Approvisionnement responsable en solutions d'intelligence artificielle (IA)

Fiche de bonnes pratiques environnementales

ceimia futureearth IVADO

Résumé exécutif

L'intelligence artificielle (IA) génère des impacts environnementaux significatifs tout au long de son cycle de vie, les principaux étant : les émissions de gaz à effet de serre liées à la consommation d'énergie, la consommation d'eau, l'épuisement des ressources minérales, l'occupation des sols, et la toxicité liée à une mauvaise gestion des déchets électroniques. Les impacts, à la fois directs et indirects, sont croissants et directement liés aux choix technologiques faits par les organisations qui acquièrent ou déploient des solutions d'IA. Le secteur public, en position d'acheteur important, dispose d'un levier concret pour orienter le marché vers des pratiques plus responsables.

Ce document s'adresse aux spécialistes en approvisionnement et aux équipes impliquées dans l'acquisition de solutions d'IA au sein des organismes publics — et, plus largement, à toute organisation souhaitant intégrer des critères environnementaux dans ses processus d'achat. Il présente 48 bonnes pratiques organisées en six étapes séquentielles, allant de la définition du besoin jusqu'à la mesure de la performance :

1. Définir et remettre en question le besoin pour l'IA
2. Établir un cadre de gouvernance responsable
3. Anticiper les impacts potentiels de l'IA
4. Préparer le cahier des charges
5. Former les équipes au cadre d'usage
6. Mesurer et divulguer la performance de la solution

L'accessibilité de ces pratiques varie selon les moyens, la maturité organisationnelle et la littératie numérique propres à chaque organisation. **De ce point de vue, ce document ne vise pas la perfection, mais la progression :** chaque pratique mise en œuvre représente un gain réel. L'objectif, ici, est d'avancer à la mesure de ses capacités, en priorisant les actions à fort impact et faible effort, puis en élargissant progressivement la démarche.

Une matrice synthétique et une liste de contrôle sont disponibles en annexe. La matrice présente l'ensemble des bonnes pratiques avec, pour chacune, un indice de facilité de mise en œuvre et un indice de réduction potentielle des impacts environnementaux. Les organisations qui souhaitent établir un ordre de priorité sont encouragées à s'y référer pour identifier les premières actions les plus adaptées à leur contexte.

À cet égard, chaque étape du processus comporte des pratiques dont le rapport effort-impact est particulièrement favorable — autant de points de départ concrets pour les organisations qui souhaitent progresser sans se sentir submergées. En voici quelques exemples :

Définir le besoin (§ 3.1)

Énoncer et consigner clairement la problématique à laquelle on cherche à répondre (§ 3.1.1) est une action simple, mais structurante : elle évite les acquisitions superflues, qui sont elles-mêmes une source d'impact évitable. Évaluer les effets positifs et négatifs de différentes solutions avant de se fixer sur l'IA (§ 3.1.4) est également possible et permet d'éviter de déployer une technologie là où une solution plus légère suffirait.

Établir un cadre de gouvernance (§ 3.2)

Établir un registre clair des responsabilités — qui est responsable de quoi dans la chaîne d'acquisition et de déploiement — (§ 3.2.4) est une pratique peu exigeante techniquement, mais déterminante pour assurer un suivi cohérent dans la durée.

Anticiper les impacts (§ 3.3)

Vérifier que la solution d'IA envisagée est compatible avec les équipements informatiques existants, afin d'éviter un renouvellement matériel prématuré (§ 3.3.2), est un réflexe accessible à toute organisation, avec un potentiel de réduction d'impact élevé. Procéder à une évaluation qualitative des impacts environnementaux dès la phase de conception (§ 3.3.1) n'exige pas de ressources spécialisées importantes et pose les bases d'une prise de décision éclairée.

Préparer le cahier des charges (§ 3.4)

Exiger que la solution soit hébergée dans une région au mix électrique faiblement carboné (§ 3.4.5) est un critère simple à formuler dans un appel d'offres, sans nécessiter d'expertise technique approfondie, et dont le levier environnemental est parmi les plus élevés de toute la fiche.

Former les équipes (§ 3.5)

Inviter les personnes utilisatrices à minimiser le nombre de requêtes soumises à l'IA (§ 3.5.3) et les sensibiliser aux fonctionnalités les plus énergivores — comme la génération d'images ou les tâches de raisonnement complexe — (§ 3.5.4) sont des actions immédiatement accessibles, sans prérequis, dont l'effet cumulatif à l'échelle d'une organisation peut être considérable.

Mesurer et divulguer la performance (§ 3.6)

S'assurer que le modèle d'IA reste sobre et performant une fois déployé, en définissant des indicateurs de suivi avec le fournisseur (§ 3.6.1), et prévoir un processus de remise en question du besoin au fil du temps pour éviter la dépendance à des outils devenus surdimensionnés (§ 3.6.3) sont deux pratiques d'effort modéré qui ancrent la démarche dans la durée.

Crédits

Coordination

IVADO
Antoine Congost

Future Earth
Nilufar Sabet-Kassouf

CEIMIA
Antoine Glory, Stefan Janusz, Stephanie King

Rédaction

CIRAIG
Lucie Leignier et Flavien Binet

ECPAR
Laurence Pageau et Aure Adell

Comité consultatif

Deborah Ahenkorah (Université Concordia), Elise Bastille-Lavigne (CEST), Soumaya Cherkaoui (Polytechnique Montréal), Martin Deron (Chemins de Transition), Martial Kiemde (Gouvernement du Québec), Jean-Noé Landry (IRIU), Stéphanie H. Leclerc (Université McGill), Alexandre Thève (Davidson Consulting), Patrice Roy (Gouvernement du Québec), Nabila Yasmine Saidji (Gouvernement du Québec), Karim Jawad (MELCCPBFP)

Graphisme

Stéphanie Hauschild, avec le soutien de l'équipe des communications d'IVADO.

Ce document a été élaboré en collaboration avec le Gouvernement du Québec.

Table des matières

Abréviations et acronymes **5**

Introduction **6**

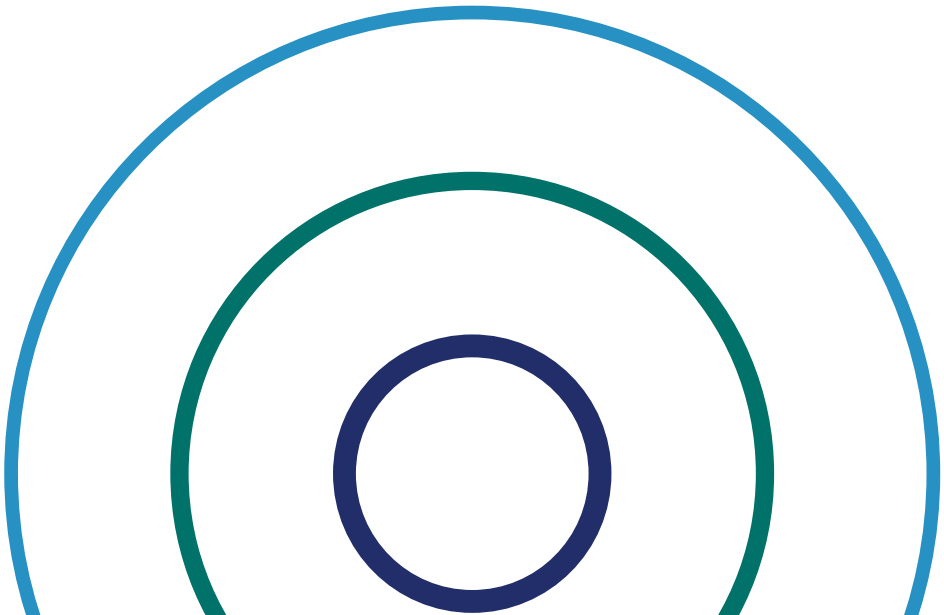
1. Comprendre l'IA et ses impacts **7**
1.1. Définition et concepts de l'IA 7
1.2. Cycle de vie de l'IA 7

2. Impacts environnementaux et sociaux de l'IA **9**
2.1. Impacts environnementaux directs de l'IA 9
2.2. Impacts environnementaux indirects de l'IA 12
2.3. Impacts sociaux et éthiques de l'IA 15

3. Bonnes pratiques d'approvisionnement en solutions d'IA **17**
3.1. Définir et remettre en question le besoin pour l'IA 17
3.2. Établir un cadre de gouvernance responsable 21
3.3. Anticiper les impacts potentiels de l'IA 23
3.4. Préparer le cahier des charges 25
3.5. Former les équipes au cadre d'usage 35
3.6. Mesurer et divulguer la performance de la solution 36

Bibliographie **38**

Annexes **43**
Matrice 43
Liste de contrôle 48
Références 53



Abréviations et acronymes

ACV	Analyse du cycle de vie
AIE	Agence internationale de l'énergie
CEST	Commission de l'Éthique en Science et en Technologie
CNIL	Commission Nationale de l'Informatique et des Libertés
CPU	*Central Processing Unit* (unité centrale de traitement)
CUE	*Carbon Usage Effectiveness* (efficacité de l'utilisation du carbone)
EIA	Évaluation d'incidence algorithmique
GES	Gaz à effet de serre
GPU	*Graphics Processing Unit* (processeur graphique)
HDD	*Hard Disk Drive* (disque dur)
IA	Intelligence artificielle
IAG	Intelligence artificielle générative
ITU	Union internationale des télécommunications
ISO	Organisation internationale de normalisation
LLM	*Large Language Model* (grand modèle de langage)
OCDE	Organisation de coopération et de développement économiques
OQLF	Office québécois de la langue française
ONU	Organisation des Nations Unies
PME	Petites et moyennes entreprises
PUE	*Power Usage Effectiveness* (efficacité énergétique)
RAG	*Retrieval Augmented Generation* (génération augmentée par récupération)
REF	*Renewable Energy Factor* (facteur d'énergie renouvelable)
SSD	*Solid State Drive* (disque SSD)
SLM	*Small Language Model* (petit modèle de langage)
TI	Technologie de l'Information
WUE	*Water Usage Effectiveness* (efficacité d'utilisation de l'eau)

Introduction

L'***Énoncé de principes pour une utilisation responsable de l'intelligence artificielle par les organismes publics*** constitue l'annexe prévue par l'Arrêté 2025-02 du 3 décembre 2025 (Gouvernement du Québec, 2025). Selon le principe de la durabilité (#6) qui y est énoncé, l'utilisation responsable des systèmes d'intelligence artificielle (IA) doit s'inscrire dans la recherche d'un développement durable. On y précise que les principes de l'énoncé s'appliquent même lorsqu'un organisme public a recours à des fournisseurs et à des prestataires pour des systèmes d'intelligence artificielle ou en lien avec ceux-ci, ce qui fait du respect de l'énoncé de principes un enjeu d'approvisionnement, puisqu'il faut que la recherche d'un développement durable soit également assurée par les fournisseurs.

Afin de soutenir les organismes publics ou toute autre organisation intéressée par l'acquisition de solutions d'IA de façon durable, le présent document présente des bonnes pratiques d'approvisionnement responsable visant à atténuer les conséquences environnementales négatives du recours à l'IA. En effet, bien que cette technologie offre des possibilités importantes en termes d'efficience technologique, ce qui explique que le gouvernement du Québec a adopté sa *Stratégie d'intégration de l'IA dans l'administration publique 2021-2026* (Secrétariat du Conseil du trésor, 2021), l'IA amène en revanche des risques environnementaux tout au long de son cycle de vie qu'il convient de prendre en considération pour respecter le principe de la durabilité. Comme le mentionne la Stratégie, "la question du développement durable est au cœur des réflexions entourant l'IA, laquelle doit permettre une meilleure utilisation des ressources dans une perspective de durabilité, que ce soit au niveau de la prise en compte de l'impact environnemental dans les applications de l'IA, ou encore de l'empreinte écologique des infrastructures technologiques nécessaires" (Secrétariat du Conseil du trésor, 2021, p.9).

Les risques éthiques, sociaux, opérationnels et de gouvernance associés à l'IA sont désormais bien documentés, et plusieurs cadres de référence proposent des balises pour guider une adoption responsable. Toutefois, les impacts environnementaux des solutions d'IA sont moins souvent abordés, bien qu'ils soient significatifs, croissants et directement liés aux choix technologiques faits par les organisations. Ce document vise à combler ce manque. Il s'adresse principalement aux spécialistes en approvisionnement et aux équipes impliquées dans l'acquisition de solutions d'IA développées par des fournisseurs externes. Son objectif est d'offrir une information claire et actionnable afin de soutenir des décisions d'achat qui tiennent compte des enjeux environnementaux sur l'ensemble du cycle de vie de l'IA. Les enjeux sociaux et éthiques qui sont en lien avec l'information fournie dans le document sont abordés sommairement.

Bien que ce document soit adapté au contexte des marchés publics, l'information qui y est présentée est également pertinente pour le secteur privé.

La première section du document est dédiée à la compréhension de l'IA, de ses impacts environnementaux et sociaux. La seconde section fournit des bonnes pratiques d'approvisionnement à adopter pour réduire ces impacts. Elle peut servir de base consultative pour définir, par la suite, des critères et des pratiques d'approvisionnement responsable concrets.



1. Comprendre l'IA et ses impacts

1.1 Définition et concepts de l'IA

L'Organisation de coopération et de développement économiques (OCDE) définit l'IA comme "un système automatisé qui, pour des objectifs explicites ou implicites, déduit, à partir d'entrées reçues, comment générer des résultats en sortie tels que des prévisions, des contenus, des recommandations ou des décisions" (OCDE, 2025, p.7-8). Cette définition est celle utilisée par le Gouvernement du Québec (2025). L'Office québécois de la langue française (OQLF) précise également qu'un système d'IA est conçu pour simuler les facultés cognitives de l'intelligence humaine (OQLF, 2026).

Si l'IA a évolué rapidement ces dernières années, ses origines remontent à plusieurs décennies. Dans les années 1960 apparaît l'*IA traditionnelle* (ou *symbolique*), fondée sur des règles fixes préétablies par la programmation (statistiques, arbres de décision, systèmes experts, etc.). Dans les années 1980-1990, l'*apprentissage automatique* (ou *machine learning*) marque un tournant : le système apprend à partir de données plutôt que de règles codées, en s'adaptant et en s'améliorant. Ceci donne notamment naissance à l'*IA prédictive*, capable de prédire des évènements futurs en analysant les tendances passées. Au début des années 2010 émerge l'*apprentissage profond* (ou *deep learning*), reposant sur des réseaux de neurones organisés en couches. Cette approche a transformé de nombreux domaines, comme la vision par ordinateur et le traitement du langage naturel, mais a, en revanche, rendu les modèles de plus en plus opaques en matière d'explicabilité.

Le secteur connaît depuis un rapide développement. Ces progrès conduisent, dans les années 2020, à l'essor de l'*IA générative* (IAG), capable de créer de nouveaux contenus (textes, images, vidéos) à partir de données existantes (HEC Montréal, 2024). Depuis l'arrivée de l'IA générative, l'IA a connu un déploiement et une adoption très importants. Aujourd'hui, plus de la moitié (50%) des internautes au Québec ont déjà utilisé l'IA, dont une part importante l'utilise quotidiennement (Académie de la transformation numérique, 2025). Ces tendances pourraient se maintenir dans les prochaines années, avec l'arrivée d'un nouveau type d'IA, l'*IA agentique*, capable de raisonner et d'interagir de façon autonome avec son environnement, avec peu de supervision humaine.

En résumé, les usages de l'IA ne cessent de se transformer, et son taux d'adoption est croissant. Cependant, cette croissance s'accompagne d'impacts environnementaux et éthiques qui pourraient ne pas être soutenables dans les années à venir (The Shift Project, 2025). Le reste de cette section présente le cycle de vie de l'IA. La section 2 présente ses enjeux environnementaux et éthiques, en se concentrant sur l'apprentissage automatique et l'apprentissage profond.

1.2 Cycle de vie de l'IA

Le cycle de vie d'un système d'IA[1] peut être décrit du point de vue technologique ou physique. Le cycle de vie technologique fait référence aux étapes nécessaires au déploiement d'un système d'IA, et le cycle de vie physique fait référence au cycle de vie de l'infrastructure physique sur laquelle repose un système d'IA.

1.2.1 Cycle de vie technologique

Le cycle de vie technologique d'un système d'IA se compose de plusieurs étapes :

1. planification et conception ;
2. collecte et traitement des données ;
3. construction du ou des modèle(s) et/ou adaptation du ou des modèle(s) existant(s) à des tâches spécifiques ;
4. test, évaluation, vérification et validation ;
5. mise à disposition pour utilisation/déploiement ;
6. exploitation et suivi ; et
7. mise hors service (OCDE, 2025).

1. Ici, le terme "système d'IA" englobe uniquement les étapes du cycle de vie qui permettent d'exploiter l'intelligence artificielle. L'IA peut parfois être intégrée à d'autres services numériques, ou à des biens non numériques. Elle peut par exemple être intégrée dans un site web (un autre service numérique) ou dans une voiture autonome (un bien non numérique). Le cycle de vie de ces derniers ne fait cependant pas l'objet du présent document.

Lorsqu'il est analysé sous l'angle des impacts environnementaux, ce cycle de vie technologique peut généralement être résumé en trois grandes étapes :

- L'**entraînement,** qui a lieu au cours de la construction du modèle. L'entraînement permet au système d'IA de construire et d'adapter son modèle à partir des données apprises. Les méthodes d'entraînement diffèrent entre l'apprentissage automatique et l'apprentissage profond.
- L'**inférence,** qui a lieu en phase d'exploitation et de suivi. Cela renvoie à l'utilisation du système d'IA par les personnes utilisatrices à la suite de son déploiement. Au cours de l'inférence, des instructions sont transmises au système d'IA afin qu'il génère une réponse. Ces instructions sont appelées *requêtes* (ou *prompts*).
- Le **réentraînement,** qui a également lieu en phase d'exploitation et de suivi. Pour maintenir ses performances en phase d'utilisation, un modèle doit être réentraîné à partir de nouvelles données. Ce réentraînement peut s'effectuer de manière continue ou à intervalles réguliers, au cours de l'exploitation du modèle.

Dans la suite du document, le terme "(ré) entraînement" sera utilisé pour désigner à la fois l'entraînement et le réentraînement du modèle.

Finalement, l'IA fait appel à des volumes importants de données tout au long de son cycle de vie : des données pour développer le modèle lors de l'entraînement, d'autres pour mesurer ses performances, puis de nouvelles, recueillies pendant son utilisation, pour le mettre à jour régulièrement. Ces données doivent souvent être triées et préparées par une intervention humaine (annotation, suppression des doublons, structuration, etc.) (Union internationale des télécommunications [ITU], 2026).

1.2.2. Cycle de vie physique

La réalisation de chacune des étapes du cycle de vie technologique (entraînement, inférence, réentraînement) décrite à la section 1.2.1 repose sur une infrastructure physique qui se compose de trois niveaux : les **centres de données**, les **réseaux** et les **terminaux utilisateurs**.

Les centres de données traitent et hébergent les données numériques. Cette infrastructure regroupe différents équipements informatiques assurant le calcul et le traitement des données, leur stockage, et la connectivité réseau. Le calcul et le traitement reposent sur des serveurs composés de *Central Processing Units* (CPUs) et/ou d'autres composants spécialisés, comme les *Graphics Processing Units* (GPUs) sur lesquels s'effectuent le (ré)entraînement et l'inférence du système d'IA[2]. Les données sont quant à elles conservées dans des baies de stockage, principalement composées de disques *Solid State Drive* (SSD) et *Hard Disk Drive* (HDD).

Pour assurer son bon fonctionnement, un centre de données est également composé d'équipements dits "non informatiques". Ceux-ci incluent le(s) système(s) de refroidissement, qui dissipent la chaleur des serveurs, et d'autres infrastructures auxiliaires (p. ex. bâtiment, générateurs et batteries de secours, vidéosurveillance) (Groupement CLIQ et al., 2026).

Les réseaux, constituant le deuxième niveau du numérique, assurent la transmission des données entre les centres de données et les terminaux utilisateurs. Les principaux types de réseaux sont les réseaux fixes (câbles, Wi-Fi, fibres optiques, etc.), qui offrent un accès localisé par connexion câblée, et les réseaux mobiles (3G, 4G, 5G, etc.) qui fonctionnent avec des antennes-relais et couvrent un large territoire (ADEME, 2024).

Finalement, les terminaux utilisateurs (téléphones intelligents, ordinateurs, écrans, etc.) constituent le dernier niveau du numérique. Ils permettent d'accéder et d'exploiter un service d'IA.

Chaque niveau de l'infrastructure physique nécessite des équipements et du matériel informatique. Le cycle de vie de ces derniers, incluant leur fabrication, leur distribution, leur utilisation et leur traitement en fin de vie, est responsable de nombreux impacts environnementaux décrits à la section 2.

2. Un système d'IA peut être hébergé dans un centre de données et être accessible grâce à un équipement distinct, mais il peut aussi être hébergé et fonctionner en interne de l'équipement (IA embarquée). Dans ce document, seule la première configuration est étudiée.

2. Impacts environnementaux et sociaux de l'IA

Les impacts environnementaux de l'IA peuvent être séparés en **impacts directs et indirects**. Les impacts directs sont ceux associés à l'existence matérielle de l'infrastructure physique qui supporte l'IA sur son cycle de vie (fabrication, distribution, utilisation, fin de vie). Les impacts indirects sont liés aux transformations des usages, des comportements et des systèmes économiques que l'utilisation de l'IA engendre (Commission de l'éthique en science et en technologie [CEST], 2024).

2.1. Impacts environnementaux directs de l'IA

L'empreinte environnementale d'un système d'IA est souvent dominée par l'étape d'inférence, suivie de l'entraînement et du réentraînement (Desroches et al., 2025; Berthelot et al., 2024). L'utilisation de l'IA est responsable de différents impacts environnementaux directs, dont les principaux sont présentés dans le tableau ci-contre, et détaillés davantage dans les sous-sections suivantes. Ces impacts directs sont négatifs, par définition.

Tableau : Résumé des impacts environnementaux directs de l'IA

Type d'impact direct	Conséquence
Consommation d'énergie et changements climatiques	Négative
Déplétion des ressources minérales	
Consommation d'eau	
Occupation des sols	
Toxicité des déchets électroniques	

Plusieurs facteurs technologiques influencent l'ampleur de ces impacts environnementaux. Ceux-ci sont discutés dans les bonnes pratiques d'approvisionnement en solutions d'IA (section 3).

2.1.1. Consommation d'énergie et changements climatiques

Lors de leur fonctionnement, les centres de données consomment de grandes quantités d'électricité. Pour l'année 2025, la consommation mondiale d'électricité des centres de données était estimée à 448 TWh. Comme l'indique l'Organisation des Nations Unies (ONU), "si les centres de données formaient un pays, leur consommation d'électricité les classerait au 11e rang mondial" (Aczel et al., 2026, p.10). Cette même année, la part de la consommation d'électricité attribuable aux serveurs d'IA spécifiquement était déjà significative (environ 20%). Au rythme actuel, la consommation d'électricité des centres de données pourrait doubler d'ici 2030, l'IA comptant pour 40% de cette consommation (Aczel et al., 2026).

Cette consommation s'explique principalement par la phase d'utilisation des CPUs et des GPUs, qui domine les impacts sur l'ensemble de leur cycle de vie (environ 70 à 90% pour l'impact sur les changements climatiques) (Falk et al., 2025).

Les systèmes de refroidissement nécessaires pour dissiper la chaleur des serveurs constituent également une part significative de la consommation énergétique des centres de données (entre 7 et 30%) (Agence internationale de l'énergie [AIE], 2025). En effet, le refroidissement nécessaire pour évacuer la chaleur des serveurs est souvent assuré par des systèmes de climatisation énergivores.

Finalement, bien que moins énergivore (environ 5% de la consommation électrique d'un centre de données) (AIE, 2025), le stockage des données pourrait avoir un impact significatif sur les changements climatiques dans les années à venir. En effet, les contraintes financières et temporelles liées à leur collecte et à leur traitement (annotation, suppression des doublons, structuration) font qu'elles sont rarement supprimées. Leur volume croît donc progressivement, augmentant davantage les besoins en stockage et en électricité (Sveinsdottir et al., 2020).

Aujourd'hui, la majorité des centres de données sont situés dans des régions où l'électricité est fortement carbonée. Plus précisément, 90% des centres de données spécialisés pour l'IA sont concentrés dans deux pays, les États-Unis et la Chine, ce qui contribue aux impacts liés aux changements climatiques (Aczel et al., 2026). En considérant l'intensité carbone du mix électrique mondial moyen, les émissions générées par l'opération des centres de données spécialisés pour l'IA seraient d'environ 40 millions de tonnes de CO_2 équivalent en 2025 (Aczel et al., 2026) - excluant les autres émissions sur le cycle de vie des centres de données, ainsi que celles des réseaux et des terminaux utilisateurs - ce qui correspond à près de la moitié des émissions de gaz à effet de serre (GES) de la province du Québec (Environnement et Changement climatique Canada, 2026).

De plus, la croissance rapide des besoins énergétiques liés à l'IA pourrait engendrer des conséquences significatives sur la compétition pour l'énergie. D'une part, cette demande risque de mobiliser une part importante de l'électricité décarbonée disponible, au détriment d'autres secteurs pour lesquels l'électrification constitue un levier de décarbonation essentiel (p. ex. le secteur des transports, de l'industrie). D'autre part, afin de répondre à la hausse de la demande électrique, plusieurs pays développent ou relancent massivement des infrastructures de production d'origine fossile (gaz ou charbon), au détriment de leurs engagements climatiques (The Shift Project, 2025).

Finalement, en plus de la consommation d'énergie lors de l'utilisation, la fabrication des équipements informatiques engendre des émissions de GES significatives. Au niveau des centres de données, la production de ses composants (serveurs, baies de stockage, alimentation, etc.) représente 20% de l'empreinte carbone totale d'un centre de données (The Shift Project, 2025; Schneider Electric, 2023). Cette proportion est encore plus marquée pour les terminaux utilisateurs : dans les régions où le mix électrique est faiblement carboné, comme au Québec, la fabrication est responsable de la majorité (jusqu'à 85%) des émissions de GES générées sur le cycle de vie du terminal (ADEME, 2025; Istrate et al., 2024). Ceci s'explique principalement par la production des semi-conducteurs qui composent les puces électroniques. Aujourd'hui, Taïwan assure 90% de la production des puces électroniques destinées aux applications d'IA (Venne J.F., 2025). Or, le procédé de fabrication est particulièrement énergivore et le mix électrique de Taïwan est fortement carboné[3], entraînant d'importantes émissions de GES. Ces émissions augmentent avec la demande et la complexité technologique croissante des puces électroniques (Oueslati Y., 2026; Roussilhe G., 2024). Le processus de fabrication implique également l'utilisation de gaz fluorés à fort potentiel de réchauffement climatique (Roussilhe et al., 2024).

3. L'intensité carbone du mix électrique de Taïwan est d'environ 500 gCO_2eq/kWh, soit environ 14 fois plus élevée que celle du Québec (Electricity Maps, 2026). Cette métrique traduit les émissions de GES qui résultent de la production d'un kilowattheure d'électricité.

2.1.2. Déplétion des ressources minérales

La fabrication des équipements informatiques contribue fortement à l'épuisement des ressources minérales. De nombreux métaux entrent dans la composition des équipements informatiques (serveurs, processeurs, batteries, puces électroniques, etc.). Certains de ces métaux, appelés *métaux rares*, sont consommés en petites quantités, mais sont indispensables pour le numérique, notamment le cobalt et le lithium (dans les batteries), l'indium (dans les écrans), le gallium et le germanium (dans les semi-conducteurs), ou encore les terres rares[4] (dans les aimants).

Or, une grande partie de ces métaux sont dits « critiques » (Ollion et al., 2024), leur approvisionnement risquant d'être restreint dans les années à venir (Bureau de Recherches Géologiques et Minières, 2025). La liste des métaux jugés les plus critiques pour le numérique varie d'une étude à l'autre en fonction des facteurs de criticité (p. ex., incertitudes sur la longévité des réserves connues, disponibilité des métaux, risques géopolitiques liés à leur approvisionnement). Le fait que le secteur de l'IA, et plus largement du numérique, utilise ces métaux critiques, et que cette utilisation devrait croître à l'avenir, augmente d'autant plus le risque d'épuisement de ces ressources.

2.1.3. Consommation d'eau

Un système d'IA consomme de l'eau à trois étapes de son cycle de vie : lors de la fabrication des puces électroniques composant les équipements informatiques (pour les terminaux, réseaux et centres de données), lors de la production de l'électricité qui alimente les centres de données et lors du refroidissement des équipements informatiques dans les centres de données[5] .

Cette consommation d'eau est très élevée. L'AIE estime qu'en 2023, les centres de données ont consommé environ 560 milliards de litres d'eau[6] à l'échelle mondiale (AIE, 2025), soit l'équivalent d'environ 227 000 piscines olympiques. Environ deux tiers de la consommation sont liés à la production de l'énergie, un quart aux systèmes de refroidissement et le reste à la fabrication des puces électroniques. Cette consommation pourrait doubler d'ici 2030, avec une contribution de plus en plus importante des systèmes de refroidissement (AIE, 2025).

Si les systèmes de refroidissement par climatisation ont longtemps été privilégiés, les technologies de refroidissement par eau se développent progressivement[7] en raison de leur consommation électrique nettement inférieure. Toutefois, cette réduction de la consommation électrique accroît la pression sur les aquifères et les risques de conflits d'usage liés à l'eau. En effet, l'eau prélevée pour le refroidissement est souvent potable et est parfois puisée dans des régions où les ressources hydriques sont déjà sous tension. Aujourd'hui, environ 45 % des centres de données dans le monde seraient situés dans des bassins fluviaux présentant un risque élevé de disponibilité en eau (The Shift Project, 2025). De plus, l'eau utilisée par les systèmes de refroidissement est rejetée en sortie sous forme gazeuse, par évaporation, ou sous forme liquide. Bien qu'une partie de l'eau liquide rejetée puisse retourner dans le milieu naturel, celle-ci est souvent de moindre qualité (eau contaminée par de la poussière, des produits chimiques et des minéraux, et température trop élevée [Spindler et al., 2024]), ce qui peut affecter les écosystèmes aquatiques.

Enfin, la consommation d'eau liée à la fabrication des équipements informatiques est aujourd'hui peu documentée, malgré des conséquences déjà observables. Par exemple, Taïwan a connu en 2021 un épisode de sécheresse sévère en raison d'un manque de précipitations et d'une utilisation accrue de l'eau pour la fabrication des semi-conducteurs (Conseil national du numérique et al., 2025).

2.1.4. Occupation des sols

Bien que peu étudiée actuellement, l'occupation des sols associée à l'IA constitue également une source d'impacts environnementaux.

4. Les terres rares sont un groupe de 17 métaux. Ce sont des métaux stratégiques qui sont indispensables à la fabrication des équipements numériques. Contrairement à ce que leur nom indique, ils ne sont pas forcément rares.

5. La consommation d'eau utilisée par les systèmes de refroidissement des centres de données est appelée *consommation directe*. L'eau consommée lors de la production de l'électricité qui alimente le centre de données et pour la fabrication des équipements informatiques est appelée *consommation indirecte* (The Shift Project, 2025).

6. Cette consommation d'eau correspond à la quantité d'eau prélevée qui n'est pas restituée à sa source d'origine après utilisation (p. ex. à travers l'évaporation) (AIE, 2025).

7. Il existe plusieurs types de systèmes de refroidissement. Certaines technologies fonctionnent par air (climatisation), d'autres par liquide (p. ex., immersion des serveurs dans un liquide de refroidissement diélectrique) ou par eau (en boucle ouverte ou fermée). Le *free cooling*, qui utilise l'air extérieur pour refroidir naturellement les serveurs, se développe également (The Shift Project, 2025).

D'une part, l'extraction des ressources naturelles nécessaires à la fabrication des équipements informatiques entraîne des effets importants autour des mines, tels que la déforestation, la pollution des eaux souterraines, la destruction des habitats et la perte de biodiversité (Gouvernement du Canada, 2025).

D'autre part, l'essor de l'IA entraîne la construction de davantage de centres de données. En 2026, environ 11 500 centres de données sont actifs dans le monde (Data Center Map, 2026), dont la superficie varie de quelques dizaines à des milliers de mètres carrés par centre. Ces vastes bâtiments contribuent à l'artificialisation des sols, favorisant la dégradation de la biodiversité environnante, en particulier lorsqu'ils s'implantent à proximité de zones naturelles sensibles, telles que les milieux humides (Dugas Bourdon, P., 2024).

L'occupation des sols associée aux centres de données provient également des infrastructures de production d'énergie (extraction des combustibles, infrastructures d'énergies renouvelables, etc.). En 2025, à l'échelle mondiale, 1 400 km^2 de sol étaient occupés par ces infrastructures pour alimenter les centres de données spécifiques à l'IA, soit environ 23 fois la superficie de Manhattan (Aczel et al., 2026). Si la production d'énergies renouvelables réduit l'empreinte carbone par rapport aux énergies fossiles, elle s'accompagne souvent d'une occupation des sols plus importante.

2.1.5. Toxicité des déchets électroniques

Finalement, les services d'IA contribuent à l'accumulation de déchets électriques et électroniques (processeurs, GPUs, etc.). Bien que les impacts environnementaux de la fin de vie soient moins importants que ceux associés à la fabrication et à l'utilisation, seulement 20% environ de ces déchets sont correctement collectés et recyclés à l'échelle mondiale (Baldé et al., 2024). Par conséquent, la majeure partie des déchets est mise en décharge, incinérée, ou traitée de manière informelle, ce qui pose des risques pour l'environnement et la santé humaine, puisque ces déchets contiennent des substances toxiques (plomb, mercure, cadmium, etc.) susceptibles de contaminer le sol et l'eau (Gouvernement du Canada, 2025). Or, les nouveaux usages rendus possibles par l'IA, conjugués à la multiplication de terminaux équipés de systèmes d'IA, pourraient accélérer l'obsolescence et le renouvellement des appareils, et ainsi accentuer ces enjeux (The Shift Project, 2025).

2.2. Impacts environnementaux indirects de l'IA

L'intelligence artificielle est souvent présentée comme un levier pour la transition environnementale, notamment à travers des applications visant à optimiser l'utilisation des ressources, améliorer l'efficacité énergétique ou soutenir la prise de décision environnementale. Cette approche est désignée par le terme *IA pour la durabilité*.

Cependant, il ne suffit pas qu'une application d'IA améliore l'efficacité d'un système. D'une part, ces bénéfices potentiels doivent être mis en perspective avec les impacts environnementaux directs propres aux systèmes d'IA décrits dans la section 2.1. D'autre part, il convient de considérer les impacts indirects de l'IA et de s'assurer que ceux qui sont négatifs soient minimisés.

L'IA engendre des impacts indirects de deuxième et de troisième ordre[8], souvent plus difficiles à quantifier, qui résultent des changements que l'IA induit dans les systèmes économiques, les modes de production et les comportements de consommation (CEST, 2024). Ces impacts peuvent être positifs (c'est-à-dire apporter un gain environnemental) ou négatifs (c'est-à-dire engendrer davantage d'impacts environnementaux).

Parmi les impacts indirects, l'*effet d'efficacité* constitue le principal effet positif : il correspond aux réductions d'impacts permises lorsque l'IA améliore l'efficacité d'un système, d'un procédé ou de l'utilisation des ressources. L'*effet de substitution*, qui survient lorsqu'une technologie remplace une autre, peut être positif ou négatif selon la nature de la technologie remplacée. Les *effets d'induction*, liés à la création de nouveaux usages rendus possibles par la technologie, peuvent eux aussi être positifs ou négatifs : ils augmentent généralement la demande de ressources numériques, mais peuvent aussi donner lieu à de nouveaux usages bénéfiques pour l'environnement (Lange et al., 2023).

8. Le deuxième ordre concerne l'utilisation des technologies numériques à des fins environnementales, et le troisième ordre renvoie aux conséquences du numérique sur les modes de vie (CEST, 2024).

Ces impacts s'accompagnent plus largement de l'*effet rebond direct* (aussi appelé *paradoxe de Jevons*). Ce terme désigne initialement un phénomène économique selon lequel une amélioration de l'efficacité d'une technologie entraîne une réduction de son coût d'utilisation, ce qui conduit à une augmentation de la demande et, finalement, de la consommation globale de ressources (Luccioni et al., 2025a). Toutefois, la notion d'effet rebond est aujourd'hui considérée de manière plus large et ne se limite pas uniquement à des mécanismes économiques. Elle renvoie plus généralement aux impacts environnementaux additionnels négatifs générés par l'introduction d'une technologie dans la société, à travers l'intensification des pratiques existantes qui compensent l'amélioration de l'efficacité de la technologie. Dans le cas de l'IA générative, l'augmentation rapide des usages ne s'explique pas uniquement par une diminution des coûts d'accès à l'IA, mais également par l'émergence de nouvelles possibilités d'utilisation (p. ex. génération de contenu, assistance conversationnelle, automatisation de tâches), qui encouragent une utilisation plus fréquente et intensive de l'IA. Cette intensification des usages entraîne, à son tour, une sollicitation plus importante des infrastructures numériques. Au-delà des centres de données, les réseaux de communication, et en particulier les réseaux mobiles, doivent ainsi faire face à une hausse du trafic de données et à des exigences croissantes en termes de performances techniques, ce qui accentue leur impact environnemental négatif (The Shift Project, 2025).

L'effet rebond peut se manifester aussi de manière indirecte : il est appelé *effet rebond indirect* et est lié à la réallocation des gains d'efficacité vers d'autres formes de consommation. Par exemple, les économies réalisées grâce à l'automatisation avec l'IA peuvent être réinvesties dans des activités plus émettrices, comme le voyage (effet rebond indirect négatif), alors qu'une application d'IA visant à sensibiliser à l'impact environnemental des transports pourrait, au contraire, encourager des comportements plus favorables à l'environnement (effet rebond indirect positif). L'effet rebond indirect peut notamment se manifester par des changements comportementaux, temporels ou géographiques (Luccioni et al., 2025a). Par exemple, le développement du télétravail assisté par des outils numériques peut réduire certains déplacements, tout en augmentant parallèlement les besoins en infrastructures numériques et la consommation énergétique résidentielle.

À une échelle d'analyse plus large, l'*effet d'échelle* désigne l'augmentation des impacts résultant de la diffusion massive d'une technologie : une technologie peut être plus efficace à l'unité, mais son déploiement à grande échelle peut néanmoins accroître son impact global en raison de la croissance du volume d'utilisation. Les *effets systémiques* vont plus loin et traduisent des transformations plus profondes des systèmes économiques et sociaux (c.-à-d. l'évolution des chaînes de valeur, des modes de production ou de consommation) et demeurent particulièrement complexes à anticiper et à quantifier (CEST, 2024).

Ainsi, les systèmes d'IA peuvent engendrer à la fois des impacts environnementaux indirects positifs et négatifs, résumés dans le tableau ci-dessous. L'équilibre entre ceux-ci est difficile à établir. Ces impacts indirects, tout comme les impacts directs, doivent être pris en compte lors de l'évaluation des bénéfices environnementaux d'une solution d'IA.

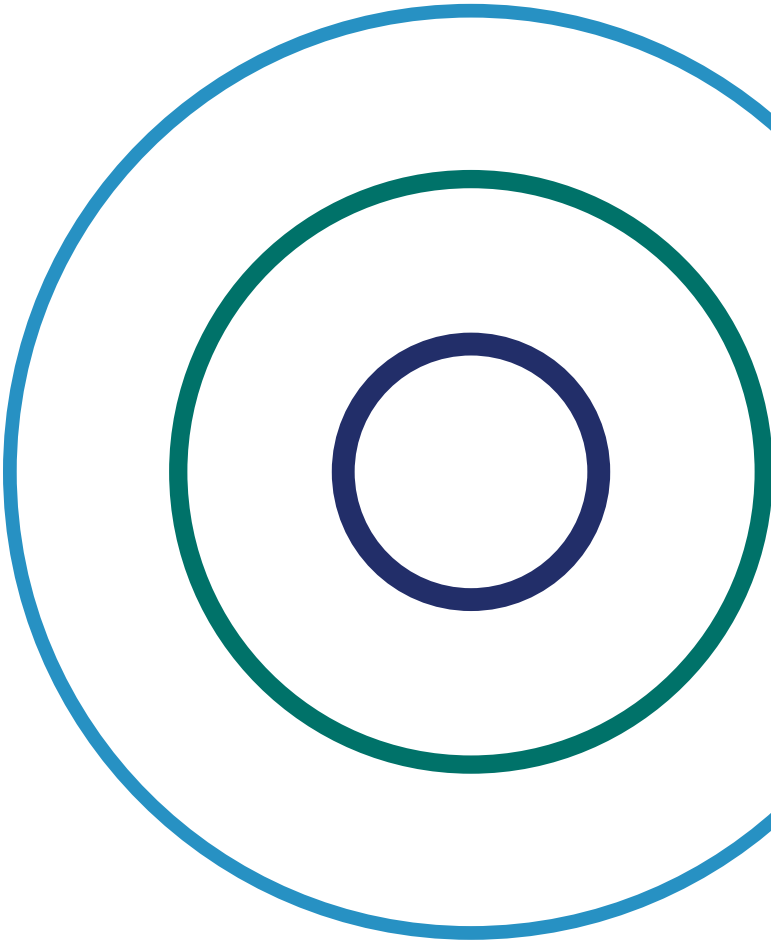


Tableau : Résumé des impacts environnementaux indirects de l'IA

Type d'impact indirect	Définition	Ordre	Conséquence
Effet d'efficacité	Réduction des impacts permise par l'amélioration de l'efficacité d'un système, procédé ou usage des ressources.	2e	Positive
Effet de substitution	Remplacement d'une technologie/activité par une autre.	2e	Positive ou négative
Effets d'induction	Création de nouveaux usages rendus possibles par la technologie.	2e	Positive ou négative
Effet rebond direct (paradoxe de Jevons)	Baisse du coût d'usage liée à un gain d'efficacité, qui entraîne une hausse de la demande et de la consommation globale.	2e	Négative
Effet d'échelle	Augmentation des impacts liée à la diffusion massive d'une technologie, même si chaque unité est plus efficace.	2e	Négative
Effet rebond indirect	Réallocation des gains d'efficacité vers d'autres formes de consommation.	3e	Positive ou négative
Effets comportementaux et temporels (*manifestation de l'effet rebond indirect*)	Réallocation du temps ou des habitudes libérées par l'IA vers d'autres activités.	3e	Négative (généralement)
Effets géographiques (*manifestation de l'effet rebond indirect*)	Transformation de l'organisation physique des activités.	3e	Positive ou négative
Effets systémiques	Transformations profondes des systèmes économiques et sociaux (chaînes de valeur, modes de production et de consommation).	3e	Positive ou négative

Une autre dimension, appelée *durabilité de l'IA*, vise à réduire l'empreinte environnementale des technologies elles-mêmes, notamment en optimisant les modèles, les infrastructures et les usages. Toutefois, la mise en œuvre de bonnes pratiques pour réduire l'empreinte environnementale d'un système d'IA ne garantit pas nécessairement que son utilisation soit durable ou alignée avec les objectifs climatiques et environnementaux, en raison de la multiplication des usages et des effets mentionnés plus haut. Par conséquent, l'évaluation de la durabilité d'une solution d'IA ne devrait pas uniquement porter sur son efficacité environnementale relative, mais également sur la pertinence du besoin auquel elle répond, les bénéfices réellement apportés et les impacts globaux susceptibles de se produire lors de son déploiement à grande échelle.

En résumé, une approche responsable de l'IA nécessite de considérer simultanément deux dimensions (l'IA pour la durabilité et la durabilité de l'IA), afin de s'assurer que les impacts environnementaux directs négatifs associés au développement et à l'utilisation de l'IA soient minimisés, justifiés et compensés par des bénéfices environnementaux plus importants.

2.3. Impacts sociaux et éthiques de l'IA

L'IA, en particulier l'IA générative, soulève d'importants enjeux éthiques et sociaux. Ces enjeux sont résumés dans le tableau ci-dessous et détaillés dans la suite de cette section.

Tableau : Résumé des impacts sociaux et éthiques de l'IA

Type d'impact social ou éthique	Conséquence
Biais et concentration du pouvoir	Négative
Fiabilité et responsabilité	
Vie privée et protection des données	
Transformation du marché du travail	
Enjeux éthiques liés aux impacts environnementaux	

2.3.1. Biais et concentration du pouvoir

Les systèmes d'IA peuvent comporter des biais, susceptibles d'accentuer les inégalités existantes. D'une part, les biais proviennent des données d'entraînement qui peuvent être incomplètes et biaisées (Gouvernement du Québec, 2024). Dans ces jeux de données, les populations marginalisées (c.-à-d. groupes de personnes subissant une discrimination liée à l'origine, le sexe, le statut socio-économique, etc.) sont souvent sous-représentées, entraînant des performances inégales des systèmes d'IA selon la population et pouvant contribuer à perpétuer les biais systémiques. D'autre part, les biais peuvent provenir des algorithmes qui pourraient privilégier certaines variables plutôt que d'autres et induire de mauvais résultats (p. ex. attribuer une importance différente selon le genre de la personne, ce qui pourrait induire des taux d'erreurs variables entre les genres) (Obvia, 2025).

Dans le domaine de l'IA générative, les coûts importants associés au développement des systèmes entraînent une concentration du pouvoir autour de quelques acteurs, qui proposent des algorithmes souvent peu transparents.

2.3.2. Fiabilité et responsabilité

Un deuxième enjeu concerne la fiabilité et la sécurité des systèmes d'IA. En plus du manque de transparence, les systèmes d'IA sont conçus pour générer des résultats, même si ceux-ci peuvent être erronés ou absurdes, ce qui augmente le risque d'hallucinations et d'erreurs (Gouvernement du Québec, 2024). Dans certains secteurs à haut niveau de risque (p. ex. secteur de la santé, des finances), des décisions prises à partir des résultats de systèmes d'IA peuvent avoir des conséquences significatives en cas d'erreurs d'algorithme, ce qui souligne la nécessité de maintenir l'implication des humains dans les processus décisionnels, et rappelle l'importance d'encadrer la responsabilité des systèmes d'IA en cas de décisions basées sur des résultats de systèmes d'IA erronés (Obvia, 2025).

2.3.3. Vie privée et protection des données

Un troisième enjeu porte sur le respect de la vie privée et la protection des données. Les modèles peuvent parfois exploiter, reproduire ou révéler des données personnelles, sensibles ou confidentielles, ou des données protégées par le droit d'auteur, sans le consentement des personnes concernées. Par exemple, des données confidentielles fournies à un système d'IA lors d'une inférence peuvent être réutilisées pour son réentraînement et réapparaître dans des inférences ultérieures générées par d'autres personnes utilisatrices (Gouvernement du Québec, 2024). Les données recueillies pourraient également être revendues par le fournisseur de la solution d'IA à des fins de promotion d'autres produits ou services. Ainsi, il est nécessaire que les fournisseurs de solutions d'IA définissent

des politiques de consentement explicites et transparentes, et que les personnes utilisatrices soient pleinement informées des usages prévus de leurs données personnelles.

2.3.4. Transformation du marché du travail

L'IA soulève également des enjeux liés aux conditions de travail. Le fonctionnement de l'IA peut reposer sur du personnel précaire (p. ex. pour l'annotation des données, la modération de contenu), sous des contrats temporaires, avec des salaires faibles et de mauvaises conditions de travail (Solaiman et al., 2023). Lors de son utilisation, l'IA générative peut entraîner des pertes d'emplois si elle remplace du personnel (p. ex. centres d'appels) (Solaiman et al., 2023). Les études démontrent que l'IA transforme le marché du travail de manière inégale : à long terme, elle accroîtra la productivité pour certains, mais fragilisera les travailleuses et travailleurs moins qualifiés (Qian et al., 2024).

2.3.5. Enjeux sociaux liés aux impacts environnementaux

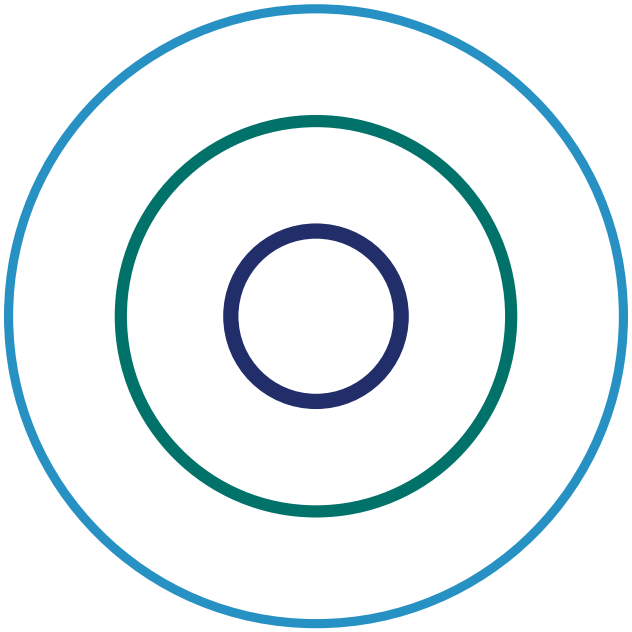
Pour terminer, certains enjeux sociaux sont liés au cycle de vie des équipements numériques (terminaux utilisateurs, réseaux, serveurs). Ces enjeux touchent particulièrement les pays du Sud (Solaiman et al., 2023).

Tout d'abord, l'extraction des métaux nécessaires à leur fabrication a souvent lieu dans des pays où la réglementation du travail dans les mines est insuffisante, entraînant des risques élevés de conditions de travail dangereuses pour la santé du personnel, de travail des enfants et d'exploitation des personnes (Carr-Wilson et al., 2024).

Ensuite, la production des puces électroniques intégrées aux équipements numériques requiert l'emploi d'environ 400 substances chimiques. Le processus génère ainsi d'importants volumes d'eaux usées chargées de polluants chimiques dangereux (métaux lourds, acides, etc.) susceptibles d'entraîner des conséquences significatives sur la santé humaine (Sandhu et al., 2025).

De plus, la fabrication des équipements numériques est fréquemment associée à des salaires bas et à des conditions d'emploi précaires (Subramanian & Yung, 2018).

Enfin, la gestion des équipements numériques en fin de vie affecte les communautés impliquées dans le recyclage informel des déchets électroniques, exposées à des conditions de travail déplorables (Traverso et al., 2024).



3.

Bonnes pratiques d'approvisionnement en solutions d'IA

En s'appuyant sur les risques environnementaux liés à l'IA identifiés et détaillés précédemment, cette section présente les bonnes pratiques d'approvisionnement à adopter. De manière générale, l'objectif est d'évaluer les impacts environnementaux de la solution IA et de démontrer la prise en considération de ces éléments dans tout le cycle de l'approvisionnement afin de respecter le principe de durabilité de l'*Énoncé de principes* (Gouvernement du Québec, 2025).

Les bonnes pratiques sont présentées en suivant les étapes de déploiement, dès la définition du besoin à l'amélioration continue :

3.1. Définir et remettre en question le besoin pour l'IA

3.2. Établir un cadre de gouvernance responsable

3.3. Anticiper les impacts potentiels de l'IA

3.4. Préparer le cahier des charges

3.5. Former les équipes au cadre d'usage

3.6. Mesurer et divulguer la performance de la solution

L'ensemble des bonnes pratiques est présenté dans un tableau résumé en annexe[9], À chaque pratique est attribuée un indice de facilité de mise en œuvre et un indice de potentiel de réduction des impacts environnementaux, afin d'aider les personnes utilisatrices de ce guide à prioriser leurs actions. Il s'agit de recommandations et d'appréciations issues de la revue de la littérature. Il est à la discrétion des organisations de décider de la façon la plus appropriée de les intégrer à leurs processus d'approvisionnement, en fonction de leur réalité et des moyens à leur disposition.

3.1. Définir et remettre en question le besoin pour l'IA

En approvisionnement responsable, une bonne définition du besoin (la raison de l'acquisition) est un facteur important de succès. Elle offre l'occasion de considérer les différentes alternatives pour répondre aux besoins identifiés, de cibler l'utilisation de l'IA sur ces besoins et d'éviter des utilisations superflues; elle facilite l'élaboration des critères d'approvisionnement par la suite, et constitue une référence tout au long du projet pour s'assurer que les décisions prises sont alignées avec les objectifs initiaux. Les bonnes pratiques présentées ci-après peuvent aussi s'appliquer lorsque l'achat d'IA est indirect (p.ex. inclus dans un logiciel plus général ou dans des services où l'IA est utilisée).

3.1.1. Énoncer et consigner clairement la problématique ou le besoin auquel répondre

Dans le contexte de l'adoption de nouvelles technologies, il est particulièrement important de prendre le temps d'en évaluer la pertinence. La première étape d'une acquisition d'IA est donc de bien définir le besoin pour la technologie. De plus, considérant la nature sensible et des dérives potentielles d'un projet lié à l'IA, une bonne documentation ne peut que favoriser l'acceptabilité sociale de la démarche. Cela permet de respecter le principe #5 de l'*Énoncé de principes pour une utilisation responsable de l'intelligence artificielle par les organismes publics* concernant l'efficience, l'efficacité et la pertinence : "Les systèmes d'IA déployés doivent être adaptés aux besoins réels

9. Le tableau a été élaboré à partir de guides, de lignes de bonnes pratiques, et de cadres réglementaires existants, principalement publiés par l'AFNOR (2024), l'Association GreenIT (2026), l'ARCEP (2024), ARCOM (2024), CESE (2024), INR (2021), Gouvernement du Québec (2025), UNEP (2024). L'ensemble des sources ayant contribué à son élaboration est toutefois indiqué dans le tableau annexe.

des organismes publics, des citoyens et des entreprises." (Gouvernement du Québec, 2025)

3.1.2. Constituer une équipe pluridisciplinaire pour soutenir la démarche

Entourez-vous de personnes possédant les connaissances et l'expérience nécessaires pour déterminer si l'IA constitue une solution souhaitable, viable et adaptée. Le développement, l'évaluation et la mise en œuvre de projets d'IA seront plus efficaces si vous disposez d'équipes diversifiées, capables de comprendre les besoins, les enjeux et les disciplines que les technologies d'IA englobent. Cela peut inclure des expertises :

- sectorielles (p. ex. santé, transport), liées à l'usage anticipé de la solution ;
- en développement durable ;
- en ingénierie des systèmes et des données ;
- en développement de modèles (par exemple, l'apprentissage profond) ;
- en éthique et protection des données ;
- en visualisation de l'information.

Si votre équipe manque d'expertise, vous pouvez faire appel à des réseaux professionnels au sein de votre organisation ou de l'administration pour recueillir des informations importantes sur le cas d'utilisation que vous envisagez.

3.1.3. Identifier des solutions (sans et avec IA) permettant de répondre à la problématique identifiée

Cela permet de mieux cerner le besoin réel et l'éventail de solutions possibles. Cet exercice peut permettre d'identifier des solutions existantes ou des processus à améliorer afin de répondre à une partie du problème sans ajouter de technologie inutile.

3.1.4. Évaluer les effets (environnementaux, sociaux, économiques) positifs et négatifs associés à ces solutions

Afin de prendre des décisions éclairées et alignées avec les objectifs de l'organisation, il faut penser aux effets potentiels pour s'assurer que la solution apporte un bénéfice clair et prévoir des mesures d'atténuation des impacts négatifs dans le choix de la solution, ses caractéristiques et les conditions d'utilisation lors de son déploiement (voir section 3.3 pour plus de détails).

3.1.5. Consulter le marché pour connaître les options potentielles et leurs caractéristiques

Une consultation du marché, dans cette phase de planification, peut aider à déterminer si, et dans quelle mesure, l'IA peut faire partie de la solution. Informez les fournisseurs de la problématique ou du défi et laissez-les proposer une solution qui réponde à vos besoins. Questionnez-les aussi sur leurs pratiques d'écoconception et de gestion durables dans tout le cycle de vie de leurs solutions afin d'évaluer leur maturité face aux exigences de développement durable (voir les sections 3.3 et 3.4 pour les sujets sur lesquels les questionner). Les enseignements tirés de cette consultation (par appel d'intérêt ou par d'autres moyens de consultation des fournisseurs) vous aideront à mieux cerner votre problématique, les options potentielles pour y répondre et leurs caractéristiques.

Outil : Fiche pratique pour l'Achat Responsable de solutions d'Intelligence Artificielle (IA)

La fiche pour l'Achat Responsable de solutions IA de la Mission Interministérielle Numérique Écoresponsable du gouvernement français sert de guide pratique pour intégrer la sobriété environnementale dans un projet d'IA. Elle permet d'abord de vérifier si l'IA est réellement nécessaire, puis d'orienter le choix du modèle et de l'architecture vers des options plus sobres. Elle fournit aussi des questions et clauses prêtes à l'emploi pour structurer un appel d'offres ou un contrat, en exigeant transparence, efficacité énergétique et maîtrise des réentraînements, et ce en fonction de l'interlocuteur. C'est un outil de référence pour consulter le marché et aligner besoins, impacts et exigences envers les fournisseurs. (Ministère de l'Économie, de l'Innovation et de l'Énergie [France], 2025)

3.1.6. Déterminer si le système (ou les fonctionnalités) d'IA constitue une solution pertinente

Après ces recherches préliminaires, il faut se demander si la solution d'IA est appropriée à notre problème, et non pas quelle utilisation on peut trouver à une solution d'IA que l'on souhaite acquérir ou que l'on se fait proposer. L'IA ne doit être déployée que si elle répond à des besoins réels. L'organisation doit démontrer en quoi l'IA apporte une valeur ajoutée par rapport à d'autres approches numériques moins impactantes et justifier l'usage de techniques avancées (comme l'apprentissage profond) au lieu de méthodes plus légères, le cas échéant. L'IA n'est pas une solution universelle; contrairement à un être humain, l'IA ne comprend pas la situation au sens humain du terme et ne peut produire un résultat qu'à partir des données introduites dans le modèle. Elle ne doit donc être déployée que lorsqu'elle constitue la réponse la plus adaptée au problème identifié.

Pour déterminer si l'IA peut vous aider à répondre aux besoins identifiés, demandez-vous si :

- des données contenant les informations dont vous avez besoin existent, même si celles-ci sont masquées ou difficilement accessibles ;
- vous disposez d'assez de données de bonne qualité pour entraîner le modèle ;
- l'utilisation des données est éthique et sûre ;
- la tâche visée est suffisamment vaste et répétitive ;
- l'utilisation de la solution fournirait des informations qu'une équipe pourrait utiliser pour obtenir des résultats concrets (Government Digital Service, 2023).

3.1.7. Déterminer le type de solution d'IA approprié aux besoins identifiés

Il peut s'agir d'une solution d'IA prête à l'emploi, d'une technologie existante avec des éléments d'IA ajoutés, de la cocréation de la solution d'IA avec les fournisseurs ou de l'externalisation du développement d'IA. Vous pourriez choisir de développer tout ou une partie de votre technologie, mais, comme il est plus environnementalement coûteux d'entraîner soi-même son modèle, il est recommandé d'opter plutôt pour une option ajustée à partir d'un modèle existant. Considérer la génération augmentée par récupération (ou *Retrieval Augmented Generation* [RAG]), qui réduit l'impact lié à l'entraînement, car si la base de données évolue, l'organisation peut simplement la mettre à jour pour maintenir la fiabilité des réponses du modèle, sans le réentraîner, contrairement aux grands modèles qui doivent régulièrement être réentraînés.

Choisir la technologie d'IA adaptée à votre défi

(traduit de : Government Digital Service, 2023)

Il n'existe pas une seule « technologie d'IA ». À l'heure actuelle, les technologies d'IA largement disponibles reposent principalement sur l'apprentissage automatique supervisé, non supervisé ou par renforcement. La technique d'apprentissage automatique qui vous offrira les meilleurs résultats dépend du problème que vous cherchez à résoudre. Le tableau suivant présente les différentes techniques d'apprentissage automatique pour vous orienter dans vos choix.

Tableau : Techniques d'apprentissage automatique

(traduit de : Government Digital Service, 2023)

Technique	Description	Exemples
Classification	Apprend les caractéristiques d'une catégorie donnée, permettant au modèle de classer des données inconnues dans des catégories existantes.	• décider si un lot de marchandises doit faire l'objet d'une inspection frontalière. • déterminer si un courriel est un pourriel ou non.
Régression	Prédit une valeur pour un point de données inconnu.	• prédire la valeur marchande d'une maison à partir d'informations comme sa taille, son emplacement ou son âge. • prévoir les concentrations de polluants atmosphériques dans les villes.
Clustering (regroupement)	Identifie des groupes de points de données similaires dans un ensemble de données.	• regrouper des clients pour identifier des sous-groupes ayant des habitudes de dépenses particulières. • regrouper des données de compteurs intelligents pour identifier des groupes d'appareils électriques.
Réduction de dimensionnalité	Réduit les données aux variables les plus pertinentes afin de rendre les modèles plus précis ou de permettre la visualisation des données.	• utilisé par les scientifiques des données lorsqu'ils évaluent ou développent d'autres types d'algorithmes d'apprentissage automatique.
Classement	Entraîne un modèle à classer de nouvelles données en se basant sur des listes déjà observées.	• retourner des pages par ordre de pertinence lorsqu'un utilisateur effectue une recherche sur un site web.

Applications courantes de l'apprentissage automatique

Certaines catégories de problèmes sont fréquemment traitées à l'aide de l'apprentissage automatique. Pour plusieurs d'entre elles, il est possible d'acheter ou d'adapter des produits commerciaux existants. Le tableau suivant présente différentes applications pour vous aguiller dans vos réflexions.

Tableau : Applications de l'apprentissage automatique
(traduit de : Government Digital Service, 2023)

Application	Description	Exemples
Traitement du langage naturel[10]	Traite et analyse le langage naturel, reconnaissant les mots, leur sens, le contexte et la structure narrative.	• convertir la parole en texte pour générer automatiquement des sous-titres. • générer automatiquement une réponse à un courriel d'un client.
Vision par ordinateur	Capacité d'une machine ou d'un programme à imiter la vision humaine.	• identification des panneaux routiers pour les véhicules autonomes. • reconnaissance faciale pour les contrôles automatisés de passeports.
Détection d'anomalies	Identifie des points de données anormaux dans un ensemble de données.	• repérer des activités frauduleuses dans le compte bancaire d'un utilisateur.
Analyse de séries temporelles	Comprendre comment les données évoluent dans le temps afin de réaliser des prévisions et du suivi.	• effectuer des analyses budgétaires. • prévoir des indicateurs économiques.
Systèmes de recommandation	Prédit la manière dont un utilisateur évaluera un élément donné afin de proposer de nouvelles recommandations.	• suggérer des pages pertinentes sur un site web en fonction des articles déjà consultés par l'utilisateur.

3.2. Établir un cadre de gouvernance responsable

Former un cadre pour une IA digne de confiance au sein de l'administration publique est essentiel au succès du projet, au respect des parties prenantes et à l'atténuation des risques. Ci-après sont présentées de grandes lignes à suivre pour établir un tel cadre dans le contexte de l'acquisition ou de l'utilisation d'une solution d'IA. Il s'agit de recommandations à haut niveau, qui peuvent être plus ou moins accessibles selon le contexte organisationnel, mais qu'il convient de tenter de mettre en œuvre le mieux possible. Au minimum, il faut appliquer le principe de proportionnalité, selon lequel les moyens déployés pour mitiger les risques identifiés doivent être à la hauteur de ces derniers. Selon le principe #9 de responsabilité de l'*Énoncé de principes*, les organismes publics demeurent entièrement responsables des décisions, des actions et des impacts générés par les systèmes d'IA, même lorsque ceux-ci fonctionnent de manière autonome. Ils doivent donc mettre en place une gouvernance adéquate ainsi que des mécanismes de reddition de comptes clairs et traçables, de manière à promouvoir la confiance du public dans l'utilisation des systèmes d'IA (Gouvernement du Québec, 2026).

10. Pour davantage de cas d'application potentiels de l'IA générative, consulter les Mesures applicables lors de l'utilisation de l'intelligence artificielle générative et leur Guide d'application. Pour documenter et analyser les cas d'usage de l'IAG, les ministères et organismes assujettis à la Loi sur la gouvernance et la gestion des ressources informationnelles des organismes publics et des entreprises du gouvernement (LGGRI) peuvent également s'appuyer sur le gabarit du *Portefeuille des occasions porteuses en IA* du Gouvernement du Québec (2026), réservé à l'usage de l'Administration publique et non accessible au public.

Pour plus d'informations sur les risques liés à la confiance et les garde-fous à considérer, consulter les travaux de l'OCDE (Gouverner avec l'intelligence artificielle; OCDE, 2025) ainsi que l'Énoncé de principes pour une utilisation responsable de l'intelligence artificielle par les organismes publics (Gouvernement du Québec, 2025). Les Mesures applicables lors de l'utilisation de l'intelligence artificielle générative (Gouvernement du Québec, 2025) et leur Guide d'application (Gouvernement du Québec, 2026) établissent pour leur part les balises de bonne gouvernance organisationnelle des systèmes d'IAG.

3.2.1. Établir un cadre de gouvernance multiacteurs

Ce cadre doit impliquer des parties prenantes internes et externes, pour soutenir la prise de décisions liée à l'acquisition et au déploiement de la solution. Cela peut inclure l'équipe pluridisciplinaire interne (voir § 3.1.2), des personnes expertes, des représentants des groupes bénéficiaires, des autorités locales, des communautés affectées et toute autre partie prenante externe pertinente.

3.2.2. Déterminer un processus pour consulter et informer le public de la mise en place du système d'IA lorsque la solution peut les concerner

L'institut AI Now propose une divulgation en quatre sections : 1) Champ d'application, 2) Avis public, 3) Autoévaluation et 4) Plan d'accès (voir encadré ci-dessous). Il faut collaborer activement avec les principales parties prenantes au moyen de mécanismes ouverts et transparents, par exemple des assemblées citoyennes, la mobilisation d'agents de la fonction publique ou l'implication des utilisateurs et utilisatrices dans la prise de décision.

OUTIL/RÉFÉRENCE :
Algorithmic Impact Assessments Report: A Practical Framework for Public Agency Accountability - AI Now Institute (Reisman, Schultz, Crawford & Whittaker, 2018).

- → **CHAMP D'APPLICATION :** définir le système de décision automatisé qu'on cherche à mettre en place de façon pratique et appropriée au contexte.
- → **AVIS PUBLIC :** informer les collectivités des systèmes susceptibles d'avoir une incidence sur leur vie quotidienne.
- → **AUTOÉVALUATION :** renforcer la capacité des organismes publics à évaluer l'équité, la justice, le respect des garanties procédurales et les effets discriminatoires. « Dans leurs autoévaluations, les agences doivent identifier les impacts potentiels sur le public et impliquer de manière proactive les communautés concernées. **[…] Une agence devrait impliquer ses communautés dès le début, avant même le processus officiel de notification et de consultation. »**
- → **PLAN D'ACCÈS PERTINENT :** Permettre aux personnes chercheuses et auditrices d'examiner les systèmes une fois mis en œuvre.

3.2.3. Mettre en place des mécanismes de contrôle du système d'IA

Les Mesures applicables lors de l'utilisation de l'intelligence artificielle générative et leur Guide d'application (Gouvernement du Québec, 2025) proposent des mécanismes de contrôle de ces systèmes, notamment l'établissement et l'application d'un processus de rétroaction structuré pour encourager les utilisateurs et utilisatrices à signaler leurs besoins, les obstacles encourus (p. ex. : hallucinations, biais) et leurs préoccupations.

3.2.4. Établir un registre des responsabilités précisant qui est responsable des différents aspects de la solution d'IA

Il est pertinent de vérifier si les modèles atteignent leurs objectifs, s'il existe un cadre de test et de surveillance clair, si les algorithmes sont robustes, impartiaux, équitables et explicables, et si le projet correspond à la manière dont les personnes citoyennes et/ou utilisatrices s'attendent à ce que leurs données soient utilisées, notamment.

3.3. Anticiper les impacts potentiels de l'IA

L'utilisation de l'IA comporte plusieurs enjeux, comme documenté dans la section 2. Ainsi, il est essentiel d'estimer les impacts potentiels et de développer des clauses et des critères d'approvisionnement qui permettent de les maîtriser.

Selon les Lignes directrices pour les approvisionnements IA du gouvernement du Royaume-Uni (Government Digital Service, 2020), une évaluation de l'impact de l'IA devrait inclure :

- Vos besoins utilisateur et le bénéfice public de votre système d'IA ;
- Les impacts humains et socio-économiques de votre système d'IA ;
- Les conséquences sur votre paysage technique et procédural existant ;
- La qualité des données et toute inexactitude ou biais potentiel ;
- Toute conséquence involontaire potentielle ;
- Les considérations liées au coût sur l'ensemble du cycle de vie, y compris les exigences de support et de maintenance continues.

De plus, les bénéfices et les impacts environnementaux de la solution sont à évaluer de manière qualitative et quantitative. Une de ces évaluations, pour les solutions avec systèmes de décision automatisés, est l'évaluation d'incidence algorithmique (voir encadré ci-contre).

3.3.1. Évaluer qualitativement les impacts environnementaux de l'IA sur l'ensemble de son cycle de vie, dès la phase de conception

Plus précisément, fournir une liste qualitative des impacts environnementaux potentiels directs et indirects négatifs (de 2e et 3e ordre), en couvrant toutes les étapes du cycle de vie de la manière la plus exhaustive possible. Si applicable, démontrer comment la solution contribue à des fins environnementales (p. ex. sensibilisation au développement durable auprès de la communauté étudiante, optimisation des trajets d'autobus, réduction du gaspillage alimentaire). Il s'agit d'une analyse sommaire en amont, pour l'étude de marché, qui sera complétée ultérieurement par une analyse de cycle de vie complète, si disponible (voir § 3.4.3).

Sans connaître les spécifications du système d'IA que vous allez acquérir, il n'est pas possible d'effectuer une évaluation complète du premier coup. Cela signifie que l'évaluation initiale doit être reconduite lorsque davantage d'informations sur la solution et le fournisseur sélectionnés seront disponibles, et ce, avant la mise en marche du système.

Évaluations d'incidence algorithmique pour les systèmes de décision automatisés

Les évaluations d'incidence algorithmique (EIA) (*algorithmic impact assessments*) s'inspirent des cadres d'évaluation d'impact dans les domaines de la protection de l'environnement, de la protection des données, de la vie privée et des politiques relatives aux droits humains, afin de permettre d'évaluer la mise en place d'un système de décision automatisé avant que l'administration ne s'engage à l'utiliser. Cela permet à l'administration et au public d'identifier les préoccupations qui pourraient nécessiter une négociation ou une autre forme de traitement avant la signature d'un contrat. C'est également à ce stade que le public et les élus peuvent s'opposer au déploiement avant que des préjudices potentiels ne se produisent.

Avant de commencer l'EIA, vous devriez vous procurer des renseignements sur:

→ l'algorithme, y compris les paramètres et les techniques de traitement des données ainsi que les résultats;

→ les données utilisées par le système, de même que les renseignements sur leur type, leur source, la méthode de collecte et leur classification de sécurité;

- → les mesures prévues ou en place pour assurer la qualité;
- → les mesures prévues pour assurer la transparence et communiquer l'information sur l'initiative à la clientèle et au public.

Outil/référence

Outil d'évaluation de l'incidence algorithmique (Gouvernement du Canada, 2026) contenu dans la Directive sur la prise de décisions automatisée du Gouvernement du Canada (2025)

L'outil d'EIA est un outil d'évaluation des risques conçu à l'appui de la Directive sur la prise de décisions automatisée du Conseil du Trésor et partagé en logiciel libre. L'EIA est accessible sous forme de questionnaire en ligne sur le Portail du gouvernement ouvert. Une fois le questionnaire rempli, un niveau d'incidence est établi en fonction des résultats. L'EIA sert à cerner les risques et à évaluer leur incidence dans un large éventail de secteurs, notamment :

- → les droits des personnes ou des collectivités;
- → la santé ou le bien-être des personnes ou des collectivités;
- → les intérêts économiques des personnes, des entités ou des collectivités;
- → la durabilité des écosystèmes.

3.3.2. Analyser l'impact des solutions d'IA sur les équipements informatiques

Il convient de s'assurer que le modèle d'IA peut fonctionner avec le matériel existant et, quand pertinent, via une connexion à bas débit ou hors connexion, afin d'éviter le renouvellement prématuré du matériel. Cela inclut les terminaux, les systèmes d'exploitation et les navigateurs web.

3.3.3. Déterminer des points et des seuils de décision associés à l'évaluation des impacts

Les risques associés et leurs stratégies d'atténuation respectives doivent être fournis et approuvés dans le cadre de l'évaluation des impacts et devraient inclure des points de décision clés « feu vert/feu rouge » lorsque cela est applicable. Prévoir des points de décision pour passer en revue l'évaluation des impacts ainsi que chaque fois qu'un changement substantiel à la conception du système d'IA est effectué.

3.3.4. Vérifier que l'implantation de la technologie ne compromet pas les engagements climatiques et de développement durable de l'organisation

L'IA déployée doit être cohérente avec les stratégies climatiques et de développement durable publiques, y compris les objectifs d'atténuation et d'adaptation aux changements climatiques. Les stratégies locales, nationales et internationales ne doivent pas être remises en cause par la croissance incontrôlée de l'IA et, en particulier, de la demande énergétique. Avec les estimations d'émissions de GES réalisées (et la consommation d'énergie sous-jacente), vérifier si cette consommation et les émissions additionnelles ne sont pas incohérentes avec les cibles de décarbonation qui concernent votre organisation.

3.4. Préparer le cahier des charges

Ajouter des critères environnementaux et sociaux dans le cahier des charges permet de sécuriser les engagements des fournisseurs en matière de développement durable. L'objectif est que le processus d'approvisionnement soit responsable afin d'atténuer les principaux risques liés à l'utilisation de l'IA. Cette section, comme l'ensemble du document, se concentre sur les critères environnementaux, mais deux encadrés à la fin proposent des considérations sociales et éthiques additionnelles.

3.4.1. Considérer des méthodes d'achat public agiles et innovantes

Si vous achetez le développement de la solution d'IA, donnez la priorité à une approche itérative du développement de produits et reflétez cela en conséquence dans l'invitation à soumissionner. Veillez à adopter une approche qui n'impose pas de charges inutiles qui dissuadent les fournisseurs, notamment les entreprises en démarrage, les petites et moyennes entreprises (PME), les fournisseurs issus du secteur associatif, communautaire et social et ceux appartenant à des groupes sous-représentés, de participer à la concurrence. Indiquer clairement, dans votre documentation relative à la passation de marché, pourquoi vous considérez que l'IA est pertinente pour le problème, tout en restant ouvert à d'autres solutions.

Quelle que soit la stratégie choisie, il faut garder à l'esprit que les technologies d'IA évoluent rapidement et que de nouvelles technologies et de nouveaux produits font sans cesse leur apparition sur le marché. Dans les appels d'offres, **définissez des critères axés sur les résultats,** en mettant l'accent sur la description des défis et des opportunités auxquels vous êtes confronté. Cela permettra aux fournisseurs de déterminer quelles technologies sont les mieux adaptées à vos besoins (voir encadré ci-contre).

Éléments clés à prendre en compte lors de la rédaction de votre cahier des charges (Government Digital Service, 2020) :

1. Commencez par définir votre problématique ;
2. Mettez en avant votre stratégie et vos exigences en matière de données ;
3. Mettez l'accent sur la qualité des données, les biais et les limites ;
4. Posez des questions pour insister sur la nécessité pour vous de comprendre l'approche du fournisseur en matière d'IA (algorithmes, modèles, source des données, etc.) ;
5. Envisagez des stratégies pour éviter la dépendance vis-à-vis d'un fournisseur et les systèmes d'IA de type « boîte noire » ;
6. Soulignez l'importance de la propriété intellectuelle ;
7. Indiquez toute intégration avec des technologies ou des services associés ;
8. Tenez compte de vos besoins en matière de support et de maintenance à long terme ;
9. Incluez des considérations relatives à la responsabilité et aux risques.

3.4.2. Identifier et prévoir les stratégies qui permettent d'atténuer les principaux risques identifiés tout au long du cycle de vie

Les conclusions des évaluations d'impact initiales doivent éclairer la passation de marché et être réexaminées aux moments clés de la prise de décision.

Les bonnes pratiques ci-après, transformées en critères d'acquisition responsable obligatoires ou d'évaluation, visent à réduire l'impact environnemental des solutions d'IA acquises.

TRANSPARENCE ENVIRONNEMENTALE

3.4.3. Privilégier un fournisseur ayant réalisé une analyse du cycle de vie (ACV) multicritère de la solution d'IA

Cette ACV doit être basée sur des standards reconnus, ou avoir fait l'objet d'une révision par une tierce partie crédible. Sinon, demander au fournisseur d'estimer l'impact environnemental de sa solution d'IA à l'aide des outils d'estimation existants. Les résultats devraient être transparents sur le type de données utilisées et les étapes du cycle de vie incluses. L'encadré ci-après propose une hiérarchie des approches qu'il est possible de demander.

Recommandations pour la mesure quantitative des impacts du cycle de vie de l'IA

Idéalement, réaliser une ACV de la solution d'IA, en suivant les normes et standards internationalement reconnus (*ISO 14040/44* [ISO, 2022; ISO 2006] , *Product Environmental Footprint* [Damiani et al., 2022], *Référentiel général pour l'IA frugale* [AFNOR, 2024], etc.) Les indicateurs environnementaux évalués devraient inclure au minimum la contribution aux changements climatiques (émissions de GES), la consommation d'eau et la déplétion des ressources (minérales et fossiles).

Si la réalisation d'une ACV n'est pas possible, estimer les impacts environnementaux à l'aide d'outils d'estimation existants: CodeCarbon (Courty et al., 2026), Green Algorithms (Lannelongue et al., 2021), ou Ecologits (calculateur en ligne) (Rincé et al., 2025). Ces outils permettent d'estimer au minimum la consommation d'énergie et la contribution aux changements climatiques. Toutefois, leurs estimations sont moins précises que les résultats d'une ACV, ce qui explique pourquoi ils doivent être considérés en seconde option.

Dans les deux cas, la quantification des impacts devrait idéalement être réalisée avec des données mesurées (p. ex. consommation d'énergie, durée d'entraînement, etc.) ou calculées à partir de celles-ci; sinon, à partir d'estimations et de calculs théoriques.

Si l'ensemble du cycle de vie ne peut être couvert, inclure au minimum dans l'étude les étapes d'entraînement et d'inférence.

Dans tous les cas, indiquer de manière transparente le type de données utilisées (mesurées, estimées) et les étapes du cycle de vie qui sont incluses dans l'évaluation, à la fois pour le cycle de vie technologique (entraînement, inférence, etc.) et physique (production du matériel, transport, utilisation, fin de vie).

Par exemple, la compagnie française Mistral AI a réalisé une ACV d'un modèle d'IA (Mistral AI, 2025). Cette étude évalue l'impact de l'entraînement du modèle Mistral Large 2, et d'une inférence moyenne associée à la génération d'une page de texte, selon trois indicateurs environnementaux (changement climatique, utilisation de l'eau, et l'épuisement des ressources (minéraux et métaux).

3.4.4. Exiger les données nécessaires à la mesure de la performance environnementale

En fonction des mesures que l'on veut suivre (voir section 3.6.2 sur les indicateurs de durabilité), le fournisseur doit obligatoirement s'engager à fournir des informations concernant l'utilisation et la performance du modèle d'IA tout au long de son cycle de vie. Les développeurs de modèles ne publient que rarement leurs données environnementales, il faut donc l'exiger. Ceci inclut la consommation d'énergie, d'eau et les autres impacts sur le cycle de vie.

HÉBERGEMENT

3.4.5. Privilégier les centres de données situés dans une région au mix électrique faiblement carboné ou avec un facteur d'énergie renouvelable (*Renewable Energy Factor* [REF]) élevé

Le REF doit être le plus proche possible de 1, pour l'entraînement et l'inférence. Dans le cas d'utilisation d'un service infonuagique à grande échelle (p. ex. Google cloud, AWS, Microsoft Azure), il est possible de choisir la région où est effectué le traitement. La plateforme Electricity Maps (2026) permet de visualiser l'intensité carbone du mix électrique à travers le monde.

3.4.6. Privilégier les centres de données avec une bonne performance environnementale

Le PUE, WUE, CUE, REF et ERF sont des indicateurs pouvant être utilisés pour mesurer la performance des centres de données. Le tableau suivant présente ces indicateurs, en précisant leurs équations, leurs unités, la norme encadrant leurs calculs (et qui devrait être suivie pour calculer ces indicateurs, ou un équivalent), ainsi que la valeur cible à atteindre pour assurer une bonne performance environnementale. Il s'agit cependant d'indicateurs de site et non d'une mesure complète de l'impact d'une solution IA. En pratique, l'amélioration d'un indicateur peut entraîner une dégradation d'un autre, c'est pourquoi il est nécessaire de suivre ces indicateurs en parallèle.

Tableau : Indicateurs de performance pour l'hébergement

Nom de l'indicateur	Équation	Norme de calcul	Valeur visée
PUE (Power Usage Effectiveness) [1]	$PUE = \frac{\text{Énergie totale du centre de données [kWh]}}{\text{Énergie des équipements IT [kWh]}}$	ISO/IEC 30134-2	Inférieur à 1,5 (Arcep et Arcom, 2024). La moyenne mondiale est 1,54 (Taylor, P., 2026).
CUE (Carbon Usage Effectiveness) [2]	$CUE = \frac{\text{Émissions totales de } CO_2 \text{ du centre de données}}{\text{Énergie des équipements IT [kWh]}}$	ISO/IEC 30134-8	Le plus faible possible.

Tableau : Indicateurs de performance pour l'hébergement (suite)

Nom de l'indicateur	Équation	Norme de calcul	Valeur visée
WUE (Water Usage Effectiveness) [3]	$WUE = \frac{\text{Consommation d'eau totale du centre de données [L]}}{\text{Énergie des équipements IT [kWh]}}$	ISO/IEC 30134-9	Inférieur à 0,4 L/kWh (Arcep et Arcom, 2024) en incluant uniquement la consommation d'eau sur site (directe). La moyenne aux États-Unis est de 0,36 L/kWh (Shehabi et al., 2024).
REF (Renewable Energy Factor)	$REF = \frac{\text{Énergie renouvelable du centre de données [kWh]}}{\text{Énergie totale du centre de données [kWh]}}$	ISO/IEC 30134-3	le plus proche de 1 possible. La moyenne mondiale est à 27% (AIE, 2025).
ERF (Energy Reuse Factor) [4]	$ERF = \frac{\text{Énergie résiduelle réutilisée [kWh]}}{\text{Énergie totale du centre de données [kWh]}}$	ISO/IEC 30134-6	le plus proche de 1 possible

Note : Pour les centres de données de plus de 2 ans, ces indicateurs devraient représenter la valeur "réelle" et non la valeur théorique "à la conception", lorsque c'est possible. Ils doivent être évalués sur une période minimale d'un an, mais idéalement sur plusieurs années.

[1] : Le PUE mesure l'efficacité énergétique d'un centre de données, et particulièrement du système de refroidissement, mais ne reflète pas l'efficacité des équipements informatiques eux-mêmes : un site peut présenter un faible PUE tout en sous-utilisant les serveurs, et en consommant beaucoup d'énergie. En lien avec l'énergie, le PUE ne renseigne pas sur les émissions de GES associées à la production de l'électricité — et doit donc être complémentaire au CUE et REF — ni sur la récupération d'énergie (p. ex. récupération de la chaleur fatale des serveurs pour chauffer des bâtiments). Enfin, le PUE ne tient pas compte de la consommation d'eau. Or, le remplacement d'un système de refroidissement par climatisation par un système par évaporation d'eau permet souvent de réduire le PUE tout en augmentant la consommation d'eau, soulignant l'importance de suivre le WUE en parallèle du PUE.

[2]: La norme ISO/IEC 30134-8 définit deux catégories de CUE. La catégorie 1 ne comptabilise que les émissions de CO2 liées à la consommation d'électricité du centre de données (et s'exprime en kg CO2/kWh). La catégorie 2 comptabilise les émissions de tous les GES liés à la consommation de toutes les formes d'énergie (et s'exprime en kg eq. CO2/kWh).

[3] : Dans le calcul du WUE, la consommation d'eau correspond à l'eau prélevée qui n'est pas restituée à sa source d'origine (incluant notamment l'eau perdue lors du fonctionnement des systèmes de refroidissement et de l'humidification des salles informatiques). Cet indicateur présente plusieurs limites. D'une part, il ne distingue pas l'origine de l'eau prélevée (eau potable, eau recyclée, eau de surface, etc.), ni la région de prélèvement, alors que l'impact environnemental du prélèvement d'un même volume d'eau varie fortement selon le niveau de stress hydrique local. Par ailleurs, le WUE de type 1 de la norme ISO/IEC 30134-9, le plus évalué, se limite à la consommation directe d'eau sur site, et n'intègre pas la consommation d'eau indirecte liée à la production de l'électricité consommée par le centre de données, ou à la fabrication des équipements informatiques. Enfin, le WUE ne prend pas en compte les volumes prélevés puis rejetés dans le milieu naturel ni la qualité de ces rejets, ce qui limite son aptitude à évaluer les impacts écologiques des rejets.

[4] : Tous les lieux ne sont pas techniquement ou économiquement adaptés à la récupération et à l'utilisation de la chaleur résiduelle générée par les serveurs, surtout dans les centres de données plus petits ou décentralisés. Comme la ERF dépend fortement des conditions locales et des paramètres de mesure, les comparaisons isolées sont d'une utilité limitée.

3.4.7. Privilégier un hébergeur engagé dans une démarche de réduction de l'impact environnemental

Plusieurs critères peuvent s'appliquer :

- Avoir fait son bilan de GES selon le GHG Protocol ou une méthodologie reconnue et disposer d'un plan de réduction des émissions de GES ;
- Avoir fait une évaluation des risques liés aux changements climatiques du service et disposer d'un plan de réduction et d'adaptation aux risques liés au climat ;
- Disposer d'un programme d'achat responsable pour les équipements informatiques (serveurs, baies de stockage, etc.) ;
- Disposer d'un programme de gestion durable des équipements en fin de vie qui respecte la hiérarchie des 3RV (réduction, réemploi, recyclage et valorisation) ;
- Posséder une/des certifications de gestion durable : *ISO 14001, ISO 50001*, certification TCO Certified cloud, ou similaire.

Outil/référence

La fiche d'approvisionnement responsable en services infonuagiques (ECPAR, 2024) de l'ECPAR propose des clauses pour réduire l'empreinte environnementale de l'hébergement.

COLLECTE ET GESTION DES DONNÉES

3.4.8. Limiter les données aux seules informations nécessaires sur le cycle de vie de l'IA

Il s'agit d'identifier, dès la phase de conception, les données réellement indispensables pour atteindre la performance attendue (Commission Nationale de l'Informatique et des Libertés [CNIL], 2024; CNIL, 2025). Cela évite de collecter ou de traiter des données non pertinentes ou personnelles et permet de réduire l'empreinte environnementale (et autres risques) (Gouvernement du Québec, 2025). Le fournisseur devrait expliquer quels types de données sont utilisés (texte, image, audio, etc.), justifier leur utilisation, et comment les données pertinentes sont sélectionnées, tout au long du cycle de vie du système (entraînement, test, inférence).

3.4.9. Privilégier la qualité des données plutôt que la quantité

Des données fiables, bien triées et annotées permettent d'entraîner un modèle avec un volume de données moins important (Sajith & Kathala, 2024; AFNOR, 2024), tout en conservant un bon niveau de performance. Ainsi, moins de ressources sont consommées et moins d'impacts sont générés. Les jeux de données en libre accès sont une source à privilégier en premier, car ils ont souvent déjà été triés, nettoyés. Le fournisseur devrait justifier la provenance des données sur le cycle de vie de l'IA et expliquer pourquoi ces sources ont été retenues.

3.4.10. Optimiser le nettoyage, l'annotation et la préparation des données

Un bon traitement préalable (suppression des doublons, correction des incohérences, annotation adaptée à la finalité, etc.) améliore les performances du modèle et évite un entraînement excessivement long. Cette rigueur devrait s'étendre à l'ensemble de l'infrastructure de données du fournisseur, pas seulement aux jeux d'entraînement : la tenue d'un inventaire exhaustif et à jour des données utilisées, incluant leurs métadonnées, permet de repérer et

d'éliminer les versions obsolètes et les données dupliquées à travers le cycle de vie complet du système, réduisant d'autant le volume de données stockées (Gouvernement du Québec, 2025). Le fournisseur devrait documenter sa procédure de préparation des données, y compris la tenue de cet inventaire, et expliquer comment il garantit la qualité des données préparées.

3.4.11. Adapter les stratégies de stockage en fonction de l'usage des données

Les données fréquemment consultées et nécessitant un accès rapide (parfois appelées *données chaudes*) doivent être stockées sur des supports de stockage performants (p. ex. disques SSD, mémoires volatiles), mais énergivores, tandis que les données rarement consultées (parfois appelées *données froides*) peuvent être stockées sur d'autres supports moins performants (p. ex. disques HDD) (Seagate, s. d.), mais qui consomment moins d'énergie. Une bonne pratique consiste à réduire le volume des données froides (p. ex. compression, réduction de la précision). Le fournisseur devrait être en mesure d'expliquer sa stratégie de stockage et ses choix technologiques.

3.4.12. Limiter la conservation des données collectées sur le cycle de vie du système d'IA, particulièrement pendant l'utilisation du modèle

Mettre en place un processus régulier pour vérifier la pertinence de ces données, fixer une durée de conservation des données, et supprimer celles qui ne sont plus nécessaires (CNIL, 2024; CNIL, 2025). Le fournisseur devrait prévoir un tel processus. Au Québec, cette obligation s'appuie notamment sur le calendrier de conservation prévu par la *Loi sur les archives* (RLRQ, c. A-21.1), auquel les organismes publics doivent se conformer pour établir leurs processus de conservation et de suppression des données (Gouvernement du Québec, 2025).

CARACTÉRISTIQUES DU MODÈLE

3.4.13. Privilégier les modèles existants, affinés, plutôt que bâtis de toutes pièces

L'affinage est le "processus par lequel un modèle d'intelligence artificielle préentraîné sur un vaste ensemble de données générales est adapté pour une tâche, un domaine ou un contexte plus particulier en ajustant plus amplement ses paramètres" (OQLF, 2025). Cela permet de mutualiser les coûts d'entraînement et d'autres impacts avec d'autres organisations utilisatrices.

3.4.14. Choisir le modèle le plus léger possible (c.-à-d. avec le moins de paramètres)

Comparés aux grands modèles de langage (ou *Large Language Models* [LLM]), les petits modèles de langage (ou *Small Language Models* [SLM]) ont moins de paramètres - allant d'un million jusqu'à 10 milliards de paramètres (Johnson, J., 2025), même si cette définition varie selon les sources - et consomment ainsi moins de ressources, mais peuvent être tout aussi performants que des LLM grâce à l'affinage. Ce choix est aussi recommandé par le Gouvernement du Québec (2024).

Cependant, les SLM sont dédiés à l'exécution de tâches limitées, et ne sont pas destinés à des tâches diverses, complexes et variées (Jeanquartier et al., 2026). Dans la mesure du possible, choisir un modèle spécialisé plutôt que généraliste dans le type de tâches à effectuer. Privilégier les modèles transparents sur le nombre de paramètres.

Outil/référence

La plateforme française Compar:IA (Ministère de la Culture [France], s. d.) donne des informations sur la taille des grands modèles afin d'éclairer la prise de décision.

Outil/référence

Pour guider dans le choix d'un petit modèle, SLM-Benchmark (Pham et al., 2025) est un benchmark en libre accès qui évalue les SLM selon leur précision, leur efficacité et leur durabilité, sur la réalisation de différentes tâches. D'après ce benchmark, Mistral-7B est le modèle qui offre le meilleur score vis-à-vis de ces 3 indicateurs, ce qui en fait un bon candidat pour de nombreuses applications.

3.4.15. Privilégier les modèles économes en énergie

Recommandé par le Gouvernement du Québec (2024), il est possible d'utiliser les benchmarks et classements comparatifs existants sur l'efficacité énergétique des modèles (p. ex. AI Energy Score [Luccioni et al., 2025b], Compar:IA, ML.ENERGY Leaderboard [Chung et al., 2025]). Ces benchmarks aident à identifier la consommation d'énergie des modèles pour un même service, et ainsi à choisir des modèles à plus faible consommation d'énergie.

3.4.16. Optimiser les performances du modèle grâce à des techniques de compression

Les techniques les plus courantes et efficaces sont : l'élagage (ou *pruning*), qui consiste à supprimer les connexions ou les neurones redondants ou peu utiles, la distillation des connaissances (ou *distillation*), visant à entraîner un modèle plus petit qui imite un modèle plus grand et complexe, et la quantification (ou *quantization*), consistant à réduire la précision numérique des paramètres (Berenstein & Charpentier, 2025). Elles visent à diminuer le besoin en ressources de calcul et en mémoire, mais d'autres techniques existent.

3.4.17. Réduire la consommation énergétique de l'inférence en privilégiant l'inférence par lots

L'*inférence par lots* consiste à regrouper plusieurs requêtes et à les traiter ensemble, plutôt qu'une par une et en temps réel (ce qui correspond à l'*inférence continue*). L'inférence par lots permet de réduire la fréquence de sollicitation du matériel informatique (GPUs et/ou CPUs), de l'utiliser plus efficacement , et ainsi de diminuer la consommation d'énergie (Hugging Face, s. d.). Cette approche est particulièrement pertinente lorsque les requêtes ne nécessitent pas de réponse instantanée.

3.4.18. Mettre en cache les fragments de requêtes répétées pour éviter des calculs inutiles

Le *prompt caching* consiste à mémoriser les parties de requêtes qui reviennent souvent (p. ex. une longue séquence de contexte répétée dans chaque requête) (Association GreenIT, 2026). Cela permet d'éviter de mobiliser le matériel informatique plusieurs fois pour traiter des éléments de prompts identiques, réduisant ainsi la charge de calcul, particulièrement pour les modèles d'IA générative. Le fournisseur devrait expliquer les stratégies de mise en cache appliquées, et préciser approximativement la part des séquences de requêtes mises en cache (Association GreenIT, 2025).

(RÉ)ENTRAÎNEMENT

3.4.19. Pendant le (ré) entraînement, trouver l'équilibre entre les ressources mobilisées pour l'entraînement, pour améliorer l'efficacité du modèle d'IA, et les gains de performance obtenus

Autrement dit, il convient de déterminer le niveau de performance optimal du modèle, afin d'éviter de gaspiller des ressources dans l'entraînement d'un modèle dont la performance a déjà atteint un plateau. Le fournisseur peut expliquer les techniques ou approches mises en place pour évaluer cet équilibre et stopper l'apprentissage du modèle lorsqu'il atteint son niveau optimal de performance.

Pour l'entraînement, le niveau optimal de quantité de données d'entraînement à collecter peut être déterminé par rapport à la taille du modèle. La "loi d'échelle de Chinchilla" indique qu'un modèle atteint sa performance optimale lorsque le nombre de *tokens* d'entraînement est environ 20 fois le nombre de paramètres du modèle (Hoffmann et al., 2022). Cette loi suggère que certains grands modèles existants (p. ex. GPT-3) auraient été entraînés de manière sous-optimale, en raison d'un nombre de paramètres trop élevé par rapport au volume des données d'entraînement. Elle montre également que l'augmentation du nombre de paramètres d'un modèle n'améliore pas nécessairement ses performances. Un modèle comportant moins de paramètres peut être plus performant, si la taille du modèle et le volume des données sont équilibrés.

3.4.20. Privilégier les phases d'entraînement sur des périodes de sous-utilisation des infrastructures ou sur des périodes pendant lesquelles l'intensité carbone du réseau électrique est faible

L'intensité carbone de l'électricité variant selon le jour et le moment de la journée, les charges de calcul importantes (p. ex. l'entraînement) devraient, lorsque c'est possible, être effectuées ou programmées pour être exécutées durant les périodes où l'électricité est moins carbonée. Une option consiste à démarrer l'entraînement du modèle lorsque l'intensité carbone de l'électricité dans la région des centres de données est inférieure à un certain seuil, puis à suspendre l'entraînement lorsque l'intensité carbone dépasse un certain seuil. La plateforme Electricity Maps donne des estimations de l'intensité carbone en temps réel de la majorité des pays.

3.4.21. Minimiser le nombre de réentraînements du modèle

La nécessité de phases de réentraînement doit être justifiée, et celles non essentielles, évitées. Le processus d'évaluation et les critères justifiant la nécessité d'un réentraînement doivent être expliqués (p. ex. des retours d'utilisateurs ou d'utilisatrices qui indiquent que le contenu généré n'est pas conforme aux besoins), par l'organisation ou par le fournisseur (en cas d'acquisition d'une solution d'IA). Il est également nécessaire d'indiquer le nombre de réentraînements du modèle prévus sur toute sa durée de vie (une plage peut être indiquée), lequel devrait être le plus faible possible, ainsi que de fournir des rapports sur les améliorations (qualitatives et quantitatives) apportées par le réentraînement et les impacts environnementaux associés.

INFÉRENCE (REQUÊTES)

3.4.22. Privilégier les modèles qui sont exécutables en local sur les terminaux utilisateurs.

Cela permet principalement d'éviter le transfert des données sur les réseaux, et donc de réduire leur impact associé, à condition que ces modèles soient petits (notamment de type SLM) et n'induisent pas une obsolescence et un remplacement des terminaux utilisateurs. Dans ce contexte, le fournisseur devrait garantir la compatibilité de la solution d'IA avec les terminaux déjà déployés, y compris lors de l'introduction

d'améliorations au modèle. Il devrait également offrir la possibilité de conserver une version antérieure afin de ne pas imposer la mise à niveau ou le remplacement des appareils.

3.4.23. Donner la possibilité d'activer ou de désactiver l'appel à l'IA

Cela permet de limiter le nombre de requêtes non essentielles. Il peut également être demandé à un utilisateur de confirmer son besoin d'avoir recours à l'IA avant d'effectuer une requête (Association GreenIT, 2026). Cela peut impliquer de modifier les options d'affichage par défaut, c'est-à-dire que le système d'IA ne soit pas activé sans intervention de l'utilisateur (*opt-in*).

3.4.24. Programmer le modèle pour limiter la taille des réponses générées

Cela permet de réduire les impacts puisque la taille de la réponse (*output tokens*) a une grande influence sur la quantité de ressources utilisées (Luccioni et al., 2024).

AUTRES PRATIQUES D'ÉCOCONCEPTION

3.4.25. Adopter ou exiger d'autres pratiques d'écoconception

Consulter le Référentiel général de l'écoconception des services numériques (Arcep & Arcom, 2024) pour trouver des mesures additionnelles appropriées à votre contexte et à votre cas d'usage et articuler des exigences potentielles pour les fournisseurs. Ces pratiques vont au-delà des modèles d'IA et s'appliquent aux solutions numériques.

Considérations sociales additionnelles

Pour le volet social, il est pertinent de demander aux fournisseurs de l'information sur leur chaîne d'approvisionnement et leurs pratiques numériques socio-responsables, par exemple :

- → Pouvez-vous fournir de l'information sur la chaîne de production de votre solution d'IA (conception, entraînement, test, déploiement, amélioration continue), notamment les entreprises impliquées et leur localisation ?
- → Votre entreprise est-elle engagée dans une démarche de numérique responsable sur le volet social (respect des droits de la personne et du travail, éthique des algorithmes et non-discrimination, accessibilité, protection des données personnelles, etc.) ?
- → Quelles mesures sont mises en place pour minimiser l'impact social négatif de votre solution (sur les droits de la personne et du travail, l'éthique des algorithmes et la non-discrimination, l'accessibilité, la protection des données personnelles, etc.) dans les différentes étapes de son cycle de vie (conception, entraînement, test, déploiement, amélioration continue) ?

Ces éléments s'inscrivent dans les principes de respect des personnes et de la règle de droit ainsi que d'inclusion et d'équité de l'*Énoncé de principes* du Gouvernement du Québec (2025).

Considérations éthiques additionnelles

Pour le volet éthique, il est important de questionner, de prendre en considération et d'inclure des clauses contractuelles sur les pratiques des fournisseurs pour assurer la transparence des modèles, la gouvernance des données, le contrôle humain, le niveau d'exactitude et de robustesse, etc. À cet effet, les clauses contractuelles de l'Union européenne présentées ci-après sont très utiles. Le modèle de clauses pour la documentation technique est idéal pour inclure des informations relatives à l'empreinte environnementale du système d'IA.

Outil/référence

Clauses contractuelles types mises à jour de l'UE relatives aux modèles d'IA (Commission européenne, 2025).

Le 13 juin 2024, l'Union européenne a adopté la loi sur l'intelligence artificielle (loi européenne sur l'IA), première législation mondiale réglementant l'IA de manière globale. Afin de favoriser des achats responsables en matière d'IA, la Communauté de pratique a publié, le 29 septembre 2023, des clauses contractuelles types de l'UE, mises à jour en mars 2025.

Cette ressource contient des clauses contractuelles pour les sujets suivants :

- → Exigences essentielles applicables au Système d'IA
  - o Système de gestion des risques
  - o Données et gouvernance des données
  - o Documentation technique* et notice d'utilisation
  - o Enregistrement (des événements)
  - o Transparence
  - o Contrôle humain
  - o Exactitude, robustesse et cybersécurité
- → Obligations du Fournisseur en ce qui concerne le Système d'IA
  - o Conformité
  - o Système de gestion de la qualité
  - o Analyse d'impact sur les droits fondamentaux
  - o Mesures correctives
  - o Obligation de fournir une explication des prises de décision individuelles
- → Droits relatifs à l'utilisation des jeux de données
- → Registre de Systèmes d'IA et audit
- → Coûts

Additionnellement, l'accès libre aux modèles peut aussi être un critère éthique à intégrer dans le choix ou le développement de la solution d'IA.

Outil/référence

Le site Compar:IA (Ministère de la Culture [France], s. d.) indique si les grands modèles existants sont à code en libre accès ou propriétaire.

3.5. Former les équipes au cadre d'usage

Les organisations doivent considérer la maîtrise de l'IA responsable comme un engagement à long terme, en l'intégrant dans leurs stratégies de développement des ressources humaines (formation) et en prévoyant des mécanismes de contrôle appropriés afin de permettre une surveillance rigoureuse des systèmes d'IA tout au long de leur cycle de vie.

3.5.1. Inclure dans la formation des personnes utilisatrices des notions liées au développement durable et aux risques environnementaux (sociaux et éthiques) de l'IA

Le personnel opérationnel ou de service doit disposer de connaissances suffisantes sur le système d'IA pour comprendre comment l'utiliser de manière responsable et donner suite à ses résultats. Cela permet de respecter le principe #10 de l'*Énoncé de principes*, à savoir celui de la compétence : "Le personnel des organismes publics doit être sensibilisé à l'utilisation, aux bonnes pratiques et aux enjeux pouvant survenir tout au long du cycle de vie des systèmes d'IA [...]" (Gouvernement du Québec, 2025). Il est important de s'assurer que le transfert de connaissances et la formation fassent partie des exigences du contrat et que les explications soient accessibles aux non-spécialistes qui pourraient avoir besoin de comprendre le système d'IA. Les Mesures applicables lors de l'utilisation de l'intelligence artificielle générative et leur Guide d'application fournissent des exemples de plans de formation pour l'IAG (Gouvernement du Québec, 2026). Le Gouvernement du Québec met également à la disposition des organismes publics, sur les plateformes BRIO et Agora, six formations soutenant une utilisation responsable de l'IA, dont Introduction à l'intelligence artificielle pour l'adminis-tration publique et Intelligence artificielle générative : opportunités, risques et défis (parties 1 et 2).

3.5.2. Préciser les profils et les cas d'utilisation autorisés afin d'éviter les utilisations non essentielles

Dans les Mesures applicables lors de l'utilisation de l'intelligence artificielle générative et leur Guide d'application (Gouvernement du Québec, 2026), il est demandé aux organismes publics d'adopter de telles règles internes et on fournit des exemples de cas d'utilisation qui peuvent être autorisés ou interdits (Gouvernement du Québec, 2026). Les services des technologies de l'information (TI) internes devraient définir des profils d'utilisation de l'IA et en permettre l'accès uniquement au personnel qui en a réellement besoin, selon les besoins identifiés dans la phase de planification.

3.5.3. Inviter les personnes utilisatrices à minimiser le nombre de requêtes adressées à l'IA

Regrouper les tâches afin d'effectuer un minimum de requêtes (*prompt engineering*).

> HEC Montréal a créé un Catalogue de requêtes pour aider sa communauté à écrire des requêtes de qualité. Cela permet d'obtenir une réponse satisfaisante avec un moins grand nombre d'itérations.

3.5.4. Sensibiliser le personnel aux fonctionnalités intensives en ressources

Les tâches de génération (génération de texte, de résumé, correction de texte, etc.) sont plus énergivores et plus impactantes que les tâches discriminatives (classification, détection d'erreurs, etc.) (Luccioni et al., 2024) et celles qui impliquent des images sont également plus impactantes que les tâches qui manipulent uniquement du texte. Alors, les tâches les plus énergivores devraient être réservées aux situations où elles apportent une réelle valeur ajoutée. De plus, les tâches complexes nécessitant des étapes de

raisonnement précises devraient être effectuées avec l'approche par *chaîne de raisonnement* (ou *chain-of-thought*), afin d'identifier et de corriger les erreurs potentielles en effectuant un nombre limité de requêtes, mais cette approche, plus consommatrice d'énergie, ne devrait pas être utilisée pour les tâches plus simples.

3.5.5. Informer la personne utilisatrice de l'impact environnemental de sa requête

L'informer minimalement de la consommation d'énergie et de la contribution aux changements climatiques. Certains outils s'intègrent dans la base de code informatique, mesurent directement la consommation d'énergie lors de l'exécution de la requête, puis estiment l'empreinte carbone (p. ex. CodeCarbon, Ecologits (répertoire Python)). Lors de l'utilisation de certains grands modèles d'IA (Claude, Gemini, ChatGPT, etc.), l'impact environnemental moyen d'une requête peut être estimé à l'aide d'Ecologits (calculateur en ligne). Informer les utilisateurs et utilisatrices de l'impact environnemental de leur requête permet de faire des choix d'utilisation de l'IA plus éclairés.

3.6. Mesurer et divulguer la performance de la solution

Pour les solutions alimentées par l'IA, les plans de mise en œuvre, les méthodes d'évaluation durables en continu, ainsi que les mécanismes permettant de réinjecter les informations dans le modèle de données sont cruciaux pour garantir une utilisation responsable. De plus, le fonctionnement et les conséquences des systèmes d'IA peuvent ne pas être apparents dans le processus d'approvisionnement et ne deviennent souvent évidents qu'au moment du déploiement, nécessitant une communication et un partage d'informations prolongés entre l'acheteur et le fournisseur qui doivent faire partie des conditions contractuelles.

3.6.1. S'assurer que le modèle reste performant, précis et sobre une fois déployé, et que les pratiques d'exploitation évoluent dans le bon sens

Il est nécessaire de tester le modèle de manière continue afin d'en garantir la précision, car un modèle imprécis peut conduire à des décisions erronées ayant des conséquences négatives. Exiger certains indicateurs de performance ou définir avec le fournisseur comment l'efficacité du modèle sera contrôlée une fois celui-ci déployé. Quelques indicateurs sont proposés dans le tableau suivant.

Tableau : Indicateurs de performance de la solution IA

Élément mesuré	Étape du cycle de vie	Indicateur	Outil/source
Effort de calcul par usage	Inférence	Nombre de tokens traités par requête	Logs du fournisseur / API
Volume de données stockées	Stockage	Gigaoctets stockés (chaud/froid)	Logs de l'hébergeur
Fréquence de réentraînement	Entraînement	Nombre de réentraînements par mois ou par an	Fournisseur
Taux de réutilisation des prompts (caching)	Inférence	% de tokens évités grâce au cache	Logs du fournisseur

3.6.2. Mettre en place un suivi et une reddition de comptes d'indicateurs de durabilité

Comme recommandé par le Gouvernement du Québec (Mesures applicables lors de l'utilisation de l'IA générative, 2025), il faut mettre en place ce suivi des indicateurs de durabilité et l'intégrer dans la gouvernance des projets. Mettre en place des outils pour mesurer les performances tout au long du cycle de vie des systèmes d'IA et rendre des comptes publics à cet effet. Des indicateurs qui peuvent être inclus dans le cahier des charges sont présentés dans le tableau ci-après.

Tableau : Indicateurs de performance environnementale de la solution IA (adaptés de : Ministère de l'Économie, de l'Innovation et de l'Énergie [France], 2025)

Élément mesuré	Étape du cycle de vie	Indicateur	Outil/source
Coût énergétique du réentraînement	Réentraînement	kWh par réentraînement	CodeCarbon Green Algorithms
Efficacité énergétique par token	Inférence	kWh/1000 tokens	Estimation interne EcoLogits
Émissions de GES	Entraînement et inférence	CO_2éq. émis par entraînement ou requête	CodeCarbon Green Algorithms
Impact par requête (API externes)	Inférence	Énergie, émissions, eau, matériel mobilisé	EcoLogits
Intensité carbone du mix électrique	Hébergement	gCO_2/kWh	Hébergeur
Efficacité du centre de données	Hébergement	PUE, CUE, WUE, ERF	Hébergeur
Taux d'énergie renouvelable	Hébergement	% d'énergie renouvelable utilisée (REF ou taux du mix national)	Hébergeur

3.6.3. Éviter de créer une dépendance à des modèles trop lourds et remettre en question le besoin de maintenir les outils d'IA au fil du temps

Le besoin initial peut évoluer, et un modèle adéquat au départ peut devenir surdimensionné, notamment si l'usage réel est plus limité que prévu ou si un modèle plus léger répond désormais au même besoin. Il convient donc de prévoir, dès le cahier des charges, une révision périodique de la pertinence de la solution et de ses bénéfices au regard de ses impacts, afin de migrer au besoin vers un modèle plus sobre ou de mettre fin à son utilisation.

3.6.4. Mettre en place une stratégie de réutilisation et de seconde vie du modèle d'IA en interne et/ou dans la communauté

Les modèles d'IA peuvent enrichir une bibliothèque interne partagée entre les organisations collaboratrices, ou peuvent être partagés à code source ouvert à la communauté, en vue d'éviter des réentraînements coûteux et énergivores. Ces modalités doivent être prévues dans l'appel au marché. La réutilisation et la mutualisation des systèmes d'IAG et des requêtes sont des recommandations du Gouvernement du Québec (Mesures applicables lors de l'utilisation de l'IA générative, 2025) qui doivent être prévues dans le cahier des charges dès le début.

Bibliographie

Académie de la transformation numérique. (2025). *NETendances 2025 : Intelligence artificielle générative.* Université Laval. https://transformation-numerique.ulaval.ca/wp-content/uploads/2026/03/netendances25-intelligence-artificielle-generative.pdf

Aczel, M., Chamanara, S., Martin, M., Farsi, A., Marwala, T., & Madani, K. (2026). *Environmental Cost of AI's Energy Use : Carbon, Water and Land Footprints.* United Nations University Institute for Water, Environment and Health (UNU-INWEH). https://doi.org/10.53328/INR26RMA002

ADEME. (2024). *Evaluation de l'empreinte environnementale de la fourniture d'accès à internet en France* (Expertises) [Etude/Recherche]. https://librairie.ademe.fr/industrie-et-production-durable/6789-evaluation-de-l-empreinte-environnementale-de-la-fourniture-d-acces-a-internet-en-france.html

ADEME. (2025). *Évaluation de l'impact environnemental du numérique en France.* (Expertises) https://librairie.ademe.fr/changement-climatique/7880-evaluation-de-l-impact-environnemental-du-numerique-en-france.html

Association Française de Normalisation (AFNOR). (2024). *Référentiel général pour l'IA frugale* (Version AFNOR SPEC 2314). https://www.afnor.org/actualites/referentiel-pour-mesurer-et-reduire-impact-environnemental-de-ia/

Agence internationale de l'énergie (AIE). (2025). *Energy and AI*. IEA. https://www.iea.org/reports/energy-and-ai

Arcep, & Arcom. (2024). *Référentiel général d'écoconception de services numériques.* https://ecoresponsable.numerique.gouv.fr/publications/referentiel-general-ecoconception/

Association GreenIT. (2025). RIA 0013 – Maximiser le « prompt caching ». Référentiel de bonnes pratiques – RIA. https://ria.greenit.fr/fr/fiches/RIA_0013-maximiser-le-prompt-caching

Association GreenIT. (2025). *Impacts environnementaux du numérique dans le monde* (Troisième édition). https://greenit.eco/nos-etudes-et-essais/impacts-environnementaux-du-numerique-dans-le-monde-2025/

Association GreenIT. (2026). *Référentiel de bonnes pratiques d'utilisation de l'intelligence artificielle générative.* https://ria.greenit.fr/fr/fiches

Baldé, C., Kuehr, R., Yamamoto, T., McDonald, R., D'Angelo, E., Althaf, S., Bel, G., Deubzer, Fernandez-Cubillo, E., Forti, V., Gray, V., Herat, S., Honda, S., Iattoni, G., Khetriwal, D., Luda di Cortemiglia, L., Lobuntsova, Y., Nnorom, I., Pralat, N., & Wagner, M. (2024). *Rapport mondial sur les déchets d'équipements électriques et électroniques—2024.* https://ewastemonitor.info/wp-content/uploads/2024/12/GEM_2024_FR_28_NOV-web.pdf

Berenstein, D., & Charpentier, B. (2025). *An Introduction to AI Model Optimization Techniques.* Hugging Face. https://huggingface.co/blog/PrunaAI/introduction-to-ai-model-optimization-techniques

Berthelot, A., Caron, E., Jay, M., & Lefèvre, L. (2024). Estimating the environmental impact of Generative-AI services using an LCA-based methodology. *Procedia CIRP, 122,* 707-712. https://doi.org/10.1016/j.procir.2024.01.098

Bolón-Canedo, V., Morán-Fernández, L., Cancela, B., & Alonso-Betanzos, A. (2024). A review of green artificial intelligence: Towards a more sustainable future. *Neurocomputing, 599,* 128096. https://doi.org/10.1016/j.neucom.2024.128096

Bureau de Recherches Géologiques et Minières. (2025). *Substances critiques et stratégiques*. Minéralinfo. https://www.mineralinfo.fr/fr/securite-des-approvisionnements-pour-leconomie/substances-critiques-strategiques

Carr-Wilson, S., Pattanayak, S. K., & Weinthal, E. (2024). Critical mineral mining in the energy transition : A systematic review of environmental, social, and governance risks and opportunities. E*nergy Research & Social Science, 116,* 103672. https://doi.org/10.1016/j.erss.2024.103672

Chung, J.-W., Wu, R., Liu, J., Kweon, O. J., Xia, Y., Wu, Z., & Chowdhury, M. (2025). *The ML.ENERGY Benchmark : Toward Automated Inference Energy Measurement and Optimization.* https://neurips.cc/virtual/2025/loc/san-diego/poster/121781

Commission de l'éthique en science et en technologie (CEST). (2024). *Regard éthique sur les effets environnementaux des technologies numériques au Québec : L'impératif de la sobriété numérique.* https://leconcepteurweb.com/wp-content/uploads/2024/10/cest_avis-sur-la-sobriete-numerique.pdf

Commission européenne. (2025). *Updated EU AI model contractual clauses.* Public Buyers Community – Procurement of AI. https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/updated-eu-ai-model-contractual-clauses

Commission Nationale de l'Informatique et des Libertés [CNIL]. (2024). *IA : Tenir compte de la protection des données dans la collecte et la gestion des données.* CNIL. https://www.cnil.fr/fr/tenir-compte-de-la-protection-des-donnees-dans-la-collecte-et-la-gestion-des-donnees

CNIL. (2025). *Développement des systèmes d'IA : Que faut-il vérifier ?* [Liste de contrôle]. https://www.cnil.fr/sites/default/files/2025-07/ia_liste_de_verification.pdf

Conseil économique social et environnemental (CESE). (2024). *Impacts de l'intelligence artificielle : risques et opportunités pour l'environnement.* https://www.lecese.fr/travaux-publies/impacts-de-lintelligence-artificielle-risques-et-opportunites-pour-lenvironnement

Conseil national du numérique, ADEME, Arcep, IGN, ECOLAB (Commissariat Général au Développement Durable, Direction générale des entreprises (ministère de l'Économie), DReDIS de l'Institut de recherche juridique de la Sorbonne, Inria, Sciences Po, & Thomas Le Goff (Telecom Paris). (2025, mai). *IA & Environnement* [Module pédagogique]. https://cafeia.org/ia-et-environnement/

Courty, B., Schmidt, V., Goyal-Kamal, inimaz, MarionCoutarel, Blanche, L., Feld, B., Lecourt, J., Saboni, A., LiamConnell, Berenstein, D., LLORET, P., SabAmine, supatomic, cianc, Léval, M., Cruveiller, A., ouminasara, Zhao, F., … rosekelly6400. (2026). *mlco²/codecarbon : V3.2.8* (Version v3.2.8) [Logiciel]. Zenodo. https://doi.org/10.5281/zenodo.20584960

Damiani, M., Ferrara, N., & Ardente, F. (2022). *Understanding Product Environmental Footprint and Organisation Environmental Footprint methods.* Publications Office of the European Union. https://doi.org/doi/10.2760/11564

Data Center Map. (2026). *Data Centers* [Jeu de données]. https://www.datacentermap.com/datacenters/

Desroches, C., Chauvin, M., Ladan, L., Vateau, C., Gosset, S., & Cordier, P. (2025). *Exploring the sustainable scaling of AI dilemma: A projective study of corporations' AI environmental impacts.* https://doi.org/10.48550/arXiv.2501.14334

Dugas Bourdon, P. (2024). *Québec a autorisé le géant Microsoft à détruire l'équivalent de neuf terrains de football en milieux humides.* Le Journal de Montréal. https://www.journaldemontreal.com/2024/11/01/microsoft-compte-detruire-lequivalent-de-neuf-terrains-de-football-de-milieux-humides-avec-laccord-du-ministere-de-lenvironnement

Gouvernement du Canada. (2025). *Comprendre les effets environnementaux de l'IA.* https://www.csps-efpc.gc.ca/tools/articles/ai-environmental-effects-fra.aspx

Electricity Maps. (2026). *Carbon Intensity Data* (Version 03 juillet 2026) [Jeu de données]. Electricity Maps. https://app.electricitymaps.com/map/72h/hourly

Environnement et Changement climatique Canada. (2026). *Rapport d'inventaire national 1990-2024 : Sources et puits de gaz à effet de serre au Canada* [Jeu de données]. https://www.canada.ca/fr/environnement-changement-climatique/services/indicateurs-environnementaux/emissions-gaz-effet-serre.html

Espace de concertation des pratiques d'approvisionnement responsable [ECPAR] (2024). Fiche d'approvisionnement responsable. Services infonuagiques. https://ecpar.org/cms/documents/89/O_Fiche_achat_responsable_Services_Infonuagiques_2024.pdf

Falk, S., Ekchajzer, D., Pirson, T., Lees-Perasso, E., Wattiez, A., Biber-Freudenberger, L., Luccioni, S., & Wynsberghe, A. van. (2025). *More than Carbon: Cradle-to-Grave environmental impacts of GenAI training on the Nvidia A100 GPU* (arXiv:2509.00093). arXiv. https://doi.org/10.48550/arXiv.2509.00093

Government Digital Service. (2020). *Guidelines for AI procurement.* GOV.UK. https://www.gov.uk/government/publications/guidelines-for-ai-procurement/guidelines-for-ai-procurement

Government Digital Service. (2023). *Assessing if artificial intelligence is the right solution.* GOV.UK. https://www.gov.uk/guidance/assessing-if-artificial-intelligence-is-the-right-solution

Gouvernement du Canada (2025). Directive sur la prise de décisions automatisée. Gouvernement du Canada. https://www.tbs-sct.canada.ca/pol/doc-fra.aspx?id=32592

Gouvernement du Canada (2026). Outil d'évaluation de l'incidence algorithmique. Gouvernement du Canada. https://www.canada.ca/fr/gouvernement/systeme/gouvernement-numerique/innovations-gouvernementales-numeriques/utilisation-responsable-ai/evaluation-incidence-algorithmique.html

Groupement CLIK, De Montenay, L., Petit, B., Diguet, C., & Fourboul, E. (2026). *Prospective d'évolution des consommations des data centers à court, moyen et long terme de 2024 à 2060* (Expertises) [Etude/Recherche]. ADEME. https://librairie.ademe.fr/energies/8910-10774-prospective-d-evolution-des-consommations-des-data-centers-a-court-moyen-et-long-terme-de-2024-a-2060.html#product-features

HEC Montréal (s.d.). Catalogue de requêtes. https://alfred.hec.ca/video/DAIP/intelligence_artificielle/catalogue-de-requetes/index.html

HEC Montréal. (2024, octobre). *Concepts clés en IA* [Parcours IA]. https://ia.hec.ca/parcours/quest-ce-que-ia/concepts-cles-en-ia/

Hoffmann, J., Borgeaud, S., Mensch, A., Buchatskaya, E., Cai, T., Rutherford, E., de Las Casas, D., Hendricks, L. A., Welbl, J., Clark, A., Hennigan, T., Noland, E., Millican, K., van den Driessche, G., Damoc, B., Guy, A., Osindero, S., Simonyan, K., Elsen, E., … Sifre, L. (2022). Training compute-optimal large language models. *Proceedings of the 36th International Conference on Neural Information Processing Systems, NIPS '22*. https://doi.org/10.48550/arXiv.2203.15556

Hugging Face. (s. d.). *Batch inference.* Hugging Face. https://huggingface.co/docs/diffusers/en/using-diffusers/batched_inference

Institut du Numérique Responsable (INR). (2021). Ministère de la transition écologique. *Guide pratique pour des achats numériques responsables.* https://ecoresponsable.numerique.gouv.fr/publications/guide-pratique-achats-numeriques-responsables/

International Telecommunication Union (ITU). (2026). *Recommendation ITU-T L.1801 (02/2026)—Guidelines for assessing the environmental impact of artificial intelligence systems.* https://www.itu.int/epublications/publication/itu-t-l-1801-2026-02-guidelines-for-assessing-the-environmental-impact-of-artificial-intelligence-systems

ISO. (2006). *ISO 14044:2006 Management environnemental—Analyse du cycle de vie—Exigences et lignes directrices.* https://www.iso.org/fr/standard/38498.html

Istrate, R., Tulus, V., Grass, R. N., Vanbever, L., Stark, W. J., & Guillén-Gosálbez, G. (2024). The environmental sustainability of digital content consumption. *Nature Communications, 15*(1), 3724. https://doi.org/10.1038/s41467-024-47621-w

Jeanquartier, F., Jean-Quartier, C., Rieder, P., Misirlić, V., Pasero, C., Hohensinner, R., Müller, H., & Holzinger, A. (2026). Assessing the carbon footprint of language models: Towards sustainability in AI. *Resources, Conservation and Recycling, 226,* 108670. https://doi.org/10.1016/j.resconrec.2025.108670

Johnson, J. (2025). *Small Language Models (SLM): A Comprehensive Overview.* Hugging Face. https://huggingface.co/blog/jjokah/small-language-model

Lange, S., Frick, V., Gossen, M., Pohl, J., Rohde, F., & Santarius, T. (2023). The induction effect: Why the rebound effect is only half the story of technology's failure to achieve sustainability. *Frontiers in Sustainability, Volume 4-2023.* https://doi.org/10.3389/frsus.2023.1178089

Lannelongue, L., Grealey, J., & Inouye, M. (2021). Green Algorithms: Quantifying the Carbon Footprint of Computation. *Advanced Science, 8*(12), 2100707. https://doi.org/10.1002/advs.202100707

Luccioni, S., Jernite, Y., & Strubell, E. (2024). Power Hungry Processing: Watts Driving the Cost of AI Deployment? *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency, FAccT '24*, 85-99. https://doi.org/10.1145/3630106.3658542

Luccioni, A. S., Strubell, E., & Crawford, K. (2025a). From Efficiency Gains to Rebound Effects : The Problem of Jevons' Paradox in AI's Polarized Environmental Debate. *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency, FAccT '25,* 76-88. https://doi.org/10.1145/3715275.3732007

Luccioni, S., Gamazaychikov, B., Strubell, E., Hooker, S., Jernite, Y., Mitchell, M., & Chamberlin, S. (2025b). *AI Energy Score Leaderboard.* Hugging Face. Hugging Face. https://huggingface.co/spaces/AIEnergyScore/Leaderboard

Ministère de l'Économie, de l'Innovation et de l'Énergie. (2025). *Fiche : IA frugale.* Programme Numérique écoresponsable. Gouvernement de France. https://ecoresponsable.numerique.gouv.fr/docs/2025/MINUMECO_fiche_IA_frugale_VF.pdf

Ministère de la Culture. (s. d.). *Compar:IA - Le comparateur d'IA conversationnelles* [Outil d'évaluation d'intelligence artificielle]. Gouvernement de France. https://comparia.beta.gouv.fr

Gouvernement du Québec (2025). Mesures applicables lors de l'utilisation de l'intelligence artificielle générative. https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/cybersecurite_numerique/Publications/Strategie_IA/Indication_app_IAG.pdf

Gouvernement du Québec (2025). Guide d'application des mesures applicables lors de l'utilisation de l'intelligence artificielle générative. https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/cybersecurite_numerique/Publications/Strategie_IA/GU_lignes_dir_IAG_2025.pdf

Gouvernement du Québec. (2025). *Arrêté numéro 2025-02 du ministre de la Cybersécurité et du Numérique en date du 3 décembre 2025 : Énoncé de principes pour une utilisation responsable de l'intelligence artificielle par les organismes publics.* Publications du Québec. https://www.publicationsduquebec.gouv.qc.ca/fileadmin/gazette/pdf_encrypte/lois_reglements/2025F/86931.pdf

Mistral AI. (2025). Notre contribution pour la création d'un standard environnemental mondial pour l'IA. https://mistral.ai/fr/news/our-contribution-to-a-global-environmental-standard-for-ai/

Obvia. (2025). *État de la situation sur les impacts sociétaux de l'intelligence artificielle et du numérique—2025.* https://doi.org/10.61737/HCLZ2424

Office québécois de la langue française (OQLF). (2026). *Intelligence artificielle.* Le grand dictionnaire terminologique. https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/8385376/intelligence-artificielle

OQLF. (2025). *Affinage.* Le grand dictionnaire terminologique. https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/26579046/affinage

Ollion, L., Drapeau, D., Boztepe, L., De Villers, S., Bouyer, G., Vergari, L., Fangeat, E., & Kergaravat, O. (2024). *Besoins en métaux dans le secteur numérique* (Expertises). https://librairie.ademe.fr/economie-circulaire-et-dechets/7713-etude-numerique-et-metaux.html

Organisation de coopération et de développement économiques [OCDE]. (2024). *Recommandation du Conseil sur l'intelligence artificielle* (OECD/LEGAL/0449). Consulté 2 juin 2026, à l'adresse https://legalinstruments.oecd.org/fr/instruments/OECD-LEGAL-0449?_ga=2.124697866.1898590258.1621541868-1852956558.1620326402

OCDE. (2025). Gouverner avec l'intelligence artificielle : État des lieux et perspectives pour les fonctions essentielles de l'État, Éditions OCDE, Paris, https://doi.org/10.1787/6816434b-fr

Organisation Internationale de Normalisation (ISO). (2022). *ISO 14040:2006 : Management environnemental—Analyse du cycle de vie—Principes et cadre.* https://www.iso.org/fr/standard/37456.html

Oueslati, Y. (2026). *Quelle est l'empreinte écologique réelle de la fabrication des infrastructures numériques, et comment elle s'est rendue plus intense à l'ère de l'IA générative ?* Digital HEC Montréal. https://digital.hec.ca/blog/quelle-est-lempreinte-ecologique-reelle-de-la-fabrication-des-infrastructures-numeriques-et-comment-elle-sest-rendue-plus-intense-a-lere-de-lia-generative/

Pham, N. T., Kieu, T., Nguyen, D. M., Xuan, S. H., Duong-Trung, N., & Le-Phuoc, D. (2025). SLM-Bench: A Comprehensive Benchmark of Small Language Models on Environmental Impacts. Dans C. Christodoulopoulos, T. Chakraborty, C. Rose, & V. Peng (Éds.), *Findings of the Association for Computational Linguistics : EMNLP 2025* (p. 21369-21392). Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.findings-emnlp.1165

Qian, Y., Siau, K. L., & Nah, F. F. (2024). Societal impacts of artificial intelligence : Ethical, legal, and governance issues. *Societal Impacts, 3,* 100040. https://doi.org/10.1016/j.socimp.2024.100040

Reisman, D., Schultz, J., Crawford, K. & Whittaker, M. (2018). *Algorithmic Impact Assessments Report: A Practical Framework for Public Agency Accountability.* AI Now Institute. https://ainowinstitute.org/publications/algorithmic-impact-assessments-report-2

Rincé, S., Banse, A., Defour, V., & Jean-Pierre, C. (2025). EcoLogits: Evaluating the Environmental Impacts of Generative AI. *Journal of Open Source Software, 10*(111), 7471. https://doi.org/10.21105/joss.07471

Roussilhe, G. (2024). *Taïwan, histoires mêlées et futurs orageux de la numérisation.* https://gauthierroussilhe.com/articles/taiwan-histoires-glorieuses-et-futurs-sombres-de-la-numerisation

Roussilhe, G., Pirson, T., Xhonneux, M., & Bol, D. (2024). From silicon shield to carbon lock-in? The environmental footprint of electronic components manufacturing in Taiwan (2015–2020). Journal of Industrial Ecology, 28, 1212–1226. https://doi.org/10.1111/jiec.13487

Sajith, A., & Kathala, K. (2024). *Is Training Data Quality or Quantity More Impactful to Small Language Model Performance?* https://doi.org/10.48550/arXiv.2411.15821

Sandhu, S., Zumeit, A., Tian, Z., Vinciguerra, V., & Dahiya, R. (2025). Semiconductor manufacturing wastewater challenges and the potential solutions via printed electronics. *iScience, 28*(10), 113576. https://doi.org/10.1016/j.isci.2025.113576

Seagate. (s. d.). *Cold Data Storage Explained* [Blog]. Seagate. https://www.seagate.com/ca/en/blog/what-is-cold-data-storage/

Schneider Electric. (2023). *Quantifying Data Center Scope 3 GHG Emissions to Prioritize Reduction Efforts—White Paper 99.* https://www.se.com/ww/en/download/document/SPD_WP99_EN/

Secrétariat du Conseil du trésor. (2021). *Stratégie d'intégration de l'intelligence artificielle dans l'administration publique 2021-2026.* Gouvernement du Québec. https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/cybersecurite_numerique/Publications/Strategie_IA/Strategie_integration_IA_2021_2026.pdf

Shehabi, A., Smith, S.J., Hubbard, A., Newkirk, A., Lei, N., Siddik, M.A.B., Holecek, B., Koomey, J., Masanet, E., Sartor, D. 2024. 2024 United States Data Center Energy Usage Report. Lawrence Berkeley National Laboratory (LBNL), Berkeley, California. LBNL-2001637. https://eta-publications.lbl.gov/sites/default/files/2024-12/lbnl-2024-united-states-data-center-energy-usage-report_1.pdf

Solaiman, I., Talat, Z., Agnew, W., Ahmad, L., Baker, D., Blodgett, S. L., Chen, C., Daum'e, H., Dodge, J., Evans, E., Friedrich, F., Gohar, U., Hooker, S., Jernite, Y., Luccioni, A. S., Lusoli, A., Mickel, J., Mitchell, M., Newman, J. C., … Subramonian, A. (2023). Evaluating the Social Impact of Generative AI Systems in Systems and Society. *ArXiv.* https://doi.org/10.48550/arXiv.2306.05949

Spindler, W., Atamian Hahn-Petersen, L., Hosseini, S., & Ketelslegers, J. (2024). *Why circular water solutions are key to sustainable data centres.* World Economic Forum. https://www.weforum.org/stories/2024/11/circular-water-solutions-sustainable-data-centres/

Subramanian, K., & Yung, W. K. C. (2018). Modeling Social Life Cycle Assessment framework for an electronic screen product – A case study of an integrated desktop computer. *Journal of Cleaner Production, 197,* 417-434. https://doi.org/10.1016/j.jclepro.2018.06.193

Sveinsdottir, T., Troullinou, P., Aidlinis, S., Delipalta, A., Finn, R., Loukinas, P., Muraszkiewicz, J., O'Connor, R., Petersen, K., Rovatsos, M., Santiago, N., Sisu, D., Taylor, M., & Wieltschnig, P. (2020, décembre). *The Role of Data in AI.* Zenodo. https://doi.org/10.5281/zenodo.4312907

Taylor, P. (2026). *Data center average annual power usage effectiveness (PUE) worldwide from 2007 to 2025* (Statista) [Jeu de données]. https://www.statista.com/statistics/1229367/data-center-average-annual-pue-worldwide/?srsltid=AfmBOoolk_7OmxZoD7iczSqII_t2K7-CJTBj4ZSa6jmYJJ_4mG8PVvci

The Shift Project. (2025). *Intelligence artificielle, données, calculs : Quelles infrastructures dans un monde décarboné ?* https://theshiftproject.org/publications/intelligence-artificielle-centres-de-donnees-rapport-final/

Traverso, M., Mankaa, R., Pedalá, M. C., & Covais, A. (2024). Social Hotspot analysis of the e-waste sector in Ghana and Nigeria. *Waste Management, 183,* 271-277. https://doi.org/10.1016/j.wasman.2024.05.023

United Nations Environment Programme (2024). *Artificial Intelligence (AI) end-to-end: The environmental impact of the full AI lifecycle needs to be comprehensively assessed.* Nairobi. https://wedocs.unep.org/items/5f3afe87-5419-439b-a6ee-14240b2605e9

Venne, J.-F. (2025). *La guerre des puces.* Gestion HEC Montréal. https://www.revuegestion.ca/la-guerre-des-puces

Vries-Gao, A. de. (2026). The carbon and water footprints of data centers and what this could mean for artificial intelligence. *Patterns, 7*(1), 101430. https://doi.org/10.1016/j.patter.2025.101430

World Economic Forum & Accenture. (2025). Advancing Responsible AI Innovation: A Playbook. World Economic Forum. https://reports.weforum.org/docs/WEF_Advancing_Responsible_AI_Innovation_A_Playbook_2025.pdf

Annexes



Matrice

Échelle de bonne pratique

Nb. section	Bonne pratique	Effort de mise en œuvre	Réduction potentielle des impacts environ-nementaux	Somme
3.1.	**Définir et remettre en question le besoin pour l'IA**			
3.1.1	Énoncer et consigner clairement la problématique ou le besoin auquel répondre.	5 (Facile)	4 (Modéré-élevé)	9
3.1.2	Constituer une équipe pluridisciplinaire pour soutenir la démarche.	4 (Plutôt facile)	3 (Modéré)	7
3.1.3	Identifier des solutions (sans et avec IA) permettant de répondre à la problématique identifiée.	3 (Moyen)	5 (Élevé)	8
3.1.4	Évaluer les effets (environnementaux, sociaux, économiques) positifs et négatifs associés à ces solutions.	4 (Plutôt facile)	5 (Élevé)	9
3.1.5	Consulter le marché pour connaître les options potentielles et leurs caractéristiques.	3 (Moyen)	s.o.	s.o.
3.1.6	Déterminer si le système (ou les fonctionnalités) d'IA est une solution pertinente.	3 (Moyen)	5 (Élevé)	8
3.1.7	Déterminer le type de solution IA approprié pour votre besoin.	4 (Plutôt facile)	4 (Modéré-élevé)	8
3.2.	**Établir un cadre de gouvernance responsable**			
3.2.1	Établir un cadre de gouvernance multi-acteurs.	2 (Plutôt difficile)	4 (Modéré-élevé)	6
3.2.2	Déterminer un processus pour consulter et informer le public de la mise en place du système d'IA.	3 (Moyen)	3 (Modéré)	6
3.2.3	Mettre en place des mécanismes de contrôle du système d'IA.	3 (Moyen)	4 (Modéré-élevé)	7
3.2.4	Établir un registre des responsabilités précisant qui est responsable des différents aspects de la solution d'IA.	4 (Plutôt facile)	3 (Modéré)	7

Nb. section	Bonne pratique	Effort de mise en œuvre	Réduction potentielle des impacts environnementaux	Somme
3.3.	**Anticiper les impacts potentiels de l'IA**			
3.3.1	Évaluer qualitativement les impacts environnementaux de l'IA sur l'ensemble de son cycle de vie, dès la phase de conception.	3 (Moyen)	5 (Élevé)	8
3.3.2	Analyser l'impact des solutions d'IA sur les équipements informatiques.	4 (Plutôt facile)	5 (Élevé)	9
3.3.3	Déterminer des points et seuils de décision associés à l'évaluation des impacts.	3 (Moyen)	5 (Élevé)	8
3.3.4	Vérifier que l'implantation de la technologie ne compromet pas les engagements climatiques et de développement durable de l'organisation.	2 (Plutôt difficile)	5 (Élevé)	7
3.4.	**Préparer le cahier des charges**			
3.4.1	Considérer des méthodes d'achat public agiles et innovantes. Définir des critères axés sur les résultats.	3 (Moyen)	s.o.	s.o.
3.4.2	Identifier et prévoir les stratégies qui permettent d'atténuer les principaux risques identifiés tout au long du cycle de vie.	4 (Plutôt facile)	5 (Élevé)	9
3.4.3	**THÈME : TRANSPARENCE ENVIRONNEMENTALE** Privilégier un fournisseur ayant réalisé une ACV multicritère de la solution d'IA.	1 (Difficile)	5 (Élevé)	6
3.4.4	**THÈME : TRANSPARENCE ENVIRONNEMENTALE** Exiger les données nécessaires à la mesure de la performance environnementale.	5 (Facile)	5 (Élevé)	10
3.4.5	**THÈME : HÉBERGEMENT** Privilégier les centres de données situés dans une région au mix électrique faiblement carboné ou avec un facteur d'énergie renouvelable (Renewable Energy Factor [REF]) élevé.	5 (Facile)	5 (Élevé)	10
3.4.6	**THÈME : HÉBERGEMENT** Privilégier les centres de données avec une bonne performance environnementale.	3 (Moyen)	4 (Modéré-élevé)	7
3.4.7	**THÈME : HÉBERGEMENT** Privilégier un hébergeur engagé dans une démarche de réduction de l'impact environnemental.	3 (Moyen)	4 (Modéré-élevé)	7

Nb. section	Bonne pratique	Effort de mise en œuvre	Réduction potentielle des impacts environ-nementaux	Somme
3.4.	**Préparer le cahier des charges**			
3.4.8	THÈME : COLLECTE ET GESTION DES DONNÉES Limiter les données aux seules informations nécessaires sur le cycle de vie de l'IA.	1 (Difficile)	3 (Modéré)	4
3.4.9	THÈME : COLLECTE ET GESTION DES DONNÉES Privilégier la qualité des données plutôt que la quantité.	3 (Moyen)	5 (Élevé)	8
3.4.10	THÈME : COLLECTE ET GESTION DES DONNÉES Optimiser le nettoyage, l'annotation et la préparation des données.	1 (Difficile)	5 (Élevé)	6
3.4.11	THÈME : COLLECTE ET GESTION DES DONNÉES Adapter les stratégies de stockage selon l'usage des données.	3 (Moyen)	3 (Modéré)	6
3.4.12	THÈME : COLLECTE ET GESTION DES DONNÉES Limiter la conservation des données collectées sur le cycle de vie du système d'IA, particulièrement pendant l'utilisation du modèle.	5 (Facile)	5 (Élevé)	10
3.4.13	THÈME : CARACTÉRISTIQUES DU MODÈLE Privilégier les modèles existants, affinés, plutôt que bâtis de toutes pièces.	3 (Moyen)	5 (Élevé)	8
3.4.14	THÈME : CARACTÉRISTIQUES DU MODÈLE Choisir le modèle le plus léger possible (c.-à-d. avec moins de paramètres).	3 (Moyen)	5 (Élevé)	8
3.4.15	THÈME : CARACTÉRISTIQUES DU MODÈLE Privilégier les modèles économes en énergie.	5 (Facile)	5 (Élevé)	10
3.4.16	THÈME : CARACTÉRISTIQUES DU MODÈLE Optimiser la performance du modèle grâce à des techniques de compression.	3 (Moyen)	3 (Modéré)	6
3.4.17	THÈME : CARACTÉRISTIQUES DU MODÈLE Réduire la consommation énergétique de l'inférence en privilégiant le traitement par lots (batch).	3 (Moyen)	3 (Modéré)	6
3.4.18	THÈME : CARACTÉRISTIQUES DU MODÈLE Mettre en cache les fragments de requêtes répétitifs pour éviter des calculs inutiles.	3 (Moyen)	3 (Modéré)	6

Nb. section	Bonne pratique	Effort de mise en œuvre	Réduction potentielle des impacts environ-nementaux	Somme
3.4.	**Préparer le cahier des charges**			
3.4.19	THÈME : (RÉ)ENTRAÎNEMENT Pendant le (ré)entraînement, trouver l'équilibre entre les ressources mobilisées pour l'entraînement, pour améliorer l'efficacité du modèle d'IA, et les gains de performance obtenus.	1 (Difficile)	3 (Modéré)	4
3.4.20	THÈME : (RÉ)ENTRAÎNEMENT Privilégier les phases d'entraînement sur des périodes de sous-utilisation des infrastructures ou sur des périodes pendant lesquelles l'intensité carbone du réseau électrique est faible.	1 (Difficile)	3 (Modéré)	4
3.4.21	THÈME : (RÉ)ENTRAÎNEMENT Minimiser le nombre de réentraînements du modèle.	3 (Moyen)	5 (Élevé)	8
3.4.22	THÈME : INFÉRENCE (REQUÊTES) Privilégier les modèles qui sont exécutables en local sur les terminaux utilisateurs.	3 (Moyen)	4 (Modéré-élevé)	7
3.4.23	THÈME : INFÉRENCE (REQUÊTES) Donner la possibilité d'activer ou de désactiver l'appel à l'IA.	5 (Facile)	5 (Élevé)	10
3.4.24	THÈME : INFÉRENCE (REQUÊTES) Programmer le modèle pour limiter la taille des réponses générées par l'IA (output token).	5 (Facile)	3 (Modéré)	8
3.4.25	THÈME : AUTRES PRATIQUES D'ÉCOCONCEPTION Adopter ou exiger d'autres pratiques d'écoconception.	3 (Moyen)	5 (Élevé)	8
3.5.	**Former les équipes au cadre d'usage**			
3.5.1	Inclure dans la formation des personnes utilisatrices des notions liées au développement durable et aux risques environnementaux, sociaux et éthiques de l'IA.	5 (Facile)	5 (Élevé)	10
3.5.2	Préciser les profils et les cas d'utilisation autorisés afin d'éviter les utilisations non essentielles.	3 (Moyen)	4 (Modéré-élevé)	7
3.5.3	Inviter les personnes utilisatrices à minimiser le nombre de requêtes demandées à l'IA.	5 (Facile)	5 (Élevé)	10

Nb. section	Bonne pratique	Effort de mise en œuvre	Réduction potentielle des impacts environ-nementaux	Somme
3.5.	**Former les équipes au cadre d'usage**			
3.5.4	Sensibiliser le personnel aux fonctionnalités intensives en ressources.	5 (Facile)	5 (Élevé)	10
3.5.5	Informer la personne utilisatrice de l'impact environnemental de sa requête.	1 (Difficile)	3 (Modéré)	4
3.6.	**Mesurer et divulguer la performance**			
3.6.1	S'assurer que le modèle reste performant, précis et sobre une fois déployé, et que les pratiques d'exploitation évoluent dans le bon sens.	3 (Moyen)	5 (Élevé)	8
3.6.2	Mettre en place un suivi et une reddition de comptes d'indicateurs de durabilité.	1 (Difficile)	5 (Élevé)	6
3.6.3	Éviter de créer une dépendance à des modèles trop lourds et remettre en question le besoin de maintenir les outils IA au fil du temps.	3 (Moyen)	5 (Élevé)	8
3.6.4	Mettre en place une stratégie de réutilisation et de seconde vie du modèle d'IA en interne et/ou dans la communauté.	2 (Plutôt difficile)	5 (Élevé)	7

Liste de contrôle

Section	Liste de contrôle pour des acquisitions (éco)responsables de solutions d'IA	
3.1.	**Lors de la définition du besoin, avez-vous...**	
3.1.1	...énoncé et consigné clairement la problématique ou le besoin à combler (la raison de l'acquisition) ?	☐
3.1.2	...mis en place une équipe pluridisciplinaire pour vous soutenir pendant toute la démarche ? (dans la définition du besoin et des fonctionnalités, les enjeux à prendre en compte, les critères d'acquisition, la contribution pendant la conception et déploiement, etc.)	☐
3.1.3	...identifié différentes solutions (sans et avec IA) qui pourraient répondre à la problématique ou combler le besoin identifié ?	☐
3.1.4	...fait une évaluation des effets (environnementaux, sociaux, économiques) positifs et négatifs associés à ces solutions afin de prendre des décisions éclairées et alignées avec les objectifs de l'organisation ?	☐
3.1.5	...consulté le marché pour connaître des options potentielles et leurs caractéristiques et pratiques responsables ?	☐
3.1.6	...determiné si le système (ou les fonctionnalités) d'IA est une solution pertinente ?	☐
3.1.7	...déterminé le type de solution d'IA approprié pour votre besoin ?	☐
3.2.	**Pour la bonne gestion et gouvernance dans l'acquisition d'une solution d'IA, avez-vous...**	
3.2.1	...établi un cadre de gouvernance multi-acteurs avec des parties prenantes internes et externes (intégrant l'équipe pluridisciplinaire interne) pour soutenir la prise de décision liées à l'acquisition et déploiement de la solution ?	☐
3.2.2	...déterminé un processus pour consulter et informer le public de la mise en place du système IA, si la solution peut les concerner ?	☐
3.2.3	...mis en place des mécanismes de contrôle du système IA ?	☐
3.2.4	...établi un registre clair des responsabilités qui précise qui est responsable des différents aspects de l'IA de la solution à acquérir ?	☐

Section	Liste de contrôle pour des acquisitions (éco)responsables de solutions d'IA	
3.3.	**Pour bien cerner les impacts potentiels de l'IA et pouvoir les éviter ou mitiger, avez-vous…**	
3.3.1	…évalué les impacts environnementaux des solutions d'IA sur tout leur cycle de vie afin de prendre des décisions éclairées et alignées avec les politiques de l'organisation ? (impacts de 1er, 2e et 3e ordre)	☐
3.3.2	…analysé l'impact de la solution d'IA sur les équipements informatiques afin d'éviter le renouvellement prématuré du matériel ? (connexion à bas débit ou hors connexion, effet sur les terminaux, systèmes d'exploitation, navigateurs web)	☐
3.3.3	…déterminé les points et seuils de décision associés à l'évaluation des impacts à respecter si des changements substantiels à la conception du système d'IA sont effectués ?	☐
3.3.4	…vérifié que l'implantation de la solution ne compromette pas les engagements climatiques et de développement durable de l'organisation ?	☐
3.4.	**Dans la préparation du cahier des charges, avez-vous…**	
3.4.1	…considéré des méthodes d'achat agiles, innovantes et/ou basées sur les résultats (outcomes-based) dans les conditions d'acquisition et d'exécution du contrat ?	☐
3.4.2	…identifié et prévu les stratégies qui permettent d'atténuer les principaux risques identifiés tout au long du cycle de vie ?	☐
	…défini des critères liés à la transparence environnementale du fournisseur ? Notamment :	
3.4.3	Privilégier un fournisseur ayant réalisé une ACV multicritère de la solution IA.	☐
3.4.4	Exiger la transparence de la performance environnementale de la solution d'IA tout au long de son cycle de vie.	☐
	…défini des critères pour réduire les impacts environnementaux de l'hébergement de la solution d'IA ? Notamment :	
3.4.5	Le facteur d'énergie renouvelable de(s) site(s) d'hébergement (mix électrique renouvelable de la région ou REF calculé selon l'ISO/IEC 30134-3 ou équivalent).	☐
3.4.6	La performance environnementale de(s) site(s) d'hébergement (indicateurs PUE, CUE, WUE, ERF, autres) (ou au moins la capacité des soumissionnaires de fournir l'information).	☐
3.4.7	Les actions du soumissionnaire (ou l'hébergeur si sous-traité) de réduction de l'impact environnemental de(s) site(s) d'hébergement (en lien à l'emplacement des sites et les impacts sur les écosystèmes et la biodiversité, à la consommation d'énergie et d'émission de GES, à la consommation d'eau et son retour au milieu en conditions de qualité, à l'acquisition et gestion responsable des équipements pour allonger leur vie utile, etc.).	☐

Section	Liste de contrôle pour des acquisitions (éco)responsables de solutions d'IA	
3.4.	**Dans la préparation du cahier des charges, avez-vous…**	
	…défini des critères pour réduire les impacts environnementaux en lien avec la collecte et la gestion des données ? Notamment :	
3.4.8	La limitation des données aux informations nécessaires pour atteindre la performance attendue de la solution (type de données, utilisation pendant le (ré)entraînement, test et inférence, critères de sélection).	☐
3.4.9	La sélection des données pour prioriser la qualité à la quantité (provenance des données sur le cycle de vie, justification).	☐
3.4.10	Les actions de nettoyage, d'annotation et de préparation des données pour optimiser l'entraînement et éviter des entraînements excessivement longs (procédure de préparation des données, justification).	☐
3.4.11	Les stratégies de stockage différenciées des données en fonction de leur fréquence d'utilisation (stratégies, choix technologiques, ajustement du volume des données chaudes et froides, etc.).	☐
3.4.12	Les actions pour vérifier la pertinence des données générées pendant l'usage du modèle (métadonnées, registres, etc.) et limiter leur collecte et leur conservation.	☐
	…défini des critères liés aux caractéristiques du modèle pour réduire ses impacts environnementaux ? Notamment :	
3.4.13	L'affinage de modèles existants plutôt que le développement d'une solution de toutes pièces.	☐
3.4.14	L'utilisation de modèles légers (c.-à-d. avec moins de paramètres) (petit modèle de langage [SLM] plutôt qu'un grand modèle de langage [LLM]).	☐
3.4.15	La sélection de modèles qui ont une meilleure efficacité énergétique comparée à d'autres (justification du modèle à l'appui d'outils de classement comparatif existants).	☐
3.4.16	Les actions d'optimisation du modèle (explication des techniques utilisées, justification, collaboration avec les utilisateurs).	☐
3.4.17 et 3.4.18.	Les actions pour réduire la consommation énergétique de l'inférence, telles que le traitement par lots, la mise en cache de fragments de requêtes répétées ou d'autres (explication des techniques utilisées, justification).	☐
	…défini des critères pour réduire les impacts environnementaux pendant le (ré)entraînement du modèle ? Notamment :	
3.4.19	Les actions (techniques ou approches) mises en place pour arrêter l'entraînement du modèle lorsqu'il a atteint le niveau de performance optimal établi pour celui-ci.	☐
3.4.20	Les mesures pour privilégier l'entraînement pendant les périodes de sous-utilisation des infrastructures ou de faible intensité carbone du réseau électrique.	☐

Section	Liste de contrôle pour des acquisitions (éco)responsables de solutions d'IA	
3.4.	**Dans la préparation du cahier des charges, avez-vous…**	
	…défini des critères pour réduire les impacts environnementaux pendant le (ré)entraînement du modèle ? Notamment :	
3.4.21	Les besoins de réentraînement du modèle et les stratégies pour les limiter (fréquence, raisons justifiant chaque réentraînement, mécanismes de suivi et d'évaluation de l'efficacité des réentraînements, évaluation qualitative et quantitative des améliorations, impacts environnementaux associés).	☐
	…défini des critères pour réduire les impacts environnementaux de l'inférence ? Notamment :	
3.4.22	La vérification que les modèles puissent être exécutés en local sur les terminaux utilisateurs, sans que ceux-ci soient bonifiés ?	☐
3.4.23	La possibilité d'activer ou de désactiver l'appel à l'IA en fonction de différents profils d'utilisation établis par l'organisation.	☐
3.4.24	La possibilité de limiter la taille des réponses générées par l'IA (output token).	☐
	…exploré la possibilité d'intégrer d'autres pratiques d'écoconception d'une solution numérique ?	
3.4.25	L'adoption ou l'intégration des critères en écoconception basés sur des référentiels existants.	☐
3.5.	**Pour assurer une utilisation responsable de la solution (ou fonctionnalités) d'IA, avez-vous…**	
3.5.1	…inclus dans la formation des personnes utilisatrices des notions liées au développement durable et aux risques environnementaux (sociaux et éthiques de l'IA) ?	☐
3.5.2	…précisé les profils et les cas d'utilisation autorisés afin d'éviter les utilisations non essentielles ?	☐
3.5.3	…invité les personnes utilisatrices à minimiser le nombre de requêtes demandées à l'IA ?	☐
3.5.4	…sensibilisé le personnel aux fonctionnalités intensives en ressources ?	☐
3.5.5	…assuré que la personne utilisatrice soit informée de l'impact environnemental de sa requête ?	☐

Section	Liste de contrôle pour des acquisitions (éco)responsables de solutions d'IA	
3.6.	**Pour assurer la mesure et la divulgation de la performance de la solution d'IA et de son utilité, avez-vous...**	
3.6.1	...défini des exigences pour s'assurer que le modèle reste performant, précis et sobre une fois déployé, et que les pratiques d'exploitation évoluent dans le bon sens ?	☐
3.6.2	...mis en place un suivi et une reddition de comptes d'indicateurs de durabilité ?	☐
3.6.3	...prévu un processus de révision et de remise en question du besoin afin d'éviter la dépendance à des modèles trop lourds ou aux outils d'IA ?	☐
3.6.4	...mis en place une stratégie de réutilisation et de seconde vie du modèle d'IA en interne ou au sein de la communauté ?	☐

Références utilisées pour la matrice de bonnes pratiques

Arcep, & Arcom. (2024). Référentiel général d'écoconception de services numériques. https://ecoresponsable.numerique.gouv.fr/publications/referentiel-general-ecoconception/

Association Française de Normalisation (AFNOR). (2024). Référentiel général pour l'IA frugale (Version AFNOR SPEC 2314). https://www.afnor.org/actualites/referentiel-pour-mesurer-et-reduire-impact-environnemental-de-ia/

Association GreenIT. (2026). Référentiel de bonnes pratiques d'utilisation de l'intelligence artificielle générative. https://ria.greenit.fr/fr/fiches

Commission européenne. (2026). Le code de bonnes pratiques de l'IA à usage général. https://digital-strategy.ec.europa.eu/fr/policies/contents-code-gpai

Commission Nationale de l'Informatique et des Libertés [CNIL]. (2024). IA : Tenir compte de la protection des données dans la collecte et la gestion des données. CNIL. https://www.cnil.fr/fr/tenir-compte-de-la-protection-des-donnees-dans-la-collecte-et-la-gestion-des-donnees

CNIL. (2025). Développement des systèmes d'IA : Que faut-il vérifier? [Liste de contrôle]. https://www.cnil.fr/sites/default/files/2025-07/ia_liste_de_verification.pdf

Conseil économique, social et environnemental. (2024). Impacts de l'intelligence artificielle : Risques et opportunités pour l'environnement. https://www.lecese.fr/travaux-publies/impacts-de-lintelligence-artificielle-risques-et-opportunites-pour-lenvironnement

Conseil national du numérique, ADEME, Arcep, IGN, ECOLAB (Commissariat Général au Développement Durable, Direction générale des entreprises (ministère de l'Économie), DReDIS de l'Institut de recherche juridique de la Sorbonne, Inria, Sciences Po, & Thomas Le Goff (Telecom Paris). (2025, mai). IA & Environnement [Module pédagogique]. https://cafeia.org/ia-et-environnement/

Data For Good. (2023, juillet). Les grands défis de l'IA générative. https://issuu.com/dataforgood/docs/dataforgood_livreblanc_iagenerative_v1.0

Gouvernement du Canada. (2024). Principes directeurs pour l'utilisation de l'IA au gouvernement. https://www.canada.ca/fr/gouvernement/systeme/gouvernement-numerique/innovations-gouvernementales-numeriques/utilisation-responsable-ai/principes.html

Gouvernement du Canada. (2025). Comprendre les effets environnementaux de l'IA (DDN2-A60). Gouvernement du Canada. https://www.csps-efpc.gc.ca/tools/articles/ai-environmental-effects-fra.aspx

Institut du Numérique Responsable (INR). (2021). Ministère de la transition écologique. Guide pratique pour des achats numériques responsables. https://ecoresponsable.numerique.gouv.fr/publications/guide-pratique-achats-numeriques-responsables/

Jeanquartier, F., Jean-Quartier, C., Rieder, P., Misirlić, V., Pasero, C., Hohensinner, R., Müller, H., & Holzinger, A. (2026). Assessing the carbon footprint of language models: Towards sustainability in AI. Resources, Conservation and Recycling, 226, 108670. https://doi.org/10.1016/j.resconrec.2025.108670

Lefèvre, L., Logizat, A.-L., Trystram, D., Bouveret, S., Bugeau, A., Combaz, J., Frenoux, E., Guennebaud, G., Lefèvre, J., & Nicolaï, J.-P. (2022). Proposition de document de cadrage Évaluation environnementale de projets impliquant des méthodes d'IA (Version 1). EcoInfo. https://hal.science/hal-03853135

Ministère de la Culture. (s. d.). Compar:IA - Le comparateur d'IA conversationnelles [Outil d'évaluation d'intelligence artificielle]. https://comparia.beta.gouv.fr

Gouvernement du Québec. (2025). Énoncé de principes pour une utilisation responsable de l'intelligence artificielle par les organismes publics. https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/cybersecurite_numerique/Publications/Strategie_IA/Enonce_IA.pdf

National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1

OCDE. (2024). Principes de l'IA. https://www.oecd.org/fr/themes/principes-de-l-ia.html

Pham, N. T., Kieu, T., Nguyen, D. M., Xuan, S. H., Duong-Trung, N., & Le-Phuoc, D. (2025). SLM-Bench: A Comprehensive Benchmark of Small Language Models on Environmental Impacts. Dans C. Christodoulopoulos, T. Chakraborty, C. Rose, & V. Peng (Éds.), Findings of the Association for Computational Linguistics: EMNLP 2025 (p. 21369-21392). Association for Computational Linguistics. https://doi.org/10.18653/v1/2025.findings-emnlp.1165

United Nations Environment Programme (UNEP). (2024). Artificial intelligence (AI) end-to-end: The environmental impact of the full AI life cycle needs to be comprehensively assessed. https://wedocs.unep.org/rest/api/core/bitstreams/07b3c8fc-bd30-4b92-b5f4-d665e927b59d/content


ceimia


futurearth


IVADO