologies spécifiques. Les résultats générés par les systèmes d'IA sont accompagnés d'explications détaillées, fournies sous forme de visuels ou de texte, afin de permettre aux utilisateurs de comprendre comment ces résultats ont été obtenus. L'utilisation de visuels permet de présenter de manière visuelle les éléments clés qui ont influencé la décision du système, tels que les variables qui ont un poid dans le modèle, les principaux motifs ou schémas détectés par l'algorithme, ou encore le positionnement de chaque membre de l'échantillon au regard du modèle généré. Ces actions permettent aux utilisateurs d'avoir une vision claire et détaillée du fonctionnement du système d'IA, en leur fournissant les informations nécessaires pour comprendre les résultats obtenus et prendre des décisions éclairées.

Le code est-il ouvert (open source) ?

En raison de considérations commerciales, les algorithmes utilisés dans nos systèmes d'IA ne sont pas accessibles en open source. Cette décision a été prise pour protéger la propriété intellectuelle et préserver notre avantage concurrentiel. Cependant, nous accordons une grande importance à la transparence et à la confiance de nos clients, c'est pourquoi nous nous engageons à fournir des informations détaillées sur notre approche méthodologique et les principes sur lesquels reposent nos modèles d'IA.

Garantir la qualité et la transparence du système au cours de son utilisation

Assurer la qualité du traitement

Une analyse automatique des logs de journalisation existe-t-elle afin d'alerter l'utilisateur et/ou la personne en cas de défaillance, de fonctionnement anormal ou critique ?

Nous avons mis en place un système de logs qui enregistre les activités. Cela nous permet de remonter à l'origine d'un problème en cas de dysfonctionnement. Les logs nous fournissent des informations détaillées sur les opérations effectuées, et facilitent l'identification et la résolution des problèmes.

Un contrôle de la qualité et de la correspondance des données collectées en environnement réel avec les données d'apprentissage et de validation est-il maintenu au cours de l'utilisation du système d'IA ?

Non applicable.

La qualité des sorties du système d'IA est-elle contrôlée au cours du cycle de vie ?

Oui, la qualité des sorties du système d'IA est constamment contrôlée tout au long de son cycle de vie. AssessFirst met en place des processus et des mécanismes de contrôle pour évaluer la performance et la précision des sorties du système, afin de garantir leur qualité. Des évaluations régulières sont effectuées pour mesurer la performance du système d'IA. Cela permet de vérifier si les sorties du système sont cohérentes, précises et fiables. Des analyses statistiques et des mesures de performance sont utilisées pour évaluer la qualité des sorties du système, en tenant compte de différentes métriques telles que la précision, le rappel, la F-score, etc. Ces mesures permettent d'identifier d'éventuels problèmes ou biais dans les sorties du système et de prendre les mesures nécessaires pour les corriger. De plus, des retours et des feedbacks des utilisateurs sont également pris en compte pour évaluer la qualité des sorties du système. Les retours des utilisateurs, qu'ils soient positifs ou négatifs, sont analysés et utilisés pour améliorer les performances du système et résoudre d'éventuels problèmes. En résumé, AssessFirst accorde une grande importance à la qualité des sorties du système d'IA et met en œuvre des contrôles réguliers pour assurer leur fiabilité et leur précision tout au long du cycle de vie du système.

Garantir la qualité et la transparence du système au cours de son utilisation

Les risques spécifiques

Dans le cas d'un apprentissage continu, la qualité des données utilisées pour l'entraînement est-elle contrôlée tout au long du cycle de vie du système ?

Oui, dans le cas d'un apprentissage continu, la qualité des données utilisées pour l'entraînement du système d'IA est constamment contrôlée tout au long de son cycle de vie. AssessFirst met en place des mécanismes de contrôle et de suivi pour s'assurer de la qualité des données utilisées et de leur pertinence pour l'apprentissage continu du système. Notamment les données collectées de manière continue sont soumises aux mêmes normes et procédures que les données initiales. Elles sont recueillies au moyen de tests psychométriques dont la validité, la fidélité, la sensibilité et l'équité ont été reconnues et démontrées. Pour obtenir des informations détaillées, nous vous recommandons de consulter notre manuel psychométrique qui fournit des informations supplémentaires sur ces tests et leur utilisation. Ce manuel est une ressource précieuse pour comprendre les principes et les standards appliqués dans la collecte continue des données.

Dans le cas d'un apprentissage fédéré, des mesures sont-elles prises pour lutter contre le caractère non-indépendant et identiquement distribué des données ? (par ex. : apprentissage fédéré d'un modèle de reconnaissance vocale sur des assistants vocaux).

Le traitement effectué ne repose pas sur un apprentissage fédéré.

Dans le cas où un système d'IA est utilisé pour collecter des données personnelles par exploration des données ou data mining, une vérification permet-elle de s'assurer que seules les données des personnes concernées sont collectées ?

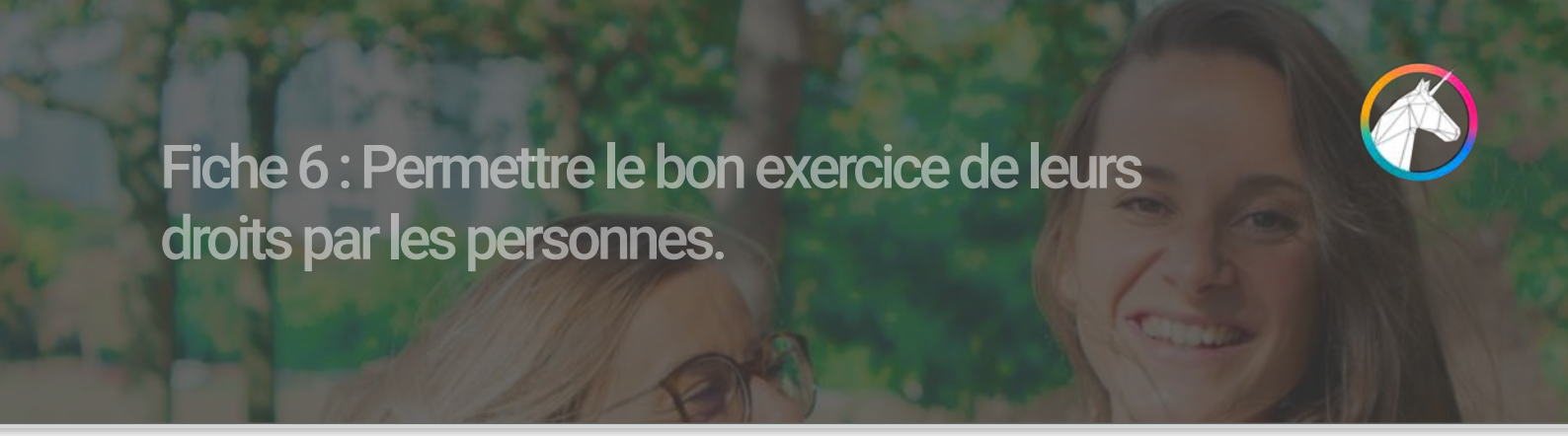
Non applicable.

Une vérification permet-elle de s'assurer de leur intégrité et de leur véracité (par la vérification de l'authenticité des sources, et de la qualité de l'extraction des données par exemple) ?

AssessFirst garantit le respect de la vie privée de ses utilisateurs en n'utilisant pas de techniques de data mining lors de la collecte des données personnelles. Les informations sont exclusivement fournies par les utilisateurs eux-mêmes, directement dans l'application AssessFirst.

Dans le cas d'un algorithme de recommandation de contenu, les suggestions pourraient-elles influencer les opinions des personnes (politiques par exemple) ? Pourraient-elles aller à l'encontre des intérêts de certaines personnes ?

Nos algorithmes de recommandation ne sont pas intrinsèquement conçus pour manipuler les opinions des individus. Leur objectif principal est de fournir des suggestions pertinentes et adaptées aux intérêts et préférences des utilisateurs. Aussi le contenu de ces suggestions ne va pas à l'encontre des intérêts de certaines personnes, physiques ou morales.



Fiche 6 : Permettre le bon exercice de leurs droits par les personnes.

Promouvoir la transparence et les droits de l'utilisateur final

Mesurer l'impact du traitement sur les droits fondamentaux des personnes

Les conséquences du traitement sur les droits fondamentaux des personnes (droits à la liberté d'expression, à la liberté de pensée, de conscience et de religion, de circuler librement, au respect de sa vie privée et familiale, etc.) ont-elles été prises en compte ?

Oui, les conséquences du traitement sur les droits fondamentaux des personnes sont prises en compte par AssessFirst. AssessFirst attache une grande importance à la protection des droits tels que la liberté d'expression, la liberté de pensée, de conscience et de religion, le respect de la vie privée et familiale, ainsi que d'autres droits fondamentaux. Lors du développement et de l'utilisation de ses systèmes d'IA, AssessFirst veille à ce que les droits des personnes soient respectés et que les conséquences potentielles soient évaluées. Cette démarche éthique est la base de nos décisions et de nos pratiques.

Les conséquences sur les groupes de personnes fondés sur la base du genre, de l'origine, ou encore des opinions politiques ou religieuses, sur la société et la démocratie en général ont-elles été considérées ?

Oui, AssessFirst reconnaît l'importance de prendre en compte les conséquences sur les groupes de personnes fondées sur le genre, l'origine, les opinions politiques ou religieuses, ainsi que d'autres caractéristiques protégées. Il est essentiel que les conséquences sur les groupes de personnes et sur la société dans son ensemble soient prises en compte afin de garantir l'égalité des chances, le respect des droits fondamentaux et la préservation de la démocratie. AssessFirst s'engage à promouvoir des pratiques responsables et éthiques dans l'utilisation de ses systèmes d'IA, afin de minimiser les risques de discrimination ou d'influence préjudiciable sur la société et la démocratie. Aussi, AssessFirst produit très régulièrement des études sur le sujet, afin de démontrer, notamment, que les résultats proposés ne sont pas impactés par le genre, l'âge, l'origine ou le handicap. Ces actions permettent de neutraliser les biais, comme le démontre nos dernière études, menées en Juillet et Décembre 2022 :

- Que cela soit en termes de genre, de catégories d'âge, d'origine ethnique, ou de handicap, il n'existe pas de différences majeures entre les résultats obtenus par les différents groupes aux questionnaires SWIPE, DRIVE et BRAIN d'AssessFirst. En somme, les résultats obtenus permettent de soutenir l'hypothèse selon laquelle les questionnaires proposés ne discriminent aucun public, et s'avèrent équitables. Les effets les plus marqués sont relatifs aux handicaps (notamment handicaps mentaux et cognitifs) - même si les effets mentionnés sont très faibles ou faibles, et explicables par d'autres interactions. De même, les effets les plus marqués concernent le genre et uniquement quelques facettes. Aussi, ces effets mentionnés sont très faibles ou faibles, potentiellement explicables par d'autres interactions ou échantillonnages, et trouvent tous une justification conceptuelle ;

- L'usage des algorithmes proposés par AssessFirst s'avère non-discriminant et équitable : les caractéristiques (distribution) des personnes recommandées - ou non recommandées, par l'algorithme (sortie du process), sont similaires à celles en entrée du process, au regard du genre et de l'âge. Autrement dit, pour un poste spécifique, les profils qui seront recommandés par l'algorithme au recruteur présentent des caractéristiques démographiques similaires à l'ensemble des candidats qui ont postulé à ce poste.

Un cadre formel a-t-il été utilisé afin de réaliser cette étude d'impact (AIPD, autre grille d'analyse) ?

AssessFirst a effectué une AIPD globale, mise à jour en mars 2023. Toutefois, AssessFirst n'a pas conduit d'AIPD unique dans le cadre du système d'IA.

Promouvoir la transparence et les droits de l'utilisateur final

Informer les personnes pour leur rendre contrôle

Les personnes ont-elles été informées du traitement de manière claire et concise ?

AssessFirst a mis en place une politique de protection des données personnelles, qui englobe des sujets essentiels tels que la confidentialité des données, la transparence, la responsabilité et la protection de la vie privée. Cette politique sert de référence pour toutes nos pratiques en matière de gestion des données, garantissant ainsi un engagement fort envers la protection des informations sensibles de nos utilisateurs.

Les informations sont-elles aisément accessibles ?

Oui, AssessFirst s'efforce de rendre les informations relatives au traitement des données personnelles facilement accessibles aux personnes concernées. Cela inclut la mise à disposition d'une politique de protection des données personnelles, qui est disponible sur notre site web et peut être consultée par les utilisateurs.

Les personnes sont-elles informées de la collecte, que celle-ci soit réalisée directement auprès des personnes ou non ?

Oui, AssessFirst s'engage à informer les personnes concernées de la collecte de leurs données personnelles, que celle-ci soit réalisée directement auprès d'elles ou non. Lorsque les données sont collectées directement auprès des personnes, des informations claires et concises leur sont fournies sur la finalité de la collecte, les types de données collectées, ainsi que sur leurs droits en matière de protection des données.

Les personnes ont-elles conscience d'interagir avec une machine ?

Oui, AssessFirst s'efforce de fournir aux utilisateurs une transparence totale sur le fait qu'ils utilisent un système d'intelligence artificielle (IA). Lorsqu'une personne utilise les services d'AssessFirst, elle a conscience que les résultats sont générés par une machine et non avec un être humain. Ces informations permettent aux utilisateurs de comprendre que leurs résultats sont générés par un système d'IA, basé sur des algorithmes et des modèles prédictifs.

Dans le cas où les utilisateurs interagissent avec une machine (bot) ou du contenu généré automatiquement (par exemple deepfake), cela est-il clairement indiqué à l'utilisateur ?

AssessFirst n'utilise pas d'IA qui nécessite "d'interagir", au sens social du terme, avec une machine. De même, AssessFirst n'utilise pas d'IA générative.

Si le responsable de traitement est une administration, l'information des personnes relative à un traitement algorithmique auquel elles seraient soumises est-elle prévue conformément au code des relations entre le public et l'administration et à la loi pour une République numérique ?

Non applicable.

Si le responsable de traitement est une administration, le code source de l'algorithme est-il rendu public à priori ?

Non applicable.

À défaut, sera-t-il transmis aux personnes qui en feront la demande via la Commission d'accès aux documents administratifs (CADA) ?

Non applicable.

L'impact sur la santé mentale et le comportement des personnes a-t-il été étudié ?

L'impact sur la santé mentale et le comportement des personnes n'a pas été étudié. Toutefois, ces considérations sortent du scope du type d'IA développées par AssessFirst, et de ses cas d'utilisation, de ses finalités et de ses objectifs.

L'utilisation de méthodes visant à influencer les choix des personnes est-elle envisagée ?

Non, l'utilisation de méthodes visant à influencer les choix des personnes n'est pas envisagée dans les systèmes d'IA d'AssessFirst. L'objectif principal de ces systèmes est de fournir des informations, des évaluations et des recommandations basées sur des critères objectifs tels que la personnalité, les motivations et les capacités de raisonnement. L'accent est mis sur l'apport d'une valeur ajoutée aux utilisateurs en les aidant à prendre des décisions éclairées et à mieux comprendre leurs propres profils. Il est important de souligner que les décisions finales restent entre les mains des utilisateurs. Les recommandations et les informations fournies par les systèmes d'IA d'AssessFirst sont destinées à être utilisées comme des outils d'aide à la décision, mais les choix et les actions finales sont de la responsabilité de chaque individu. AssessFirst s'engage à respecter les principes éthiques et à promouvoir la transparence, l'autonomie et le respect des droits des utilisateurs. L'objectif est de fournir des informations objectives et impartiales pour faciliter la prise de décision, sans chercher à influencer ou à manipuler les choix des personnes.

Le système pourrait-il mener à une utilisation exacerbée d'un produit, similaire à une addiction ?

Les systèmes d'IA d'AssessFirst ne sont pas conçus pour encourager une utilisation excessive ou addictive d'un produit ou d'un service. Leur objectif est de fournir des évaluations, des recommandations et des informations basées sur des critères objectifs, tels que la personnalité, les motivations et les capacités de raisonnement. Il est important de noter que l'utilisation d'un produit

ou d'un service de manière excessive ou addictive relève de la responsabilité individuelle de chaque utilisateur. Les systèmes d'IA d'AssessFirst ne cherchent pas à influencer ou à manipuler les comportements des utilisateurs dans le but d'encourager une utilisation excessive. AssessFirst encourage une utilisation responsable des technologies et met en place des mesures pour promouvoir la transparence, le consentement éclairé et le respect des droits des utilisateurs. Il revient aux individus de faire preuve de discernement et de prendre des décisions équilibrées en fonction de leurs propres besoins et de leurs valeurs.

Pourrait-il faciliter le harcèlement ?

Les systèmes d'IA d'AssessFirst ne sont pas conçus pour encourager ou faciliter des pratiques de harcèlement. Cependant, il est important de noter que le comportement des individus ne relève pas du contrôle direct d'AssessFirst. Les utilisateurs doivent toujours faire preuve de respect et d'éthique lorsqu'ils interagissent avec d'autres personnes, que ce soit dans un contexte professionnel ou social. AssessFirst encourage ses utilisateurs à respecter les principes d'égalité, de non-discrimination et de respect des droits fondamentaux. Si des cas de harcèlement ou d'abus sont signalés, AssessFirst prendra les mesures appropriées pour enquêter et agir conformément à sa politique de protection des utilisateurs.

Promouvoir la transparence et les droits de l'utilisateur final

Fournir un cadre propice à l'exercice des droits liés à la protection des données

Quelles mesures permettent aux personnes d'exercer leurs droits ?

Toute personne ayant un compte personnel AssessFirst peut se rendre sur son compte pour y exercer ses droits. Notre FAQ Legal permet d'en savoir plus.

Sont-elles informées de comment et auprès de qui ces droits peuvent être exercés ?

Oui, dans la politique de confidentialité. Notre DPO est joignable sur l'adresse email privacy@assessfirst.com en cas de questions supplémentaires concernant l'exercice des droits.

Le droit d'opposition au traitement de ses données peut-il facilement être exercé par l'individu pour les phases d'entraînement comme de production ? Est-il possible pour les individus d'exercer ce droit à tout moment (avant et après la collecte) ?

Il n'y a pas de droit d'opposition au sens strict du terme pour le système d'intelligence artificielle. En revanche, nous rappelons que la personne passant les questionnaires AssessFirst peut supprimer son compte à tout moment.

Le droit à l'effacement et le droit d'accès peuvent-ils facilement être exercés par l'individu ?

Le droit à l'effacement et le droit d'accès peuvent être exercés facilement par l'individu en se connectant sur son compte AssessFirst. En cas de question, l'individu peut joindre le DPO sur privacy@assessfirst.com.

Le droit à la limitation du traitement et le droit à la rectification peuvent-ils facilement être exercés par l'individu ?

Le droit à la rectification peut être exercé facilement par l'individu en se connectant sur son compte AssessFirst. L'individu peut joindre le DPO sur privacy@assessfirst.com en cas de question sur son droit à la rectification ou en cas de demande de droit à la limitation du traitement.

Dans le cas où le système d'IA est utilisé pour établir un profil des personnes, le profil prédit par le système est-il utilisé à titre indicatif lors du traitement ou est-il un résultat en soi ? Si le profil constitue un résultat en soi, le droit à la rectification peut-il facilement être exercé ?

Les systèmes d'IA AssessFirst ne sont pas utilisés pour établir le profil d'une personne.

Promouvoir la transparence et les droits de l'utilisateur final

Encadrer les décisions automatisées

Le système conduit-il à une décision fondée exclusivement sur un traitement automatisé, par conception ou dans les faits, et produisant des effets juridiques la concernant ou l'affectant de manière significative de façon similaire ?

Non, le système d'AssessFirst n'est pas conçu pour prendre des décisions fondées exclusivement sur un traitement automatisé qui produirait des effets juridiques significatifs pour les individus concernés. Les résultats et les informations fournis par le système sont destinés à être utilisés comme outil d'aide à la décision par les utilisateurs, mais les décisions finales restent de la responsabilité des utilisateurs. L'objectif d'AssessFirst est d'apporter des informations et des recommandations basées sur l'analyse de données, mais ces informations ne doivent pas être considérées comme des décisions définitives prises par le système lui-même. Les utilisateurs conservent leur rôle décisionnel et peuvent prendre en compte d'autres facteurs pertinents avant de prendre une décision.

Même si certains systèmes ne semblent pas produire un effet sur les personnes, des conséquences peuvent-elles exister dans les faits (par ex. : un algorithme de tri des CV pour le recrutement qui ordonnerait certains CV en bas d'une file trop longue pour qu'un humain puisse vérifier la pertinence de la décision prise par l'algorithme) ?

Oui, il est possible que des conséquences se produisent dans les faits, même si elles ne sont pas directement produites par le système d'IA lui-même. C'est pour cela qu'AssessFirst met en place des mécanismes de contrôle et de suivi pour identifier et atténuer ces conséquences potentielles. Une supervision humaine, des évaluations régulières de la performance et de l'équité du système, ainsi que des processus de révision et de correction peuvent contribuer à minimiser ces effets indésirables. Aussi, l'accent est mis sur la transparence, la responsabilité et l'évaluation continue des systèmes d'IA afin de garantir qu'ils fonctionnent de manière éthique, équitable et respectueuse des droits des individus.

Quelle est la base légale sur laquelle repose la prise de décision automatisée ?

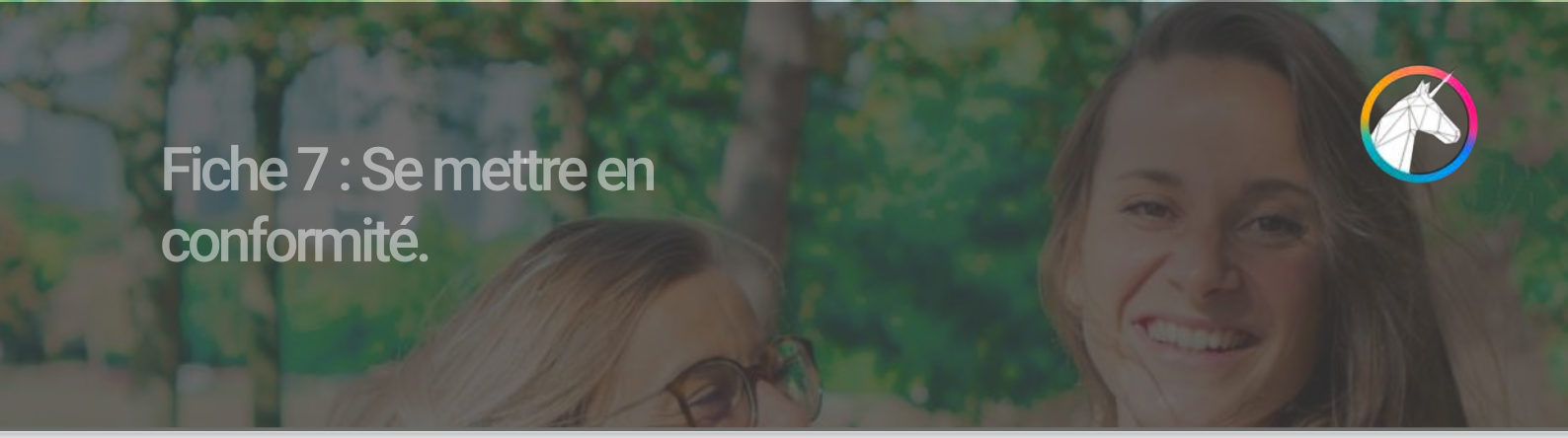
Le système d'AssessFirst n'est pas conçu pour prendre des décisions automatisées.

Les personnes peuvent-elles facilement s'opposer à une décision fondée exclusivement sur un traitement automatisé ? Si oui, par quels moyens ?

Le système d'AssessFirst n'est pas conçu pour prendre des décisions automatisées.

Les individus décidant de ne pas avoir recours au système d'IA en application de l'article 22 bénéficient-ils des mêmes avantages et possibilités que ceux utilisant le système ?

Non applicable.



Fiche 7 : Se mettre en conformité.

Attribuer les responsabilités et documenter son traitement

Respecter les normes, certifications et code de bonne conduite comme preuves

L'IA a-t-elle été certifiée par un organisme tiers (LNE, bureau d'étude, etc.) ou par une autorité (HAS, ANSM, AMF, etc.) ?

À ce jour, les systèmes d'IA d'AssessFirst n'ont pas encore été certifiés par un organisme tiers. Cependant, nous attachons une grande importance à la certification de nos systèmes et nous sommes engagés dans des démarches en cours en vue d'obtenir cette certification. Nous reconnaissons l'importance de l'évaluation indépendante de nos systèmes pour garantir leur fiabilité, leur transparence et leur conformité aux normes éthiques et de qualité.

Le système d'IA respecte-t-il certains codes de conduite ou bonnes pratiques ?

Oui, le système d'IA d'AssessFirst est conçu pour respecter certains codes de conduite et bonnes pratiques. Nous nous engageons à adopter une approche éthique et responsable dans le développement et l'utilisation de nos systèmes d'IA. Voici quelques principes que nous suivons :

- Respect de la vie privée : Nous accordons une grande importance à la protection des données personnelles et nous nous conformons aux lois et réglementations en vigueur en matière de confidentialité des données ;
- Explicabilité : Nous nous efforçons de rendre nos systèmes d'IA explicables en fournissant des informations compréhensibles sur leur fonctionnement, leurs limites et leurs implications ;
- Équité : Nous nous engageons à éviter les biais discriminatoires dans nos systèmes d'IA et à assurer un traitement équitable de toutes les personnes concernées ;
- Responsabilité : Nous assumons la responsabilité de nos systèmes d'IA et de leurs résultats. Nous mettons en place des mécanismes de contrôle et de surveillance pour identifier et corriger les éventuels problèmes ou erreurs ;
- Collaboration : Nous encourageons la collaboration et le dialogue avec les parties prenantes, y compris les chercheurs, les experts en éthique de l'IA et les utilisateurs, afin de bénéficier de leurs connaissances et de leurs points de vue dans l'amélioration continue de nos systèmes ;
- Performance : Nous accordons une grande importance à la performance et à la fiabilité de nos systèmes d'IA. Nous nous efforçons de développer des solutions technologiques robustes et précises, capables de fournir des résultats cohérents et fiables. Nous effectuons des tests rigoureux et des évaluations continues pour mesurer et améliorer la performance de nos systèmes, en tenant compte des retours des utilisateurs et des évaluations externes lorsque cela est possible. Notre objectif est de fournir des résultats de haute qualité qui répondent aux attentes des utilisateurs et aux besoins spécifiques de chaque cas d'utilisation.

Ces codes de conduite et bonnes pratiques sont intégrés dans nos processus de développement et de déploiement des systèmes d'IA, afin de garantir que nos solutions respectent des normes éthiques et sociales élevées.

Attribuer les responsabilités et documenter son traitement

Effectuer une AIPD et l'évaluer

Une analyse d'impact relative à la protection des données (AIPD) a-t-elle été menée ?

AssessFirst a effectué une AIPD globale, mise à jour en mars 2023. Toutefois, AssessFirst n'a pas conduit d'AIPD unique dans le cadre du système d'IA.

Un outil d'évaluation de l'impact du système d'IA a-t-il été utilisé ?

À l'heure actuelle, AssessFirst n'a pas utilisé d'outil d'évaluation de l'impact du système d'IA.

Attribuer les responsabilités et documenter son traitement

Documenter pour fiabiliser

Une documentation concernant les modalités de collecte et de gestion des données d'entraînement et de production utilisées, l'algorithme, la qualité des sorties du système, les outils utilisés, la journalisation, les mesures de sécurité existe-elle ?

Une documentation en cours de conception.

Cette documentation est-elle partagée avec toutes les personnes ayant à en connaître pour assurer une analyse et une maîtrise des risques efficaces (utilisateurs du système d'IA, comité d'éthique, service qualité et gestion des risques, personnes concernées, etc.) ?

Une documentation en cours de conception.

Une approche en sources ouvertes a-t-elle été adoptée afin d'impliquer une communauté de tiers dans la conception et l'amélioration du système ?

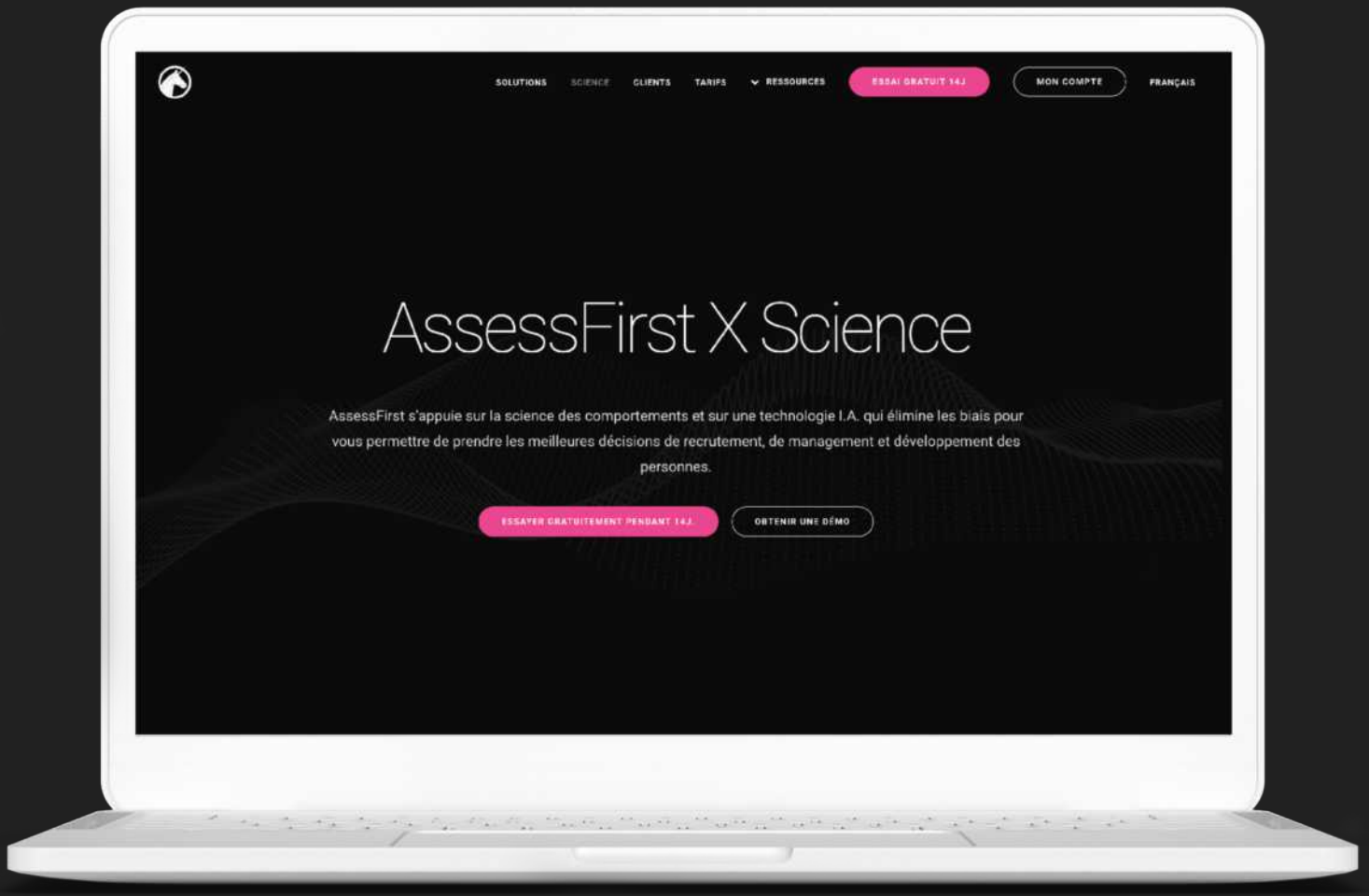
En raison de considérations commerciales, les algorithmes utilisés dans nos systèmes d'IA ne sont pas accessibles en open source. Cette décision a été prise pour protéger la propriété intellectuelle et préserver notre avantage concurrentiel. Cependant, nous accordons une grande importance à la transparence et à la confiance de nos clients, c'est pourquoi nous nous engageons à fournir des informations détaillées sur notre approche méthodologique et les principes sur lesquels reposent nos modèles d'IA.

Conclusion

Chez AssessFirst, nous sommes fermement convaincus que le véritable progrès ne peut se faire qu'en servant équitablement les intérêts des candidats et des entreprises. C'est cette croyance qui a façonné notre module d'IA, permettant à chacun de mieux se comprendre et de prendre des décisions plus judicieuses. Aujourd'hui, ce sont plus de 3500 entreprises réparties dans 40 pays et plus de 5 millions de personnes qui utilisent AssessFirst pour bâtir un monde plus efficace et plus juste. En plaçant l'éthique et le contrôle des biais au cœur de nos préoccupations, nous nous assurons que notre technologie reste une force pour le bien. Les algorithmes d'AssessFirst sont conçus avec la plus grande vigilance, chaque étape de la création et de l'utilisation de notre IA étant minutieusement surveillée. Notre objectif est simple : développer une IA éthique et dénuée de biais. Chaque analyse, chaque algorithme, est ainsi le fruit d'un travail d'équipe assidu et réfléchi. Nos équipes travaillent sans relâche pour garantir l'équité des analyses prédictives et pour éviter que nos algorithmes, lorsqu'ils sont utilisés dans les processus décisionnels, n'engendrent des discriminations par la voie de biais algorithmiques. Notre engagement pour l'équité est reconnu dans la communauté scientifique à travers plusieurs publications de recherche, attestant ainsi de notre dévouement à construire un avenir où la technologie et l'éthique peuvent coexister en harmonie.

Emeric Kubiak
Head of Science

Lucile Whitbeck
Head of Legal & DPO



ASSESSFIRST

AssessFirst a développé une solution de recrutement prédictif permettant aux entreprises de prédire dans quelle mesure les candidats et collaborateurs réussiront et prospéreront dans leur travail. La solution AssessFirst analyse les données de plus de 5 000 000 de profils, qu'ils soient candidats, salariés ou professionnels du recrutement. Plus de 3 500 entreprises utilisent la solution pour augmenter leurs performances jusqu'à 25 %, diminuer leurs coûts de recrutement de 20 % et réduire le taux de turnover de leurs employés de 50 %.