

POUR UNE UTILISATION CHOISIE ET NON SUBIE DE L'INTELLIGENCE ARTIFICIELLE

Un cadre éthique et responsable pour la métropole.

Depuis l'irruption des intelligences artificielles génératives en 2023, l'IA connaît un essor sans précédent, bouleversant nos manières de produire, d'analyser et de diffuser l'information. Capable de traiter des volumes massifs de données et de mobiliser des puissances de calcul inédites, l'IA ouvre des perspectives considérables dans des domaines stratégiques : santé, recherche scientifique, transition écologique, action publique. Mais cette révolution technologique s'accompagne de défis majeurs. L'IA interroge la place de l'humain face à la machine, la relation au savoir, à l'apprentissage, à la création, et soulève des enjeux éthiques, sociaux et environnementaux. Les risques sont multiples : biais algorithmiques, discriminations, erreurs (« hallucinations »), atteintes à la vie privée, exploitation de données protégées, impact énergétique insoutenable. À cela s'ajoutent des questions de transparence, de souveraineté numérique et de fracture territoriale.

En engageant une réflexion sur les conditions d'utilisation des IA au service des politiques publiques locales, Bordeaux Métropole a souhaité se saisir de ces sujets avec des **lignes directrices claires, en cohérence avec ses orientations numériques et sa stratégie de numérique responsable**. L'hypothèse du recours à l'IA par Bordeaux Métropole ne peut ainsi pas se concevoir comme une finalité en soi, ni comme un objet à part, elle doit s'inscrire dans la continuité des engagements métropolitains : **agir pour un numérique choisi et non subi**. Ce credo guide l'action métropolitaine et permettra de saisir en confiance les opportunités que l'IA ouvre, telles que :

- ✓ **Pour l'efficacité de l'administration** : l'automatisation de tâches répétitives (classement, rédaction, traduction, recherche d'information), le soutien à l'analyse et la décision (aide à l'instruction, détection de signaux faibles, priorisation), et l'amélioration continue des processus dans tous les métiers, sous supervision humaine et avec une auditable des résultats,
- ✓ **Pour la proximité et relation aux citoyens** : la proposition de parcours simplifiés et plus accessibles (guichets augmentés, réponses multicanales) tout en améliorant la réactivité des services et la personnalisation des réponses, dans le respect de la transparence et de la protection des données,
- ✓ **Pour les politiques publiques et l'action territoriale** : la capacité de pilotage fondée sur les données (évaluation, indicateurs, scénarios prospectifs), l'optimisation des services urbains en regard ou encore l'accélération de réponse inédite soutenant la transition écologique et la gestion des vulnérabilités du territoire.

Bordeaux Métropole s'impose néanmoins des limites strictes quant à l'usage de l'intelligence artificielle, afin de garantir un cadre éthique et protecteur. Sont ainsi proscrits les usages suivants :

- ✓ **Surveillance individuelle** : toute forme de surveillance intrusive ou de suivi personnalisé est proscrite sauf dans des circonstances exceptionnelles qui seraient expressément encadrées par la loi,
- ✓ **Notation sociale** : l'instauration de mécanismes d'évaluation, de notation ou de classement social est formellement interdite,

- ✓ **Décisions entièrement automatisées** : les systèmes d'IA ne doivent pas rendre de décisions automatisées ayant un impact significatif sur les droits ou les conditions de vie des citoyens sans aucune validation humaine,
- ✓ **Utilisation des données biométriques** : le recours à l'intelligence artificielle pour des dispositifs de reconnaissance biométrique, notamment faciale, est interdit, sous réserve de circonstances exceptionnelles qui seraient expressément encadrées par la loi.

Le présent cadre est le fruit d'**une construction collective menée sur une période de neuf mois par l'ensemble des acteurs du territoire de Bordeaux Métropole : citoyens, représentants de la société civile, élus, agents, étudiants**. Ainsi :

- ✓ Les élus métropolitains, élus au numérique des communes du territoire ainsi que des groupes d'agents de Bordeaux Métropole et des communes du territoire métropolitain ont construit et itéré autour des principes et engagements décrits dans le présent document,
- ✓ Des citoyens, représentés par les membres du conseil de développement (C2D) de Bordeaux Métropole ont été sensibilisés et ont pu partager leurs ressentis, leurs attentes et leurs craintes quant à l'utilisation de l'IA par la collectivité,
- ✓ Des représentants de la société civile (enseignants, entrepreneurs, chercheurs) constituant les membres du Comité Consultatif du Numérique ont confronté leurs regards autour des enjeux à considérer pour Bordeaux Métropole relativement à la construction d'IA par la collectivité,
- ✓ Bordeaux Métropole a accueilli et contribué à l'organisation d'une convention citoyenne étudiante menée sur 3 journées réunissant 60 étudiant·es en science politique, informatique, ingénierie et philosophie. Elle visait à proposer des orientations et actions concrètes et applicables autour au thème « Faire de l'IA une alliée de la formation ».

Le cadre de Bordeaux Métropole autour de l'intelligence artificielle repose sur **4 principes fondamentaux** :

- ✓ **Principe n°1** : Des IA maîtrisées pour garantir sécurité, autonomie et éthique,
- ✓ **Principe n°2** : Des IA utiles pour garantir un équilibre positif entre apports et impacts,
- ✓ **Principe n°3** : Des IA de confiance au service des citoyens et du territoire,
- ✓ **Principe n°4** : Des IA en appui des agents pour renforcer le service public.

Ces principes se déclinent en **16 engagements décrivant les lignes rouges et conditions de construction et utilisation d'IA en confiance pour les acteurs du territoire**. Il s'agit d'un document cadre pour les citoyens et agents du territoire métropolitain mais aussi pour les prestataires de services et les délégataires concourant au service public. Des clauses juridiques déclineront les principes ici arrêtés, notamment dans le cadre des marchés publics passés par la collectivité.

Ce document repose sur les informations disponibles sur les enjeux de l'IA à la date de son adoption. Le rythme des innovations étant très soutenu, des évolutions seront sans doute nécessaires et ce document de référence fera l'objet de mises à jour régulières.

CADRE REGLEMENTAIRE

En 2025, Bordeaux Métropole s'inscrit dans un paysage réglementaire français en pleine structuration autour de l'intelligence artificielle, à l'intersection du droit européen et des initiatives nationales. Ce paysage s'articule autour des textes suivants :

Le Règlement européen sur l'IA (IA Act)

Entré en vigueur le 1er août 2024, l'IA Act constitue la première législation dédiée à l'encadrement des systèmes d'IA, s'appliquant à l'échelle européenne. Il repose sur une approche par les risques, classant les systèmes en quatre catégories : risque minimal, limité, élevé et inacceptable. Les systèmes à risque inacceptable (ex. : reconnaissance faciale en temps réel) sont interdits depuis février 2025. Les systèmes à haut risque (application dans les métiers de la santé, éducation, justice, sécurité, etc.) doivent respecter des exigences fortes en matière de transparence, de robustesse ou encore de supervision humaine.

Le Règlement général sur la protection des données (RGPD) et la CNIL : une articulation renforcée

Le Règlement général sur la protection des données (RGPD) est un pilier fondamental de la régulation de l'IA. Il impose des obligations en matière de traitement automatisé, de transparence algorithmique, de droit à l'explication et de protection des données personnelles. La CNIL, autorité de contrôle en France, joue un rôle central dans l'interprétation et l'application conjointe du RGPD et de l'IA Act.

Lois nationales et initiatives complémentaires

Au-delà du droit européen, la France a renforcé son cadre législatif avec plusieurs textes spécifiques:

- Loi pour une République numérique (2016), qui pose les bases de l'ouverture des données publiques (open data), essentielle pour l'entraînement des IA,
- Loi Informatique et Libertés, modifiée pour intégrer les exigences du RGPD et encadrer les traitements algorithmiques dans le secteur public,
- Loi sur la régulation numérique qui prévoit des obligations de transparence algorithmique pour les plateformes numériques et les services publics utilisant l'IA.

PRINCIPE N°1

DES IA MAÎTRISÉES POUR GARANTIR SÉCURITÉ, AUTONOMIE ET ÉTHIQUE

Bordeaux Métropole s'engage à :

Engagement n°1

Analyser la maîtrise technique et informationnelle des IA

Bordeaux Métropole s'engage à mener, pour chaque IA dont le développement ou le déploiement est envisagé, une analyse de sa gouvernabilité et de sa maîtrise technique. Il s'agira notamment de vérifier si la collectivité dispose d'informations suffisantes sur les algorithmes et les données utilisées. Les paramètres des systèmes devront être connus et la supervision des résultats rendue possible.

Engagement n°2

Garantir l'auditabilité des IA dans la durée

En amont du développement ou de l'acquisition d'une IA, Bordeaux Métropole s'engage à examiner et garantir l'auditabilité et à la réversibilité de celle-ci. Dans cette même optique, à maintenabilité, stabilité égales et à coûts comparables, la collectivité privilégiera des solutions open source. Afin de conserver la maîtrise, Bordeaux Métropole s'engage pour les IA installées à organiser un audit périodique des résultats et d'analyser d'éventuels écarts.

Engagement n°3

Protéger les données des citoyens, agents, de la collectivité et du territoire

Bordeaux métropole s'engage à protéger les données des citoyens, des agents publics et du territoire. Concernant les agents et citoyens, l'utilisation par des IA ne pourra s'envisager que dans le strict respect du règlement général sur la protection des données personnelles et des recommandations de la CNIL en matière d'IA. Concernant la collectivité et le territoire, l'accaparement par des systèmes d'IA sera contrôlé, notamment en termes de propriété intellectuelle. De façon globale, les données exploitées se limiteront au strict nécessaire.

Engagement n°4

Vérifier les biais éventuels induits par les IA exploitées

Bordeaux Métropole s'engage à vérifier que les traitements réalisés par des IA mises en œuvre n'affectent pas le respect de principes d'équité, d'inclusion, de justice sociale et de non-discrimination. L'attention sera portée tant sur les phases d'entraînement que d'exploitation.

PRINCIPE N°2

DES IA UTILES POUR GARANTIR UN EQUILIBRE POSITIF ENTRE APPORTS ET IMPACTS

Bordeaux Métropole s'engage à :

Engagement n°5

Etudier systématiquement l'équilibre apport de valeur / impacts

Bordeaux Métropole s'engage à mener systématiquement, en amont de tout potentiel projet de mise en service d'IA, une analyse des bénéfices qui sont attendus à court et à long terme au regard des impacts s'y rattachant. Les IA mises en services devront présenter des apports de valeurs supérieurs aux impacts économiques ou écologiques identifiés.

Engagement n°6

Recourir à des IA les moins impactantes possibles

Certains systèmes d'IA, entraînés localement avec des volumes de données réduits, ont un impact plus limité sur l'environnement. Bordeaux Métropole s'engage à examiner la possibilité de recourir à ces modèles et à minimiser à chaque étape les impacts.

Engagement n°7

Encadrer les phases d'expérimentation

La vérification de l'utilité d'un système d'IA peut nécessiter de passer par une phase d'expérimentation. Cette phase peut aussi s'avérer nécessaire pour favoriser l'apprentissage interne, le développement de cas d'usages pertinents et l'évaluation concrète de leur valeur ajoutée pour les politiques publiques. Bordeaux Métropole s'engage à limiter dans le temps et réaliser systématiquement un bilan documentant ces phases expérimentales. Toute expérimentation non probante conduira à l'abandon du projet.

Engagement n°8

S'assurer de la viabilité et l'apport de valeur d'une IA dans la durée

Bordeaux Métropole s'engage à vérifier dans la durée la viabilité et l'apport de valeur d'une IA en service pour s'assurer qu'elle soit utile, utilisable, utilisé. Sera également prévu dès le départ de limiter dans le temps des projets dont la valeur est temporaire ou fortement décroissante.

PRINCIPE N°3

DES IA DE CONFIANCE AU SERVICE DES CITOYENS ET DU TERRITOIRE

Bordeaux Métropole s'engage à :

Engagement n°9

Signaler et expliquer le recours à l'IA en toute transparence

Bordeaux Métropole s'engage à signaler le recours à l'IA dans ses processus de gestion et à garantir une application strictement restreinte au périmètre d'usage préalablement défini. Cette information est portée à connaissance des citoyens et agents par différents canaux (information sur des formulaires, site internet...) en précisant le type d'IA utilisé, ses principaux paramètres, les données utilisées et les résultats attendus. La métropole s'engage également à publier annuellement un registre des IA utilisés.

Engagement n°10

Garantir une alternative humaine au numérique dans les canaux de communication avec le citoyen

Dans la continuité de sa stratégie Numérique Responsable, car l'accès de tous aux services publics est une nécessité, Bordeaux Métropole garantit une alternative humaine au numérique (dont l'intelligence artificielle) dans chaque canal de communication qui en sera doté. De manière générale, la collectivité sera particulièrement attentive à ce qu'aucune initiative n'induisse, in fine, une accentuation de la fracture numérique.

Engagement n°11

Accompagner un usage éclairé de l'IA par les citoyens

Bordeaux Métropole s'engage à déployer des contenus et à mener des actions d'information et de sensibilisation sur le territoire afin de faciliter la compréhension des impacts des IA sur les processus d'apprentissage (éducation des jeunes publics, développement des compétences), sur la santé mentale ainsi que certaines conséquences démocratiques (fake news, manipulation de l'opinion publique, etc.) et impacts environnementaux induits.

Engagement n°12

Associer des représentants des citoyens au suivi des orientations concernant l'IA

Pour s'assurer de conserver un alignement fort avec les attentes et besoins des citoyens, Bordeaux Métropole s'engage à associer via le comité consultatif du numérique des représentants de citoyens via un suivi annualisé autour de : un bilan des avancées, une projection vers les perspectives envisagées par la collectivité en matière d'IA sur l'année à venir et l'alignement aux engagements posés dans le présent document.

PRINCIPE N°4

DES IA EN APPUI DES AGENTS POUR RENFORCER LE SERVICE PUBLIC

Bordeaux Métropole s'engage à :

Engagement n°13

Former et préparer les agents à une utilisation éclairée des IA

Bordeaux Métropole s'engage à sensibiliser et former les agents qui utiliseront des IA à leur fonctionnement mais aussi à leur contrôle. La compréhension fine des impacts, des opportunités et des conditions de mise en œuvre des IA permettra aux agents d'avoir un usage éclairé et maîtrisé de ces systèmes.

Engagement n°14

Mettre l'agent au cœur de la construction d'IA

Pour garantir des solutions utiles, utilisées et utilisables et soutenir l'identification d'impacts sur les métiers et activités des agents, Bordeaux Métropole s'engage à associer ceux/celles-ci à leur construction dès les phases les plus amont d'étude ou expérimentation.

Engagement n°15

Évaluer l'impact des IA sur les métiers et l'organisation du service public

Bordeaux Métropole s'engage à évaluer les impacts de l'introduction et l'utilisation d'IA sur les métiers et l'organisation du service public. Elle prendra appui sur un cadre méthodologique partagé avec les représentants du personnel.

Engagement n°16

Utiliser l'IA en tant qu'aide à la décision et non de prise de décision automatisée

Bordeaux Métropole s'engage à favoriser l'usage de l'IA par ses agents comme appui à des décisions individuelles ou collectives. Cet engagement illustre la volonté de la collectivité d'utiliser l'IA comme facilitateur, pour soutenir ses activités quotidiennes et de conserver une décision humaine.

PILOTAGE ET GOUVERNANCE

Le respect des 4 principes et 16 engagements exposés repose sur la mise en place d'une gouvernance dédiée au suivi des enjeux liés à l'IA au sein de la collectivité. Celle-ci visera à évaluer les opportunités, prévenir les risques, assurer la transparence des processus, et garantir un recours éthique et responsable à l'IA au sein de Bordeaux Métropole.

Le comité de pilotage numérique métropolitain d'ores et déjà établi dans le cadre des 7 Orientations Numériques Partagées de la métropole, intégrera à son périmètre les enjeux liés à l'Intelligence artificielle. En tant qu'organe politique, ce comité sera, annuellement, le lieu d'information et d'arbitrage des élus, à travers notamment la présentation d'éléments de bilan et de prospective stratégique.

La mise en œuvre de ce cadre autour de l'IA s'appuiera également sur une organisation administrative dédiée par la mobilisation d'un comité opérationnel IA qui assurera trimestriellement le suivi et la mise en œuvre d'une feuille de route opérationnelle. Il sera composé d'agents experts représentant les différents métiers de la métropole et des communes.

Consciente des enjeux économiques, sociaux et démocratiques soulevés par l'IA, Bordeaux Métropole souhaite inscrire sa démarche dans une logique d'ouverture et de dialogue continue.

Les instances de représentation du personnel (CST et FSSCT) seront associées ou saisies en amont de tout déploiement définitif d'un système d'IA dès lors que celui-ci pourrait impacter l'organisation du travail, les missions des agents ou les conditions de travail. Un rapport annuel leur sera communiqué pour réaliser un bilan des expérimentations réalisées et IA déployées.

Le comité consultatif du numérique, établi dans le cadre des 7 Orientations Numériques Partagées de Bordeaux Métropole et mobilisant des représentants de la société civile (entreprises, recherches, universités), sera associé annuellement au bilan des avancées de Bordeaux Métropole, perspectives envisagées et à l'alignement aux engagements posés dans le présent document.

GLOSSAIRE

Algorithme : description dans un langage formel d'une suite finie et ordonnée de processus qui, à partir de données en entrée, livre des données en sortie en un temps fini, en vue d'un objectif prédéterminé.

Biais : erreurs systématiques dans les systèmes d'IA pouvant conduire à des résultats injustes. Ils peuvent être induit par l'entraînement ou par les jeux de données de son exploitation. Il est important de les identifier et de les corriger pour maintenir la confiance des citoyens.

Données : information stockée dans un format qui permet son utilisation par un programme. Les données peuvent prendre de nombreuses formes : chiffres, texte, coordonnées géographiques, etc.

Données personnelles : toute information relative à une personne physique identifiée ou qui peut être identifiée, directement ou indirectement, par référence à un numéro d'identification ou à un ou plusieurs éléments qui lui sont propres.

Fracture numérique : écart au sein de la population face à l'accès aux technologies numériques, que ce soit par manque de matériel et service, ou par manque de connaissances et de compréhension de ces technologies.

Hallucination : contenu incorrect, inopportun ou fictif généré par une Intelligence artificielle générative et présenté de manière factuelle, comme si les informations étaient authentiques ou en adéquation avec la requête d'origine.

Intelligence artificielle (IA) : ensemble des théories et des techniques développant des programmes informatiques complexes capables de simuler certains traits de l'intelligence humaine, comme le raisonnement ou l'apprentissage.

Intelligence artificielle générative : catégorie d'intelligence artificielle en capacité de créer de l'information de manière automatique à partir de grands volumes de données existantes sur lesquelles elle est entraînée.

Modèle : représentation mathématique ou computationnelle d'un processus ou d'un système utilisé pour faire des prédictions ou des décisions.

Open source : se dit d'une solution conçue pour être accessible au public : n'importe qui peut voir, modifier et distribuer le code ou les paramètres à sa convenance.

Protection des données : ensemble des mesures visant à protéger les données personnelles et publiques contre les accès non autorisés, les modifications ou les destructions non souhaitées.

Règlement général sur la protection des données (RGPD) : règlement européen qui encadre le traitement des données personnelles sur le territoire de l'Union européenne. Il renforce et unifie la protection des données pour les individus au sein de l'UE.

Transparence : caractéristique des systèmes et processus qui sont ouverts et compréhensibles pour les utilisateurs et les parties prenantes, favorisant ainsi la confiance et l'imputabilité.