

IA digne de digne de confiance

— CONSTRUIRE UNE GOUVERNANCE
ADAPTÉE À CHAQUE ENTREPRISE

Un guide pour une gouvernance adaptée à chaque entreprise

Ce document retrace le travail d'une quinzaine d'entreprises françaises rassemblées au sein du groupe de travail IA Responsable d'Impact AI. En pleine pandémie de la Covid-19, ces organisations ont continué à se réunir virtuellement et en toute confidentialité afin d'explorer et d'analyser leurs enjeux de la gouvernance de l'intelligence artificielle. Chaque partie résume ainsi les discussions des ateliers en ligne effectués entre mars et juin 2020 afin de :

- Partager des expériences de mise en œuvre des principes éthiques avec les autorités travaillant sur des projets de réglementation ;

- Présenter des bonnes pratiques afin de soutenir les entreprises qui s'engagent sur la voie de la gouvernance de l'IA ;

- Sensibiliser plus largement d'autres organisations et entreprises aux enjeux de l'intelligence artificielle éthique.

— PRÉFACE



Renaud Vedel

— COORDONNATEUR DE LA
STRATÉGIE NATIONALE POUR
L'INTELLIGENCE ARTIFICIELLE

La France a l'ambition de devenir un acteur international majeur de l'intelligence artificielle (IA) et voici deux ans que le Gouvernement a défini la stratégie nationale qui la matérialise. Son développement s'intègre aux propres orientations de la Commission européenne, qui visent en particulier à bâtir une IA digne de confiance et plus souveraine pour l'Europe. Les entreprises, essentielles à la réussite de nos objectifs économiques, doivent également être conscientes des enjeux éthiques et sociaux de ces technologies inédites et assumer leur part de responsabilité. C'est pourquoi je soutiens les initiatives vertueuses d'Impact AI. Ce guide, fruit de la réflexion d'une quinzaine d'entreprises et organisations nationales, participe de l'action pour nous engager collectivement sur la voie d'une IA digne de confiance.

Impact AI a analysé les défis et promesses que l'introduction puis la généralisation de l'intelligence artificielle engendrent afin d'édifier une IA digne de confiance et d'apporter sa contribution aux consultations européennes.

Enjeux sociétaux, économiques et normatifs de l'IA

L'INTELLIGENCE ARTIFICIELLE S'APPUIE SUR DES DONNÉES POUR PROPOSER DES MODÈLES PRÉDICTIONNELS PORTEURS D'AVANTAGES-CLÉS POUR LES ORGANISATIONS DANS DE TRÈS NOMBREUX DOMAINES. Néanmoins, si prometteuse soit-elle, son acceptation sociale n'est pas acquise. La suspicion envers le progrès technique va croissant au sein du grand public tandis que, parallèlement, l'IA pose de légitimes questions au sein des entreprises sur la conception de ses applications, leur but et leur usage. Par exemple, l'utilisation de technologies de reconnaissance faciale pour lutter contre le terrorisme pourrait, si elle était abusive, faire émerger une société de surveillance. La confiance constitue donc le préalable indispensable à l'adoption de l'IA par tous et doit être pensée par les organisations.

Une gouvernance proportionnée aux risques

Une entreprise a tout intérêt à s'emparer de la problématique de la confiance de l'IA car une défaillance d'application ou des mésusages médiatisés peuvent entacher son image, fragiliser sa position concurrentielle, dissuader les meilleurs candidats de la rejoindre. D'autant que, si les organisations sont familières de la gestion de tels risques, l'intelligence artificielle a ceci de particulier qu'elle les amplifie considérablement et en crée de nouveaux. Aussi, pour maximiser les bénéfices de l'IA et en contenir les dangers, les organisations responsables adoptent-elles une démarche éthique by design, c'est-à-dire qui intègre ces enjeux dès la phase de conception. Les organisations s'engagent dans une gouvernance pour une IA digne de confiance proportionnée aux risques afin de corriger ou d'atténuer les effets délétères de la technologie, par des outils techniques ou des protocoles adaptés. Cependant, nombre d'entreprises, bien que conscientes de ces enjeux, manquent des ressources et méthodes nécessaires à cette démarche.

De leur côté, les autorités, en particulier européennes, ont multiplié les initiatives pour délivrer un cadre normatif éthique. Ainsi, le groupe d'experts AI HLEG¹ missionné par la Commission a notamment conçu une grille d'évaluation éthique des solutions d'IA et invité les parties prenantes à la tester. Répondant à cet appel, Impact AI a formulé des recommandations pour la rendre applicable dans les organisations et, ce faisant, participer à l'élaboration d'un cadre juridique opérationnel adapté aux enjeux des entreprises. Le collectif a identifié les besoins et dispositifs actuels de gouvernance de l'IA au sein des entreprises françaises synthétisés dans ce guide.

LES BUTS SOUHAITABLES DES CADRES NORMATIFS

Les cadres réglementaires de l'IA doivent se concentrer avant tout sur deux objectifs-clés : d'abord inciter les acteurs à adopter leurs propres normes de gouvernance pour une IA digne de confiance qui soient opérationnelles et adaptées à leurs activités ; ensuite, soutenir le développement de technologies, de systèmes et d'outils qui les aident à identifier et atténuer les risques.

1. Le High Level Expert Group on AI (AI HLEG) est un groupe de cinquante-deux experts nommés par la Commission européenne dont la mission consiste à soutenir la mise en œuvre de la stratégie européenne d'intelligence artificielle.

THALES	GOVERNANCE D'APPLICATIONS CRITIQUES
	<p>Pilotage d'avion, sécurisation de flux bancaires, reconnaissance faciale aux frontières... Les solutions IA de Thales sont critiques et interrogent l'éthique dans de multiples dimensions telles que les biais, le respect des données personnelles ou leur fiabilité. L'entreprise a bâti une gouvernance d'une IA digne de confiance en fonction de ces enjeux singuliers. Elle s'appuie sur une « Charte éthique et de transformation numérique » et sur la démarche Thales TrUE qui promeut une IA transparente où les utilisateurs ont accès aux données, une IA intelligible, qui permet d'en justifier les résultats, et une IA éthique, qui respecte protocoles standardisés, législation et droits de l'homme.</p>

ORANGE	INNOVATION ET ENGAGEMENT SOCIÉTAL
	<p>L'opérateur, qui entend tenir un rôle-clé dans le développement d'une société numérique de confiance, voit l'intelligence artificielle comme un tremplin pour accélérer sa propre transformation numérique et proposer des services nouveaux à ses clients. C'est pourquoi la question de l'éthique a d'abord été abordée sous l'angle de l'innovation afin d'anticiper le plus en amont possible la gestion des risques, tel que celui des discriminations. La signature d'une charte pour une IA inclusive, en avril 2020, a marqué la première concrétisation formelle de l'engagement d'Orange dans une gouvernance en ligne avec ses valeurs et sa promesse sociétale.</p>

MICROSOFT	CONTRATS ET BESOINS DES ACTEURS
	<p>Selon Microsoft, l'IA doit se placer au service de l'humain, sans jamais transiger sur la transparence et la confiance. C'est ainsi que, en s'appuyant sur un vaste et rigoureux travail de veille sur l'appareil normatif international et ses évolutions, l'entreprise a pu définir un principe d'accès et de partage des données : « aussi ouvert que possible, aussi fermé que nécessaire ». Quatre types de contrats ont été élaborés pour sécuriser juridiquement les initiatives des acteurs selon leurs besoins : production d'applications ouvertes à partir d'informations publiques (O-UDA), données publiques utilisées à des fins de calculs algorithmiques (C-UDA), données privées cédées pour concevoir des applications d'IA en open source (DUA-OAI), données émanant de protagonistes différents qui les agrègent pour créer des applications mutualisées. (DUA-DC).</p>

Impact AI a déterminé sept principes pour bâtir une IA digne de confiance puis, à l'aune de l'expérience de ses membres, a examiné les conditions techniques et organisationnelles de réussite de leur mise en œuvre et les difficultés rencontrées.

Des principes à la pratique

LES SEPT PILIERS ÉTHIQUES D'IMPACT (voir ci-après et page 5) synthétisent ceux de l'Institute of Electrical and Electronics Engineers (IEEE), de l'OCDE et de l'Union européenne. Leur implémentation dans les organisations procède d'une analyse préalable de leurs risques spécifiques. Cette évaluation s'accompagne d'une réflexion sur les moments opportuns pour vérifier le respect des règles éthiques durant cycle de vie d'une application d'intelligence artificielle, sur les solutions techniques automatisables susceptibles de contenir d'éventuelles dérives ainsi que sur les référentiels, audits et certifications les plus adaptés aux particularités de l'organisation. Les réponses apportées se heurtent parfois à des obstacles que les entreprises s'appliquent à dépasser.

Robustesse, équité, responsabilité

Pour créer les conditions de la confiance, les solutions d'IA doivent être robustes, c'est-à-dire capables de fournir des résultats fiables durant tout leur cycle de vie, en limitant les erreurs ou les incohérences des prédictions. La conception des algorithmes et l'échantillonnage des données d'apprentissage s'avèrent ici cruciaux. Ainsi, le machine learning apprend à partir d'un éventail d'exemples qui doivent être suffisamment nombreux et divers pour que les résultats soient généralisables. À défaut, l'algorithme se fondera sur des raccourcis pour établir des prédictions qui de ce fait se révéleront erronées. Or, il arrive que des exemples soient légèrement modifiés ce qui aboutit à des erreurs avec des conséquences graves pour les utilisateurs. La R&D des entreprises s'efforce de résoudre ce problème scientifique.

Alors que l'on attend des résultats d'une application d'intelligence artificielle une neutralité parfaite, ceux-ci s'avèrent parfois discriminatoires. En effet, les ensembles de données d'apprentissage, constitués dans le passé, sont quelquefois porteurs de préjugés que l'IA va alors reproduire et amplifier. L'équité consiste à mettre en place des procédures de conception et de vérification qui débarrassent l'IA de ces biais cognitifs, statistiques ou économiques. Cela requiert d'additionner cette exigence à celle de précision des résultats.

Enfin, le développement de l'intelligence artificielle s'accompagne de garde-fous qui minimisent les dommages qu'elle peut causer. Les mécanismes mis en place doivent se soumettre à l'obligation de rendre des comptes.

LES SEPT PRINCIPES D'IMPACT AI POUR UNE IA DIGNE DE CONFIANCE (VOIR EXEMPLES PAGE 5)

- <1>. Dignité, pour une IA au service d'un utilisateur maître de ses décisions
- <2>. Robustesse, pour une IA fiable dans la durée
- <3>. Gouvernance des données, pour une IA respectueuse de la vie privée
- <4>. Transparence, pour une IA interprétable et explicable
- <5>. Équité, pour une IA qui traite chaque personne de manière juste
- <6>. Développement durable et bien-être, pour une IA qui contribue à résoudre des problèmes universels
- <7>. Responsabilité, pour une IA capable de rendre des comptes sur l'efficacité des mécanismes de minimisation des risques

AXA	<p>TRANSPARENCE</p> <p>La compagnie s'est associée à l'OCDE pour expliquer les prédictions des algorithmes sur les causes des crises économiques en se basant sur le principe de la transparence. Elle s'implique également dans les travaux d'une thèse à Sorbonne Université qui étudie la manière optimale de faire comprendre aux néophytes les résultats d'une IA en se fondant sur les méthodes actuelles d'interprétabilité.</p>
AXIONABLE	<p>DÉVELOPPEMENT DURABLE & BIEN-ÊTRE</p> <p>La multitude des référentiels, la variété des définitions des normes et des périmètres d'application qui en découlent, compliquent la mesure de l'impact social et environnemental des solutions d'IA. La classification des activités dites à hauts risques - reconnaissance faciale, applications de santé... -, actuellement en discussion à la Commission européenne, posera les premiers jalons d'un cadrage bienvenu.</p>
DELOITTE	<p>GOVERNANCE DES DONNÉES</p> <p>Selon une méthodologie interne, l'entreprise ordonnance la collecte des données stratégiques de ses clients et assure leur intégrité dans le processus d'apprentissage. Les procédures et ensembles de données sont testés et documentés dans leur intégralité, à toutes les étapes du cycle de vie du système d'IA. Ces dispositions s'appliquent aux outils développés en interne mais aussi en externe et maintiennent ainsi la confiance des utilisateurs.</p>
SCHNEIDER ELECTRIC	<p>DIGNITÉ</p> <p>Outre l'objectif business, les spécifications des solutions d'IA récapitulent les besoins de formation de l'utilisateur ainsi que le niveau d'implication et de contrôle qu'il serait souhaitable de lui offrir. Il recense aussi les interfaces nécessaires à sa maîtrise des données, leur catégorisation ou leur sélection pour un éventuel réapprentissage par l'algorithme (présentation d'options et/ou d'explications, alerte en présence de données atypiques, etc.).</p>

Si le besoin de gouvernance s'avère universel, les mises en pratique varient selon les entreprises, leurs activités, leur taille, leur organisation et leur culture. Quelques lignes directrices communes se dégagent néanmoins.

Mettre en œuvre la gouvernance

LA GOUVERNANCE CONSISTE PRIORITAIREMENT À IDENTIFIER, prévenir et corriger les potentiels effets délétères de l'utilisation de l'intelligence artificielle ou à contenir l'intensification des risques traditionnels de l'entreprise qu'elle génère.

Seule la direction générale peut conférer sa légitimité au projet. C'est pourquoi la participation d'un membre du Comex¹ aux instances de gouvernance démontre l'intérêt porté aux enjeux éthiques de la technologie ; la prise en compte du développement de l'IA dans le plan stratégique témoigne quant à elle de la portée du sujet et de son caractère transversal.

La réussite de la gouvernance tient également à la sensibilisation des collaborateurs, qu'ils soient managers, concepteurs, développeurs ou simples utilisateurs. Ils doivent prendre la mesure des potentialités de l'IA sans en ignorer les risques. L'association de représentants de toutes les entités de l'entreprise à la réflexion sur le projet de gouvernance peut lancer un processus d'acculturation que des actions complémentaires des RH et de communication interne s'attacheront à diffuser dans l'entreprise.

Stratégie, organes et outils

L'intégration précoce de la gouvernance de l'IA aux missions de départements aux responsabilités comparables déjà en place dans l'organisation, assure une cohérence d'ensemble à la politique de gestion des risques. La stratégie de la gouvernance est donc portée par un Comité éthique étendu à l'IA, ou inversement, qui définit des normes, des objectifs et une feuille de route, le tout matérialisé dans une Charte. Un(e) responsable y anime les débats des représentants des principales fonctions de l'entreprise enrichis parfois du regard d'experts externes.

Un ou plusieurs autres organes se chargent de la mise en œuvre opérationnelle de cette politique, notamment en concevant une boîte à outils largement diffusée dans l'organisation. Parmi eux, les checklists de vérification du respect des principes éthiques des solutions d'IA ont fait leurs preuves. Pour assurer la crédibilité à long terme de la gouvernance, il est nécessaire de mener des audits réguliers, de mesurer l'efficacité des recommandations et de recueillir l'avis des collaborateurs sur les impacts de l'IA. Enfin, la communication externe, en particulier le rapport annuel, permettra de faire rayonner la stratégie de l'entreprise.

OUTILS DE GOUVERNANCE

Les référentiels conçus par les organismes extérieurs à l'entreprise jouent un rôle majeur pour asseoir la politique de gouvernance. Ces instruments solides structurent la démarche, l'objectivent et la rendent opposable. Trois possibilités s'offrent aux entreprises dans ce domaine : les labels volontaires, les certificats et les normes.

1. Comex : comité exécutif

CERCLE INTERELLES	MÉTHODOLOGIE CONTRE LES DISCRIMINATIONS DE GENRE
	<p>Selon cette association qui œuvre pour la féminisation et l'égalité professionnelle dans les secteurs scientifiques et technologiques, il est très utile de tester auprès d'un échantillon réduit de répondants suffisamment avertis des problématiques de l'IA, les checklists censées évaluer une application d'intelligence artificielle sous l'angle des discriminations de genre. Seront notamment mesurés le temps de réponse et l'intelligibilité des questions posées. L'utilisation de la checklist par un groupe d'individus (équipe-projet) évite aussi les automatismes et l'autosatisfaction dans les réponses. D'autant que, pour les collaborateurs, de se poser collectivement la question de leur maturité en termes d'IA accélère l'acculturation sur ce sujet.</p>

MAIF	ENGAGEMENT DU COMEX
	<p>Au sein de la mutuelle, deux directeurs généraux adjoints, dans leurs domaines respectifs, sont devenus sponsors des projets internes portant sur les données et l'IA : la direction de la Sécurité des systèmes d'information et la direction de l'Assurance de la personne, de la data et de l'actuariat produit. Ces questions sont partie intégrante de leurs responsabilités et ils sont porteurs de ces sujets au sein du Comité de direction générale. Les engagements de la MAIF sont formalisés dans la « Charte pour un monde numérique résolument humain et éthique », publiée en 2017.</p>

SUBSTRA FOUNDATION	CONCEPTION D'UN RÉFÉRENTIEL EN MODE PARTICIPATIF
	<p>Substra Foundation, organisation non lucrative qui travaille au développement d'écosystèmes de data de confiance, soutient les entreprises qui souhaitent autoévaluer leur démarche. Elle déploie une approche participative dans le but de concevoir un référentiel « data responsable et de confiance » dont les organisations se saisissent pour apprécier la maturité de leur politique d'éthique de l'IA et pour progresser en puisant dans des ressources techniques adaptées à chaque critère d'évaluation. Elles peuvent ainsi, en toute autonomie, identifier leurs points forts et concentrer leurs efforts sur les actions à mettre en œuvre pour améliorer leurs performances dans d'autres domaines.</p>

ÉCHELLE DE MATURITÉ DE LA GOUVERNANCE DE L'IA SELON IMPACT AI

À LA LUMIÈRE DES EXPÉRIENCES DE SES MEMBRES, IMPACT AI A MIS AU POINT UNE GRILLE D'ÉVALUATION DE L'AVANCEMENT DES SYSTÈMES DE GOUVERNANCE DE L'IA ET LA MET À DISPOSITION DES ENTREPRISES. LES MESURES DE MATURITÉ DEVRONT SE RÉPÉTER AU GRÉ DE L'ÉVOLUTION DU DÉVELOPPEMENT DE L'IA ET SES CRITÈRES S'Y ADAPTERONT.

THÈME	NIVEAU 1	NIVEAU 2	NIVEAU 3	NIVEAU 4	NIVEAU 5
PRINCIPES ÉTHIQUES	Recensement et comparaison des principes éthiques des autorités (UE, OCDE...).	Positionnement des principes éthiques internes de l'organisation par rapport à ceux définis par les autorités.	Publication en interne des principes éthiques de l'IA (charte...).	Publication à l'externe des principes éthiques de l'IA.	Audit de l'application des principes éthiques listés dans la charte.
SPONSORING	Aucun responsable n'est identifié. Les collaborateurs s'organisent spontanément pour explorer le sujet et faire remonter le besoin au Comex.	Le Comex confirme son sponsoring et nomme un(e) responsable du développement d'une gouvernance de l'IA.	Un(e) responsable de la gouvernance a été nommé(e) mais ne dispose pas de la légitimité ou des ressources nécessaires pour mobiliser les parties prenantes pertinentes.	La personne responsable de cette mission dispose de la légitimité et des ressources nécessaires pour coconstruire la gouvernance de l'IA.	La personne responsable de la gouvernance de l'IA rapporte directement au Comex avec le CTO (Chief Technical Officer).
MODÈLE DE GOUVERNANCE	Aucun modèle de gouvernance n'est choisi.	Recensement et comparaison des modèles de gouvernance actuels / ajout de thématiques sur l'IA digne de confiance à une gouvernance interne préexistante.	Création d'un comité de gouvernance interne afin de piloter la mise en place de la charte.	Le comité de gouvernance représente tous les métiers et zones géographiques de l'organisation, déploie des outils de gestion des risques et des formations.	Le comité de gouvernance s'ouvre aux experts externes et communique dans le rapport annuel. Il initie une réflexion sur l'évolution du modèle de gouvernance.
PROTOCOLES ET OUTILS	Aucun outil de gouvernance de l'IA n'est identifié.	Recensement et comparaison des outils du marché.	Recensement et qualification des besoins des métiers ainsi que des risques liés à l'IA, notamment à travers une checklist.	Sélection des outils choisis dans une boîte à outils de gouvernance et mise à la disposition des parties prenantes pertinentes.	Sensibilisation en interne et formation sur l'utilisation de la boîte à outils, elle-même mise à jour régulièrement.

— REMERCIEMENTS

Coordinateurs du cycle de travail sur la gouvernance d'une IA digne de confiance

Paul-Marie Carfantan AI Ethics & Governance Consultant DELOITTE FRANCE
Sarah El Marjani Research & Partnership PMO AXA



Membres du groupe de travail IA Responsable ayant participé aux différentes étapes de création du guide

Grégory Abisor	DELOITTE FRANCE
Jonathan Aigrain	AXA
Ismael Al-Amoudi	GRENOBLE EM
Teva Atani	DELOITTE FRANCE
Olivier Baes	MAIF
Daniel Bartolo	MAIF
Jean-David Benassouli	PWC FRANCE
Romain Bey	SUBSTRA
Gwendal Bihan	AXIONABLE
Emmanuel Bloch	THALES
Éric Boniface	SUBSTRA
Sandrine Bortone	CERCLE INTERELLES
Sylvie Caruso Cahn	SNCF RÉSEAU
Marcin Detyniecki	AXA
Hector Dupont de Rivoire	MICROSOFT
Richard Eudes	DELOITTE FRANCE
Émilie Sirvent-Hien	ORANGE
Hélène Jeannin	ORANGE
Anne-Marie Jonquière	CEA
Claude Le Pape Gardeux	SCHNEIDER ELECTRIC
Eneric Lopez	MICROSOFT
Nicolas Marescaux	MACIF
Louis Montanié	PWC FRANCE
Thierry Notaire	MAIF
Bernard Ourghanlian	MICROSOFT
Florence Picard	INSTITUT DES ACTUAIRES
Yves Yota Tchoffo	DELOITTE FRANCE
Jean-Jacques Thomas	SNCF RÉSEAU
Camille Vaziaga	MICROSOFT

QU'EST-CE QU'IMPACT AI ?

Créé en 2018, le collectif de réflexion et d'action Impact AI ambitionne d'éclairer les enjeux éthiques de l'intelligence artificielle. Il associe une soixantaine d'acteurs venus de tout l'écosystème numérique français : grandes entreprises, ESN (Entreprises de services numériques), sociétés de conseil, startups, institutions de recherche et du monde académique. Par ses initiatives, il nourrit le débat, formule des recommandations destinées aux pouvoirs publics, promeut des dispositifs favorisant et vérifiant l'usage responsable de l'intelligence artificielle, et soutient des projets innovants qui œuvrent dans l'intérêt général. Impact AI accompagne ainsi l'émergence d'une IA au service de tous.

