



ÉCOLE NATIONALE D'ADMINISTRATION PUBLIQUE

**L'éthique de l'intelligence artificielle : les principes et les mesures qui
pourraient inspirer l'élaboration d'un cadre éthique dans
l'administration publique québécoise**

Rapport de stage

PRÉSENTÉ À

M. GUILLAUME BERGERON ET M. CHRISTIAN BOUDREAU

COMME EXIGENCE AU COURS

ENP7969 S – Stage (9 crédits)

Activité d'intégration obligatoire

PAR

DANIELA SILVESTRE PINHEIRO

SEPTEMBRE 2021-JANVIER 2022

Table des matières

Remerciements	iii
Liste d'abréviations et sigles	iv
Liste des annexes	v
Résumé / Summary	vi
1. Introduction	7
2. Description de l'organisation et du mandat confié à l'étudiante	8
2.1. Contexte du stage et description de l'organisation	8
2.2. Mandat du stage	11
2.3. Activités prévues	11
2.4. Biens livrables	11
3. Recension des écrits	13
3.1. La définition de l'intelligence artificielle	13
3.2. Les avantages et les inconvénients	14
3.3. Une IA éthique, digne de confiance	15
3.4. IA éthique dans le monde	16
3.5. Les dimensions pour les principes éthiques en IA	20
3.6. Les outils pour une IA digne de confiance	21
3.7. Les questions de recherche	23
4. Méthodologie de recherche et d'intervention	24
5. Résultats	29
6. Discussion et/ou recommandations	39
7. Conclusion	41
8. Bibliographie	42
9. Annexes	49

Remerciements

La rédaction de ce rapport dans le cadre du stage que j'ai effectué pour la fin de ma maîtrise à l'École Nationale d'Administration Publique (ENAP) est le résultat du soutien incessant et de la contribution précieuse de plusieurs personnes que je tiens à remercier sincèrement :

M. Marc Vézina, Directeur de l'architecture d'entreprise gouvernementale, et M. Guillaume Bergeron, Conseiller pour la Stratégie d'introduction à l'intelligence artificielle et responsable du stage, pour leur accueil chaleureux et l'échange enrichissant d'expériences et de connaissances ;

Aux autres collaborateurs de la Direction de l'architecture d'entreprise gouvernementale (DAEG), pour la disposition, la sympathie et le soutien qui, avec les autres, m'ont fait me sentir bien accueillie dans l'organisation même dans le contexte du télétravail ;

Au personnel du Secrétariat du Conseil du trésor (SCT), pour le premier accueil au Sous-secrétariat du Dirigeant principal de l'information et de la transformation numérique (SSDPITN) ;

Au personnel du ministère de la Cybersécurité et du Numérique (MCN), en particulier toute l'équipe du Sous-ministériat adjoint à la transformation numérique gouvernementale (SMATNG), pour la démonstration d'engagement, de performance et de dévouement pour la bonne exécution du travail ;

M. Christian Boudreau, mon conseiller académique, pour avoir consacré son temps à me fournir des conseils et des orientations de grande valeur pour la réalisation de ce travail ;

Au personnel de l'ENAP pour sa compétence, sa responsabilité et sa courtoisie envers toute la communauté des étudiants et des professionnels ;

À mon mari Diógenes et à mes filles Giovanna et Isabella, qui se sont embarqués avec moi dans ce projet personnel et professionnel, en déménageant du Brésil au Canada, et qui ont été mon refuge où j'ai trouvé la force et le courage de continuer ;

Je suis profondément reconnaissante pour tout !

Liste d'abréviations et sigles

DAEG :	Direction de l'architecture d'entreprise gouvernementale
ENAP :	École Nationale d'Administration Publique
GEHN IA :	Groupe d'experts de haut niveau en intelligence artificielle
IA :	Intelligence artificielle
MCN :	ministère de la Cybersécurité et du Numérique
OCDE :	Organisation de coopération et de développement économiques
SCT :	Secrétariat du Conseil du trésor
SIA :	Système d'intelligence artificielle
SMATNG :	Sous-ministériat adjoint à la transformation numérique gouvernementale
SSDPITN :	Sous-secrétariat du Dirigeant principal de l'information et de la transformation numérique
STRATÉGIE IA :	Stratégie d'intégration de l'intelligence artificielle dans l'administration publique 2021-2026
UE :	Union Européenne
UNESCO :	<i>United Nations Educational, Scientific and Cultural Organisation</i>

Liste des annexes

Analyse des stratégies IA – Bloc 1

- Danemark 50
- Union Européenne 54

Analyse des stratégies IA – Bloc 2

- Finlande 58
- Lituanie 62
- Suède 65

Analyse des stratégies IA – Bloc 3

- Estonie 67
- Malte 70
- Norvège 73
- Qatar 77

Analyse des stratégies IA – Bloc 4

- Luxembourg 80
- Pays-Bas 83
- Portugal 86
- Serbie 89

Analyse des stratégies IA – Bloc 5

- Canada 91

Fiches techniques

- Déclaration de Montréal 94
- UNESCO 95
- Commission Européenne 96

Résumé / Summary

À la fin du mois de juin 2021, le gouvernement québécois a annoncé sa Stratégie d'intégration de l'intelligence artificielle dans l'administration publique pour la période 2021-2026. L'objectif est de permettre à l'administration publique d'améliorer son efficacité et d'innover dans la prestation de services au public, tout en garantissant que l'intégration de l'intelligence artificielle se fera de manière éthique, sans compromettre la vie privée et la protection des renseignements personnels.

Dans ce contexte, nous avons mené une recherche comparative à partir d'une analyse documentaire des travaux réalisés dans le domaine de l'intelligence artificielle par 13 pays sélectionnés et par l'Union européenne afin d'identifier les principes et les mesures mises en place pour le développement et l'utilisation responsables de l'intelligence artificielle qui pourraient inspirer l'élaboration d'un cadre éthique dans l'administration publique québécoise.

La présente recherche nous a permis d'identifier qu'il existe une convergence générale autour des principes éthiques suivants : transparence et explicabilité ; responsabilité ; respect de la vie privée ; justice et équité (non-discrimination) ; sûreté et sécurité (non-nuisance). Il a également été possible de noter les différentes mesures choisies pour garantir une IA digne de confiance, telles que la création de comités/conseils d'éthique, la rédaction de guides ou de lignes directrices et le développement d'outils d'évaluation des risques.

At the end of June 2021, the Quebec government announced its Strategy for the Integration of Artificial Intelligence in Public Administration for the period 2021-2026. The objective is to enable the public administration to improve its efficiency and innovate in the delivery of services to the public, while ensuring that the integration of artificial intelligence will be done in an ethical manner, without compromising privacy and the protection of personal information.

In this context, we conducted a benchmark research based on a literature review of the work done in the field of artificial intelligence by 13 selected countries and the European Union to identify the principles and measures put in place for the responsible development and use of artificial intelligence that could inspire the elaboration of an ethical framework in the Quebec public administration.

The present research has allowed us to identify that there is a general convergence around the following ethical principles: transparency and explicability; accountability; privacy; justice and equity (non-discrimination); safety and security (non-harmfulness). It was also possible to note the different measures chosen to ensure trustworthy AI, such as the creation of ethics committees/councils, the drafting of guides or guidelines and the development of risk assessment tools.

1. Introduction

La mise en œuvre de l'intelligence artificielle dans le secteur public est une étape essentielle pour les institutions gouvernementales afin de rendre leurs processus plus rapides et plus transparents, favorisant l'optimisation des processus avec une productivité et une efficacité accrue des flux d'activités et une amélioration conséquente du service aux citoyens.

D'autre part, il est nécessaire de considérer tous les défis éthiques, moraux et sociaux que l'intelligence artificielle peut présenter, et il est donc nécessaire de développer un cadre qui assure le développement et l'utilisation de ce type de technologie d'une manière responsable, éthique et transparente, augmentant ainsi les avantages que l'intelligence artificielle présente tout en atténuant ses risques.

Conscient de cette complexité, le gouvernement du Québec a lancé sa Stratégie d'intégration de l'intelligence artificielle dans l'administration publique pour la période 2021-2026. L'objectif est de permettre à l'administration publique d'améliorer son efficacité et d'innover dans la prestation de services au public, tout en garantissant que l'intégration de l'intelligence artificielle (IA) se fera de manière éthique, sans compromettre la vie privée et la protection des informations personnelles.

Le gouvernement québécois vise alors le déploiement d'un cadre pour le développement et l'utilisation responsables de l'intelligence artificielle. L'objectif de ce rapport de stage est de présenter les résultats obtenus à partir de l'analyse comparative des travaux réalisés par d'autres gouvernements et organisations dans le domaine de l'IA afin d'identifier les principes et les mesures qui pourraient inspirer l'élaboration d'un cadre éthique dans l'administration publique québécoise.

Conformément à son objectif principal, ce rapport est structuré comme suit : (1) description de l'organisation et du mandat qui a été confié à l'étudiante ; (2) recension des écrits ; (3) méthodologie de recherche et d'intervention ; (4) résultats ; (5) discussion et/ou recommandations ; (6) conclusion.

2. Description de l'organisation et du mandat

2.1. Contexte du stage et description de l'organisation

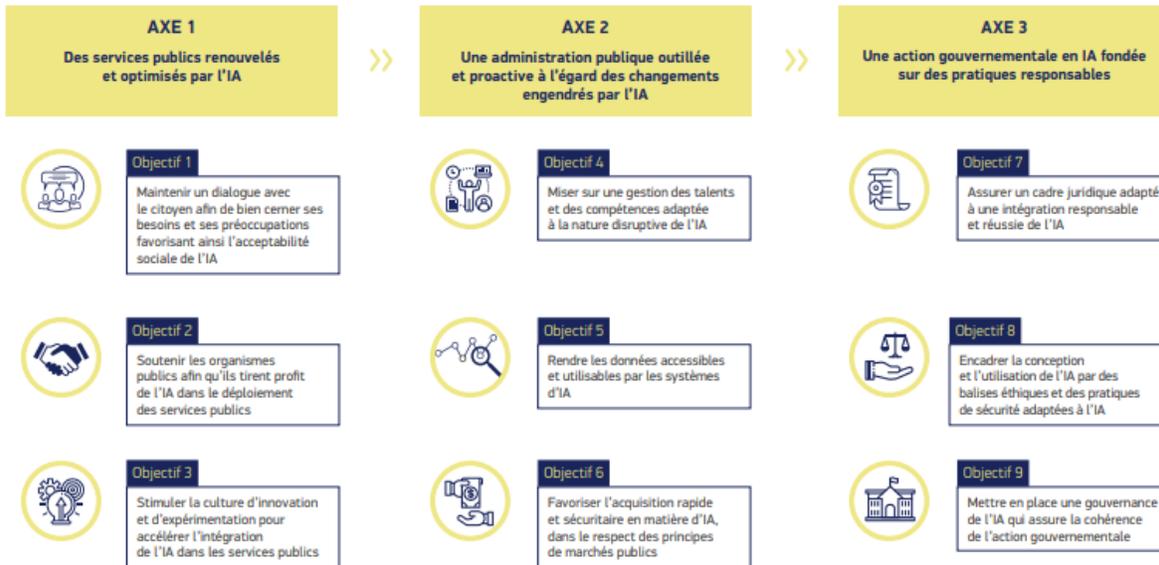
Le gouvernement du Québec a adopté, en juin 2019, la Stratégie de transformation numérique gouvernementale 2019-2023 avec l'objectif d'offrir des services publics intuitifs et faciles d'utilisation pour la population et d'améliorer l'efficacité de l'État (Québec, s.d.).

La sixième mesure clé de cette stratégie gouvernementale consiste à adopter une stratégie d'introduction à l'intelligence artificielle afin de placer les bases nécessaires à la compréhension des concepts et des technologies associées, soutenir son utilisation et baliser son usage auprès des organisations publiques (Québec, 2019, p. 4).

À la fin du mois de juin 2021, le gouvernement québécois a annoncé sa Stratégie d'intégration de l'intelligence artificielle dans l'administration publique pour la période 2021-2026. L'objectif est de permettre à l'administration publique d'améliorer son efficacité et d'innover dans la prestation de services au public, tout en garantissant que l'intégration de l'IA se fera de manière éthique, sans compromettre la vie privée et la protection des informations personnelles (Secrétariat du Conseil du Trésor, 2021).

Cette stratégie vise à positionner l'administration publique comme bénéficiaire exemplaire de l'IA en prenant notamment appui sur le leadership du Québec dans ce domaine. À cette fin, elle présente 5 principes, divisés en 3 axes, pour un total de 9 objectifs (Québec, 2021a).





Pour la réalisation de ces objectifs, la Stratégie IA comporte 19 mesures clés prioritaires à initier entre 2021-2023. De nouvelles mesures seront introduites annuellement à compter de 2023 (Québec, 2021b).

Parmi les mesures clés prioritaires, dans l'axe 3, on trouve la mesure 8.1 : Doter l'administration publique d'un cadre de développement et d'utilisation responsables de l'IA.



Objectif 8

Encadrer la conception et l'utilisation de l'IA par des balises éthiques et des pratiques de sécurité adaptées à l'IA

8.1 Doter l'administration publique d'un cadre de développement et d'utilisation responsables de l'IA

Un cadre de développement et d'utilisation responsables de l'IA sera adopté afin de doter l'action gouvernementale en matière d'IA de fondations solides favorisant notamment le respect de la vie privée, la cybersécurité, l'éthique et la non-discrimination.

Que ce soit en joignant sa voix à la Déclaration de Montréal pour un développement éthique de l'intelligence artificielle ou en adoptant des principes qui lui sont propres, l'administration publique québécoise s'inspirera des nombreux travaux réalisés sur le développement et l'utilisation éthiques de l'IA à travers le monde. L'adoption d'un tel cadre sera effectuée en concertation avec les organismes publics et sera appuyée par des experts de l'IA.

À terme, ces principes constitueront l'un des piliers de l'intégration de l'IA dans l'administration publique pour les années futures.

La responsabilité de la mise en œuvre de la stratégie IA incombait au Secrétariat du Conseil du trésor (SCT), où se trouvait le Sous-secrétariat du dirigeant principal de l'information et de la transformation numérique (SSDPITN).

Cette responsabilité incombe actuellement au ministère de la Cybersécurité et du Numérique (MCN) qui a été récemment créé par le Projet de loi n° 6 avec pour mission « de diriger et de coordonner les actions de l'État dans les domaines de la cybersécurité et du numérique, de proposer au gouvernement les grandes orientations en ces domaines, de déterminer les secteurs d'activités où il entend agir en priorité et de proposer au gouvernement des mesures en vue d'accroître l'efficacité de la lutte contre les cyberattaques et les cybermenaces au Québec » (Québec, s.d.).

En effet, le projet de loi précité, en plus de confier au ministre de la Cybersécurité et du Numérique les responsabilités qui étaient dévolues à Infrastructures technologiques Québec, lui a confié les fonctions du président du Conseil du trésor en matière de ressources informationnelles, notamment en matière de transformation numérique, de sécurité de l'information et de données numériques gouvernementales (Éditeur officiel du Québec, 2021, p. 2).

Dans ce contexte, le MCN, par la Direction de l'architecture d'entreprise gouvernementale (DAEG) qui se trouve au sein du Sous-ministériat adjoint à la transformation numérique gouvernementale (SMATNG), vise à élaborer un cadre de développement et d'utilisation responsables de l'IA en plus de positionner le gouvernement face à la Déclaration de Montréal sur l'IA responsable.

C'est dans ce cadre que s'inscrit cette proposition de stage.

2.2. Mandat du stage

Spécifiquement, la stagiaire devra, en premier lieu, prendre connaissance de la documentation liée à la Stratégie IA pour participer ensuite aux différents travaux de l'action portant sur le développement et l'utilisation responsables de l'IA, plus précisément :

- Le positionnement du gouvernement du Québec en lien avec la Déclaration de Montréal sur l'IA responsable ;
- L'élaboration d'un cadre de développement et d'utilisation responsable de l'IA dans l'administration publique québécoise, par l'entremise de balises éthiques et de pratiques de sécurité (en matière de cybersécurité, de confidentialité et de protection des renseignements personnels) adaptées à l'IA.

Ainsi, la stagiaire devra procéder à une analyse comparative des travaux réalisés par d'autres gouvernements et organisations dans le domaine de l'IA afin d'identifier les principes et les mesures qui pourraient inspirer l'élaboration d'un cadre éthique dans l'administration publique québécoise.

2.3. Activités prévues

Les activités prévues pour la réalisation du mandat consistent en :

- Produire des fiches d'analyse ;
- Participer aux rencontres d'équipe ;
- Contribuer à la rédaction des rapports ;
- Effectuer d'autres tâches spécifiques confiées par le mandant, selon le besoin du présent mandat.

2.4. Biens livrables

Trois principaux livrables sont attendus, à savoir :

- Résultat de l'analyse de la documentation liée à la Stratégie IA, notamment la Déclaration de Montréal sur l'IA responsable ;
- Compilation des données obtenues par l'analyse comparative concernant les travaux réalisés par d'autres gouvernements et organisations œuvrant dans le domaine de l'IA ;
- Le rapport final de stage.

3. Recension des écrits

Afin de mieux comprendre le contexte de notre étude, nous allons envisager la revue de la littérature au regard de plusieurs notions et dimensions clés : la définition de l'intelligence artificielle ; ses avantages et ses inconvénients ; les enjeux éthiques, moraux et sociétaux de l'IA ; les principes éthiques en IA et les instruments/outils que nous pouvons mettre en place pour parvenir à une IA digne de confiance.

3.1. La définition de l'intelligence artificielle

Bien que plusieurs définitions de l'intelligence artificielle (IA) aient vu le jour au cours des dernières décennies, nous pouvons la comprendre comme la capacité d'une machine à reproduire des comportements liés aux humains, tels que le raisonnement, l'apprentissage, la planification et la créativité (Parlement européen, 2020).

Le Pollotec (2018, p. 75) souligne que l'IA « est une discipline scientifique qui vise à simuler, en les décomposant, les différentes facultés cognitives humaines sur des ordinateurs via des algorithmes ».

Ainsi, l'IA fait référence aux « systèmes autonomes capables d'accomplir des tâches complexes que l'on croyait réservées à l'intelligence naturelle : traiter de grandes quantités d'informations, calculer et prédire, apprendre et adapter ses réponses aux situations changeantes, et reconnaître et classer des objets » (Déclaration de Montréal IA responsable, 2018, p. 7).

Le groupe d'experts indépendants de haut niveau sur l'intelligence artificielle (GEHN IA), mis en place par la Commission européenne en juin 2018, propose, quant à lui, la définition suivante pour un système d'IA, dans le document intitulé « Lignes directrices en matière d'éthique pour une IA digne de confiance » (2019, p. 48) :

Les systèmes d'intelligence artificielle (IA) sont des systèmes logiciels (et éventuellement matériels) conçus par des êtres humains et qui, ayant reçu un objectif complexe, agissent dans le monde réel ou numérique en percevant leur environnement par l'acquisition de données, en interprétant les données structurées ou non structurées collectées, en appliquant un raisonnement aux connaissances, ou

en traitant les informations, dérivées de ces données et en décidant de la (des) meilleure(s) action(s) à prendre pour atteindre l'objectif donné. Les systèmes d'IA peuvent soit utiliser des règles symboliques ou apprendre un modèle numérique, et peuvent également adapter leur comportement en analysant la mesure dans laquelle l'environnement est affecté par leurs actions préalables.

3.2. Les avantages et les inconvénients

L'IA présente de nombreux avantages tant pour le secteur privé que pour le secteur public, permettant d'accroître la prospérité humaine et d'améliorer le bien-être des individus, contribuant à une activité économique dynamique et durable, ainsi qu'à la stimulation de l'innovation et de la productivité (Déclaration de Montréal, 2018 ; OCDE, 2019 ; Commission européenne, 2019).

En effet, l'IA « constitue un progrès scientifique et technologique majeur qui peut engendrer des bénéfices sociaux considérables en améliorant les conditions de vie, la santé et la justice, en créant de la richesse, en renforçant la sécurité publique ou en maîtrisant l'impact des activités humaines sur l'environnement et le climat » (Rapport de la Déclaration de Montréal pour un développement responsable de l'intelligence artificielle, 2018, p. 7).

Les systèmes d'IA peuvent toutefois aussi présenter certains risques qui doivent être traités de manière appropriée et proportionnée (Commission européenne, 2019).

Selon le rapport de la Déclaration de Montréal pour le développement responsable de l'IA (2018, p. 7), « les machines intelligentes peuvent contraindre les choix des individus et des groupes, abaisser la qualité de vie, bouleverser l'organisation du travail et le marché de l'emploi, influencer la vie politique, entrer en tension avec les droits fondamentaux, exacerber les inégalités économiques et sociales, et affecter les écosystèmes, l'environnement et le climat ».

L'avancée de l'IA apporte donc de nombreux avantages, mais elle soulève aussi des questions éthiques, morales et sociales qui doivent être posées dès la conception d'un système d'IA (Thomasset, 2020, p. 8). Ces questions sont transdisciplinaires et doivent être analysées dans

leurs composantes robotique, juridique, sociologique, psychologique, philosophique, éthique et morale (Linden, 2020, p. 23).

Quant à la multiplication des questionnements éthiques, Thomasset (2020, p. 7) soutient que les processus de calcul et de résolution de problèmes à la base de l'IA sont particulièrement opaques et peuvent produire des biais, des discriminations et des visions du monde sans que nous puissions toujours ouvrir la « boîte noire » qui les fabrique.

Comme le précise Linden (2020, p. 16), l'IA consiste en un ensemble d'algorithmes. Ces algorithmes, même s'ils sont capables d'apprendre, travaillent sur des représentations prédéfinies du monde, des modèles toujours partiels.

Ainsi, ces enjeux peuvent générer une méfiance et une résistance à la mise en œuvre d'un système d'IA. Que faut-il faire pour résoudre ou atténuer ce problème, pour garantir le développement et l'utilisation responsables d'une IA qui inspire le plus souvent confiance auprès des parties prenantes impliquées ou concernées ? Que serait une IA digne de confiance ?

3.3. Une IA éthique digne de confiance

Pour l'OCDE, une IA éthique doit être digne de confiance et respecter les droits de la personne et la vie privée, être équitable, transparente, explicable, robuste, sûre et sécurisée, et dont les acteurs impliqués dans leur développement et leur utilisation restent responsables (OCDE, 2021, p. 6).

Selon la Commission européenne, une IA digne de confiance présente trois caractéristiques qui doivent être respectées tout au long du cycle de vie d'un système : a) elle doit être licite, en assurant le respect des législations et réglementations applicables ; b) elle doit être éthique, en assurant l'adhésion à des principes et valeurs éthiques ; et c) elle doit être robuste, sur le plan tant technique que social puisque les systèmes d'IA peuvent causer des préjudices involontaires (GEHN IA, 2018, p. 6).

En effet, il faut considérer que les considérations éthiques concernant les systèmes d'IA se rapportent à toutes les étapes du cycle de vie de ces systèmes, c'est-à-dire de la recherche, de la conception et du développement jusqu'au déploiement et à l'utilisation, en passant par la maintenance, l'exploitation, la commercialisation, le financement, le suivi et l'évaluation, la validation, la fin d'utilisation, le démantèlement et la mise hors service (UNESCO, 2021, p. 5).

Ainsi, la notion d'éthique nous conduit à une réflexion plus large que l'analyse de la conformité d'une action aux lois et normes existantes. Comme l'explique le juge à la Cour suprême des États-Unis Potter Stewart (cité dans Abid, 2021, p. 101), « l'éthique, c'est le fait de connaître la différence entre ce qu'on a le droit de faire et ce qu'il est juste de faire ».

Comme souligne Abid (2021, p. 101),

Déterminer ce qu'une organisation a le droit de faire va de soi grâce aux lois et aux cadres réglementaires ; toutefois, ce qu'il est juste de faire est plus délicat à définir. Dans le domaine de la technologie, ces deux frontières sont encore plus floues, car il existe une infinité d'utilisations possibles qui respectent le cadre juridique mais qui, parfois, ne respectent pas le système de valeurs dominant.

Avec ces considérations mentionnées ci-dessus, nous pouvons avoir une première idée des éléments qui doivent être présents pour assurer l'éthique en IA afin d'inspirer la confiance de toutes les parties impliquées dans le développement et l'utilisation d'un système d'IA.

3.4. IA éthique dans le monde

De nombreuses initiatives ont été lancées à travers le monde pour établir des principes généraux et autres paramètres éthiques à adopter par les acteurs publics et privés en la matière, au moyen de codes de conduite, de manuels de bonnes pratiques et de lignes directrices de haut niveau.

Parmi les différentes initiatives en ce sens, nous avons, par exemple, les instruments suivants, élaborés par des experts issus des secteurs les plus divers et qui sont les plus cités dans les stratégies nationales ou dans d'autres documents traitant de l'IA :

- **Les principes d'Asilomar**

Les principes d'Asilomar sont 23 directives pour la recherche et le développement de l'IA. Ce document expose les problèmes de développement de l'IA, l'éthique et les lignes directrices pour le développement de l'IA bénéfique. Les principes ont été créés lors de la conférence Asilomar sur l'IA bénéfique organisée par la *Future of Life Institute* (une organisation à but non lucratif), en 2017 à Pacific Grove, en Californie (Whatls.com, 2019). Parmi ces principes et autres considérations éthiques, nous soulignons : la sécurité, la transparence, la responsabilité, les valeurs humaines, les données personnelles, la liberté et la vie privée et le contrôle humain.

- **IEEE - *Ethically Aligned Design* (EAD)**

L'institut des ingénieurs électriciens et électroniciens (IEEE), une très large association professionnelle, présente dans le document intitulé *Ethically Aligned Design* (EAD) sa vision sur la manière dont les technologies intelligentes et autonomes peuvent être alignées sur des principes éthiques ou moraux et des recommandations qui donnent priorité au bien-être humain. En résumé, les principes généraux présentés dans le document sont les suivants (IEEE, 2017, p. 18) :

1. **Droits de la personne** : Le système d'intelligence artificielle (SIA) sera créé et géré de manière à respecter, promouvoir et protéger les droits de la personne internationalement reconnus.
2. **Bien-être** : Les créateurs de SIA doivent adopter l'amélioration du bien-être humain comme critère principal de réussite du développement.
3. **La gestion des données** : Les créateurs de SIA doivent donner aux individus la possibilité d'accéder à leurs données et de les partager en toute sécurité, afin de maintenir la capacité des personnes à contrôler leur identité.
4. **Efficacité** : Les créateurs et les exploitants d'un système d'information géographique doivent fournir des preuves de l'efficacité et de l'adéquation aux objectifs du système.
5. **Transparence** : Le fondement d'une décision particulière concernant un SIA doit toujours pouvoir être découvert.

6. **Responsabilité** : La création et l'exploitation d'un SIA doivent permettre de justifier sans ambiguïté toutes les décisions prises.

7. **Sensibilisation aux mauvais usages** : Les créateurs de SIA doivent se prémunir contre tous les mauvais usages et risques potentiels des SIA en exploitation.

8. **Compétence** : Les créateurs de SIA doivent préciser les types et les niveaux de connaissances nécessaires pour comprendre et exploiter une application donnée de SIA. Les opérateurs doivent adhérer aux connaissances et aux compétences requises pour un fonctionnement sûr et efficace.

- **Déclaration de Montréal pour un développement responsable de l'intelligence artificielle**

La Déclaration de Montréal est une initiative de l'Université de Montréal. Sa rédaction est le fruit du travail d'une équipe de scientifiques pluridisciplinaires et interuniversitaires qui s'appuie sur un processus de consultation citoyenne et sur la concertation avec des experts et des parties prenantes du développement de l'IA (Rapport de la Déclaration de Montréal pour un développement responsable de l'intelligence artificielle, 2018, p. 6).

La Déclaration compte actuellement les signatures de 2386 citoyens et 200 organismes/organisations, dont 19 sont des institutions gouvernementales nationales et internationales, parmi lesquelles nous retrouvons la Commission d'accès à l'information, le ministère des Relations Internationales du Québec et la Ville de Montréal (Déclaration de Montréal IA responsable, s.d.).

Les principes et les recommandations présentes dans ce document sont les suivants :

- Dix principes : bien-être ; respect de l'autonomie ; protection de l'intimité et de la vie privée ; solidarité ; participation démocratique ; équité ; inclusion de la diversité ; prudence ; responsabilité ; développement soutenable
- Huit recommandations en vue de l'élaboration de politiques publiques : organisme indépendant de veille et de consultation citoyenne ; politique d'audit et de certification des SIA ; encapacitation et autonomisation ; formations en éthique ; favoriser un développement inclusif de l'IA ; protéger la démocratie des manipulations politiques de l'information ; développement international de l'IA ; empreinte environnementale directe et induite des SIA

- **Les principes sur l'IA de l'OCDE – Recommandation du Conseil sur l'intelligence artificielle**

Les principes de l'OCDE sur l'IA ont été adoptés, en mai 2019, par les pays membres de l'OCDE, lorsque ces derniers ont approuvé la Recommandation du Conseil sur l'intelligence artificielle (OCDE, 2019). Ces principes énoncent des recommandations concrètes pour l'élaboration de politiques publiques et de stratégies. Étant donné leur portée générale, ils peuvent s'appliquer au développement de l'IA partout dans le monde.

- Cinq principes d'une approche responsable à l'appui d'une IA digne de confiance : croissance inclusive, développement durable et bien-être ; valeurs centrées sur l'humain et équité ; transparence et explicabilité ; robustesse, sûreté et sécurité ; responsabilité
- Cinq recommandations en vue de l'élaboration de politiques nationales et de coopération internationale : investir dans la recherche et le développement en matière d'IA ; favoriser l'instauration d'un écosystème numérique pour l'IA ; façonner un cadre d'action favorable à l'IA ; renforcer les capacités humaines et préparer la transformation du marché du travail ; favoriser la coopération internationale au service d'une IA digne de confiance

- **UNESCO – Recommandation sur l'éthique de l'intelligence artificielle**

Le 24 novembre 2021, la Recommandation sur l'éthique de l'intelligence artificielle a été adoptée par la Conférence générale de l'UNESCO lors de sa 41^e session. Ce document présente les dix principes suivants : proportionnalité et innocuité ; sûreté et sécurité ; équité et non-discrimination ; durabilité ; droit au respect de la vie privée et protection des données ; surveillance et décision humaines ; transparence et explicabilité ; responsabilité et redevabilité ; sensibilisation et éducation ; gouvernance et collaboration multipartites et adaptatives (UNESCO, 2021).

En plus de ces documents, on peut également citer la *Vision commune de Charlevoix sur l'avenir de l'intelligence artificielle*, une des huit déclarations d'engagement émises par les

dirigeants du G7, le 9 juin 2018, lors de leur sommet dans Charlevoix (Québec) (Sept, G., & Canada, 2018). Les *Principes du G20 sur l'IA* qui, issus de la Recommandation de l'OCDE, ont été accueillis par les dirigeants du G20, en juin 2019, à l'occasion du Sommet d'Osaka (G20, 2019).

Plusieurs organisations publiques, privées et civiles se sont inspirées de droits fondamentaux pour élaborer des cadres éthiques pour les systèmes d'IA en y reprenant certains principes communs comme le respect de la dignité humaine, le respect de la démocratie, de la justice et de l'état de droit, l'égalité, la non-discrimination et la solidarité puis la liberté des individus et droits de citoyens (GEHN IA, 2018, p. 12).

Si chaque entité a ses propres priorités et sa propre approche, nous pouvons trouver des points communs entre les divers principes et lignes directrices existants. En effet, il existe des travaux de chercheurs qui ont analysé les différents documents existants afin d'identifier les similitudes entre eux concernant les principes considérés comme primordiaux pour le développement et l'utilisation d'une IA digne de confiance, comme nous le verrons par la suite.

3.5. Les dimensions pour les principes éthiques en IA

Jobin et ses collègues (2019) ont cartographié et analysé l'ensemble actuel de principes et de lignes directrices sur l'éthique de l'IA présents dans un certain nombre de documents produits par des auteurs de différents secteurs afin de déterminer si un consensus général global sur ces questions est en train d'émerger. Les résultats ont révélé une convergence générale autour de cinq principes éthiques (transparence, justice et équité, non-malfaisance, responsabilité et respect de la vie privée), avec des divergences substantielles quant à la manière dont ces principes sont interprétés, pourquoi ils sont considérés importants, à quel problème, domaine ou acteurs ils se rapportent et comment ils devraient être mis en œuvre. Ces résultats ont souligné l'importance d'intégrer les efforts de développement de lignes directrices avec un examen éthique substantiel et des stratégies de mise en œuvre appropriées.

Dans un autre travail, à partir d'une analyse de contenu approfondie de 36 documents considérés comme importants sur les principes de l'IA, rédigés par différents auteurs, une équipe de chercheurs du centre de recherche *Berkman Center for Internet & Society* de l'université de Harvard a identifié 8 dimensions pour les principes éthiques de l'IA, à savoir (Fjeld et al., 2020) : vie privée (*privacy*), responsabilité (*accountability*), sûreté et sécurité (*safety and security*), transparence et explicabilité (*transparency and explainability*), équité et non-discrimination (*fairness and non-discrimination*), contrôle humain de la technologie (*human control of technology*), responsabilité professionnelle (*professional responsibility*), promotion des valeurs humaines (*promotion of human values*).

Une deuxième constatation dans ce travail est que les documents plus récents tendent à couvrir l'ensemble de ces huit thèmes (Fjeld et al., 2020).

3.6. Les outils pour une IA digne de confiance

Comme nous l'avons vu précédemment, en plus de définir les principes considérés comme primordiaux pour assurer le développement et l'utilisation responsables de l'IA, plusieurs documents formulent des recommandations concernant les mesures appropriées qui peuvent être adoptées pour atténuer les enjeux qu'un système d'IA peut présenter (Déclaration de Montréal, 2018 ; OCDE, 2019). Les discussions et documents politiques autour de l'IA sont passés de l'énonciation de principes à leur mise en œuvre (OCDE, 2021).

Or, comme Dilhac et ses collègues (2021, p. 14) le signalent pertinemment,

[L]es principes ont pour vocation de guider l'action. Ils peuvent ainsi être à la base de différentes mesures concrètes de la gestion et du développement éthique de l'IA. Il peut s'agir, par exemple : de lois ; de politiques publiques ; d'audits ou de certifications ; de formations ; d'acteurs institutionnels ; de codes de conduite ; de solutions techniques.

Il existe donc plusieurs mesures (ex : outils, instruments et méthodes de travail) pour aider à mettre en œuvre les principes pour le développement et l'utilisation responsables d'un système d'IA, en augmentant les avantages et en réduisant les risques que l'IA peut présenter.

Le Réseau d'experts de l'IA à l'OCDE (ONE AI) a élaboré un cadre pour les mesures ou instruments identifiés comme pertinents pour le développement, le déploiement et l'utilisation des systèmes d'IA digne de confiance avec l'objectif de développer une base de données interactive qui fournira aux acteurs de l'IA et aux décideurs politiques des informations sur les instruments les plus récents (OCDE, 2021).

Les instruments techniques, procéduraux et éducatifs (OCDE, 2021)

Techniques	Procéduraux	Éducatifs
<p>Visent à aborder des questions spécifiques liées à l'IA sous un angle technique, notamment la détection des biais, la transparence et l'explicabilité des systèmes d'IA, la performance, la robustesse, la sûreté et la sécurité contre les attaques adverses</p>	<p>Fournissent des conseils de mise en œuvre opérationnels ou liés aux processus</p>	<p>Englobent les mécanismes de sensibilisation, d'information, de préparation ou de perfectionnement des parties prenantes impliquées ou concernées par la mise en œuvre d'un système d'IA</p>
<ul style="list-style-type: none"> ➤ Boîtes à outils / outils logiciels ➤ Documentation technique ➤ Certification technique ➤ Normes techniques ➤ Outils de développement de produits / cycle de vie ➤ Outils de validation technique 	<ul style="list-style-type: none"> ➤ Lignes directrices ➤ Cadres de gouvernance ➤ Outils de développement de produits / cycle de vie ➤ Outils de gestion des risques ➤ Codes de conduite sectoriels ➤ Conventions collectives ➤ Certification ➤ Documentation relative aux processus ➤ Normes de processus 	<ul style="list-style-type: none"> ➤ Processus de gestion du changement ➤ Renforcement des capacités / sensibilisation ➤ Conseils en matière de conception inclusive ➤ Matériel éducatif / programmes de formation

3.7. Les questions de recherche

Après l'analyse de tous ces points jugés importants pour la compréhension du mandat du stage, nous présentons les questions auxquelles ce rapport tentera de répondre en lien avec le présent mandat et, par conséquent, le besoin de l'organisation hôte.

Quelles sont les principales stratégies adoptées par d'autres pays ou organisations qui pourraient guider l'élaboration d'un cadre éthique pour le développement et l'utilisation responsables de l'IA dans l'administration publique québécoise ? Quels sont les principes le plus souvent invoqués dans leurs stratégies ainsi que les mesures les plus porteuses à mettre en œuvre pour soutenir une IA éthique et responsable ?

4. Méthodologie de recherche et d'intervention

Pour l'accomplissement de ce mandat, des rencontres initiales ont eu lieu avec le mandant et le conseiller académique, puis avec l'équipe chargée de la mise en œuvre de la stratégie IA. Lors de ces réunions, il a été établi ce que l'on attendait comme activités à réaliser et comme biens à livrer tout au long du stage.

Il a aussi été jugé pertinent et nécessaire de procéder à une analyse comparative des initiatives réalisées dans d'autres pays et organisations dans le domaine de l'IA qui pourrait inspirer le gouvernement du Québec à concevoir son propre cadre éthique pour un développement et une utilisation responsable de l'IA au sein de l'administration publique.

Une analyse comparative (étalonnage ou *benchmarking*) est un exercice qui vise à recenser, à comparer, à comprendre et à adapter les pratiques en cours dans une organisation sur la base d'observations faites ailleurs. Le but de l'analyse comparative est de trouver des exemples de résultats supérieurs et de comprendre les processus à l'origine de ces performances afin d'adapter et d'incorporer ces meilleures pratiques dans ses propres opérations (Bruno, 2014).

Le *benchmarking* est « un processus continu de recherche, d'analyse comparative, d'adaptation et d'implantation des meilleures pratiques pour améliorer la performance des processus dans une organisation » (Wikipédia, 2021).

La méthodologie adoptée pour la réalisation de ce mandat a consisté, dans un premier temps, à collecter des données à partir de l'analyse de la documentation portant sur la stratégie de l'IA¹ (.), en vue d'extraire les principaux éléments à considérer et à retenir pour promouvoir le développement d'une IA éthique et transparente.

Ensuite, une analyse comparative des plans stratégiques lancés à ce jour par d'autres pays et organisations (Tableau 1) a été réalisée, avec une attention particulière portée aux principes

¹ En particulier la Stratégie québécoise d'intégration de l'IA dans l'administration publique 2021-2026 et les mesures clés 2021-2023 ainsi que la Déclaration de Montréal et des documents publiés par l'OCDE, UNESCO, Commission Européenne etc.

éthiques pouvant intéresser le gouvernement québécois et aux mesures mises en place pour soutenir ces stratégies pour une IA éthique et responsable.

Tableau 1
Plans stratégiques analysés

Pay	Document
Danemark	<i>National Strategy for Artificial Intelligence</i>
Estonie	<i>Report of Estonia's AI Taskforce</i>
Estonie	<i>Estonia's national artificial intelligence strategy 2019-2021</i>
Finlande	<i>Leading the way into the age of artificial intelligence: Final report of Finland's Artificial Intelligence Programme 2019</i>
Finlande	<i>Finland's Age of Artificial Intelligence: Turning Finland into a leading country in the application of artificial intelligence – Objective and recommendations for measures</i>
Finlande	<i>Work in the age of artificial intelligence: Four perspectives on the economy, employment, skills and ethics</i>
Lituanie	<i>Lithuanian Artificial Intelligence Strategy: A Vision of the Future</i>
Luxembourg	Intelligence artificielle : une vision stratégique pour le Luxembourg
Malte	<i>Malta : The ultimate AI launchpad. A Strategy and Vision for Artificial Intelligence in Malta 2030</i>
Malte	<i>Malta towards trustworthy AI: Malta's ethical AI framework</i>
Norvège	<i>National Strategy for Artificial Intelligence</i>
Pays-Bas	<i>Strategic Action Plan for Artificial Intelligence</i>
Portugal	<i>AI Portugal 2030</i>
Portugal	<i>Guia para a inteligência artificial: ética, transparente e responsável</i>
Qatar	<i>National Artificial Intelligence Strategy for Qatar</i>
Serbie	<i>Strategy for the Development of Artificial Intelligence in the Republic of Serbia for the period 2020-2025</i>
Suède	<i>National approach to artificial intelligence</i>
Union Européenne	Un plan coordonné dans le domaine de l'intelligence artificielle
Union Européenne	Lignes directrices en matière d'éthique pour une IA digne de confiance

La démarche méthodologique utilisée dans cette étude d'étalonnage en est une de recherche documentaire. De cette façon, nous avons recherché des informations concernant les principes éthiques et les mesures annoncées dans des documents publics, qu'il s'agisse de documents publiés par les gouvernements ou des organismes faisant la promotion ou s'intéressant à l'IA, qui visaient à garantir une IA digne de confiance. Nous avons également

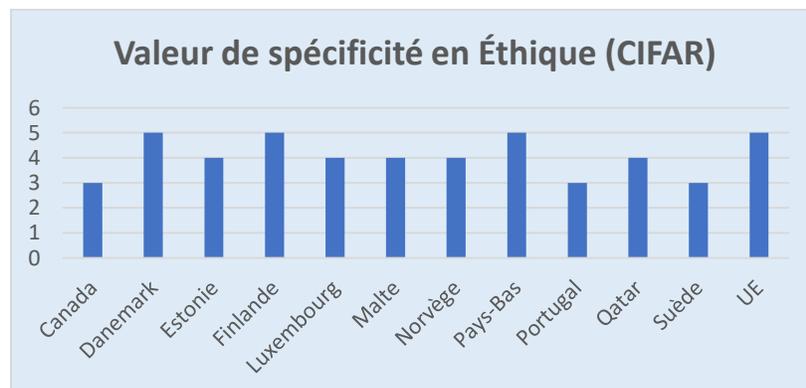
effectué des recherches dans la littérature scientifique afin de mieux comprendre les avantages et les défis éthiques que l'IA présente pour l'ensemble de la communauté.

Le choix des pays et organisations pour l'analyse comparative

Pour le choix des gouvernements et des organisations, nous nous sommes appuyés sur un rapport préparé par Johnny Kung (2020) qui a analysé un total de 28 stratégies nationales et régionales en matière d'IA. Ce rapport a produit une valeur de spécificité pour chaque domaine de politique publique en fonction du degré d'attention accordé à un domaine particulier dans les documents de stratégie énoncés :

- **4 ou 5** – pour les stratégies qui comprennent des mesures politiques détaillées et qui peuvent allouer des fonds à un domaine particulier ;
- **2 ou 3** - celles qui comprennent des objectifs politiques généraux, mais peut-être pas de mesures particulières, ou qui comprennent un plan pour développer des mesures plus détaillées ;
- **1** – celles qui reconnaissent simplement l'importance d'un domaine, mais ne précisent pas d'objectifs politiques ou de mesures concrètes.

Ainsi, en plus de l'Union européenne, nous avons sélectionné 13 pays qui ont démontré une plus grande attention à l'éthique dans leurs stratégies, de niveau 3, 4 et 5, selon Johnny Kung (2020).



La démarche adoptée pour l'analyse comparative

L'analyse des plans stratégiques choisis a été divisée en 5 blocs, comme proposé au mandant, pour un meilleur suivi de l'avancement des travaux. Chaque bloc a donc été analysé dans un délai de 1 à 2 semaines, en fonction du volume de documents à analyser. Cette division a été effectuée simplement pour une meilleure organisation et répartition du travail dans la période prévue pour le stage (4 mois), avec une présentation progressive de l'avancement de la recherche compte tenu du fait que l'organisme d'accueil organise des réunions hebdomadaires avec toute l'équipe qui la compose pour faire un suivi de la réalisation des objectifs fixés.

En annexe se trouvent les analyses qui ont été présentées au mandant tout au long du stage réalisé :

- Bloc 1 – Danemark et Union Européenne
- Bloc 2 – Finlande, Lituanie et Suède
- Bloc 3 – Estonie, Malte, Norvège et Qatar
- Bloc 4 – Luxembourg, Pays-Bas, Portugal et Serbie
- Bloc 5 – Canada

Dans un deuxième temps, une analyse comparative a été réalisée entre les entités sélectionnées, à l'exclusion de la Lituanie et de la Serbie, en raison de la difficulté à trouver des informations plus détaillées sur la mise en œuvre de leurs stratégies.

En analysant les plans stratégiques choisis, nous avons identifié les principes éthiques qui ont été pris en compte et les mesures qui ont été mises en œuvre et, à partir de ces données, nous avons construit un tableau comparatif qui sera présenté ensuite dans la section des résultats.

Il faut souligner que la Finlande et la Suède ne prévoient pas de liste de principes éthiques exprimés dans leurs stratégies. En fait, certains documents stratégiques présentent une section spécifique consacrée au domaine de l'éthique où ils annoncent expressément les principes choisis. Néanmoins, dans les plans stratégiques qui ne comportent pas de section

spécifique sur ce sujet, il est possible de trouver des valeurs/principes éparpillés dans d'autres sections du document, qui ont également été pris en compte pour le présent travail.

En outre, quelques principes ont une présence moins forte dans certains plans stratégiques gouvernementaux analysés, tel que l'utilité et le développement, ce qui ne signifie pas qu'ils sont moins importants que d'autres. Toutefois, comme ils font l'objet d'un consensus moindre, ces principes ne seront pas examinés dans le présent document, bien qu'ils soient présents dans le tableau comparatif.

Outre la difficulté de trouver des informations plus précises sur la mise en œuvre de certaines stratégies et mesures nationales, ce qui limite l'analyse documentaire, il faut également noter la limitation linguistique de la stagiaire (portugais, français et anglais) qui peut expliquer les lacunes présentes dans ce rapport.

Enfin, nous tenons à préciser que le stage a été réalisé entièrement en télétravail, en raison des directives du gouvernement du Québec compte tenu de la situation épidémiologique que nous traversons actuellement.

5. Résultats

L'analyse comparative demandée nous a permis d'identifier les principes éthiques qui ont été pris en considération pour le développement et l'utilisation responsables de l'IA dans chaque stratégie analysée ainsi que les mesures choisies pour la mise en œuvre des stratégies dans le but de garantir une IA digne de confiance. Nous présentons d'abord les constatations pour l'Union européenne, puis les résultats pour les pays sélectionnés.

5.1. Union européenne

Concernant la Stratégie de l'Union européenne, nous avons les éléments suivants :

	Union européenne 
Nom stratégie et date	Plan coordonné sur l'intelligence artificielle (2018)
Forme de l'encadrement, liste de principes et méthodes pour parvenir à une IA digne de confiance (mutuellement complémentaires ou alternatives)	<p>Les lignes directrices en matière d'éthique pour une IA digne de confiance (GEHN IA - CE)</p> <p>Quatre principes éthiques : respect de l'autonomie humaine ; prévention de toute atteinte ; équité ; explicabilité</p> <p>Sept éléments essentiels pour sa mise en œuvre : action humaine et contrôle humain ; robustesse technique et sécurité ; respect de la vie privée et gouvernance des données ; transparence ; diversité, non-discrimination et équité ; bien-être sociétal et environnemental ; responsabilité</p> <p>Méthodes techniques : l'architecture du système d'IA, éthique et état de droit dès la conception, méthodes d'explication, essais et validations, qualité des indicateurs de service</p> <p>Méthodes non techniques : réglementation, codes de conduite, normalisation, certification, cadres de gouvernance internes et externes, éducation et sensibilisation, participations des parties prenantes et dialogue social, diversité et équipes de conception inclusives</p>

Un groupe d'experts de haut niveau en IA a été désigné par la Commission européenne en juin 2018 pour préparer des lignes directrices en matière d'éthique pour une IA digne de confiance.

Le document élaboré présente 4 principes éthiques et 7 exigences que les systèmes d'IA devraient respecter tout au long de leur cycle de vie. Le document présente également des méthodes techniques et non techniques pour parvenir à une IA digne de confiance Commission européenne (2019, avril).

En 20 octobre 2020, le Parlement européen (2020) a adopté une résolution contenant des recommandations à la Commission concernant un cadre pour les aspects éthiques de l'intelligence artificielle, de la robotique et des technologies connexes (2020/2012(INL)) :

- Certificat européen de conformité éthique à être délivré par un organisme national de surveillance compétent (haut risque) ;
- Organismes de surveillance (contrôler l'application du règlement et réaliser les évaluations des risques et de la conformité et la certification).

Les États membres de l'Union européenne appuient généralement leur cadre éthique sur des directives élaborées par le groupe d'experts de haut niveau sur l'IA désigné par la Commission européenne.

5.2. Les résultats des stratégies des pays choisis

Quant aux pays sélectionnés, nous avons dressé un tableau comparatif des principes éthiques pris en compte par chacun d'entre eux et des mesures mises en œuvre pour assurer l'éthique de l'intelligence artificielle, que nous présentons ci-dessous plus en détail.

	Canada - 3 	Danemark - 5 	Estonie - 4 	Finlande - 5 	Luxembourg - 4 	Malte - 4 	Norvège - 4 	Pays-Bas - 5 	Portugal - 3 	Qatar - 4 	Suède - 3 
Nom stratégies et date	Stratégie pancanadienne en matière d'IA (2017)	Stratégie nationale sur l'intelligence artificielle (2019)	Stratégie nationale sur l'intelligence artificielle 2019-2021 (2019)	Ouvrir la voie à l'ère de l'intelligence artificielle 2019-2025 (2019)	Intelligence artificielle : une vision stratégique pour le Luxembourg (2019)	Malte : plateforme de lancement de l'intelligence artificielle — Stratégie et vision de l'intelligence artificielle à Malte 2030 (2019)	La stratégie nationale pour l'intelligence artificielle (2020)	Plan d'action stratégique pour l'intelligence artificielle (2019)	IA Portugal 2030 (2019)	Plan d'action : stratégie nationale du Qatar sur l'intelligence artificielle (2019)	Approche nationale de l'intelligence artificielle (2018)
Liste de principes	Egalité, transparence, responsabilité, légalité et équité procédurale	Autodétermination, dignité, responsabilité, explicabilité, égalité et justice, développement	Utilité, non-nuisance, autonomie, justice et clarté (transparence)	La Finlande ne prévoit pas de liste de principes, encourageant le développement de l'autorégulation - les questions éthiques doivent être résolues au cas par cas	Les lignes directrices en matière d'éthique pour une IA digne de confiance (GEHN IA - CE)	Quatre principes éthiques IA : autonomie humaine, prévention des dommages, équité et explicabilité Sept éléments essentiels , non exhaustives, pour sa mise en œuvre : agence humaine, vie privée et gouvernance des données, explicabilité et transparence, bien-être, responsabilité, équité et impartialité, performance et sécurité	Les lignes directrices en matière d'éthique pour une IA digne de confiance (GEHN IA - CE)	Les lignes directrices en matière d'éthique pour une IA digne de confiance (GEHN IA - CE)	Responsabilité, transparence, explicabilité et justice	Équité, responsabilité et transparence	La stratégie suédoise ne fournit pas de liste de principes, mais souligne que les applications de l'IA doivent être éthiques, sûres, sécurisées, fiables, transparentes et responsables, et que les systèmes d'IA doivent être soigneusement conçus pour éviter qu'ils ne causent des dommages
Forme de l'encadrement	<ul style="list-style-type: none"> • Conseil consultatif en matière d'IA • Directive sur la prise de décisions automatisée 	<ul style="list-style-type: none"> • Le conseil d'éthique des données • Des méthodes et des lignes directrices communes au secteur public (utilisation transparente des algorithmes) seront développées 	Une proposition de principes d'utilisation responsable des données sera rédigée	Comité d'éthique de l'IA	Le comité interministériel AI4Gov (composé de représentants du ministère de la Digitalisation, du Service des médias et des communications (SMC), et du Service information et presse (SIP))	<ul style="list-style-type: none"> • Un cadre éthique de l'IA (Malta Towards Trustworthy AI) • Un comité national d'éthique technologique sera créé 	Un organe consultatif sera créé	Lignes directrices pour l'application de l'analyse algorithmique des données	Un comité d'éthique sera créé	<ul style="list-style-type: none"> • Un comité d'intelligence artificielle sera créé • Des lignes directrices seront introduites 	<ul style="list-style-type: none"> • Centre de durabilité de l'IA (AI sustainability centre) • Le Comité pour l'innovation technologique et l'éthique (Komet) • Des règles, des standards, des normes et des principes éthiques seront développés
Initiatives et outils en place	<ul style="list-style-type: none"> • Livre blanc « L'intelligence artificielle responsable au sein du gouvernement du Canada » • Outil d'évaluation de l'incidence algorithmique • Groupe de travail sur la sensibilisation du public – Ateliers virtuels 	<ul style="list-style-type: none"> • Une boîte à outils d'éthique des données (lignes directrices, études de cas et outils d'évaluation d'impact) • Guide - Ethique des Données – Comment le faire • Un label commun de cybersécurité et d'éthique des données (un système d'étiquetage) sera lancé en 2021 	Élaboration d'un questionnaire d'auto-évaluation (développeurs d'IA)	<ul style="list-style-type: none"> • Registre d'IA d'Helsinki (registres d'IA ouverts – permettent de suivre la manière dont les algorithmes sont utilisés) • Livre blanc sur le registre de l'IA • Cours en ligne sur l'éthique de l'IA • Un défi éthique de l'IA (pour inciter les entreprises à contribuer à la création de principes éthiques pour l'IA) • Projet de recherche <i>Ethical AI for the Governance of the Society</i> (ETAİROS) 	- Groupe de travail interministériel (évaluation préliminaire des risques juridiques et éthiques)	Programme de certification en IA	Un bac à sable (sandbox) réglementaire pour une IA responsable	<ul style="list-style-type: none"> • Registre d'IA d'Amsterdam (registres d'IA ouverts – permettent de suivre la manière dont les algorithmes sont utilisés) • Livre blanc sur le registre de l'IA • Un guide étape par étape (<i>Non-discrimination by design</i>) 	<ul style="list-style-type: none"> • Un Guide de l'IA dans l'administration publique - GuA (en version bêta) • Un outil d'évaluation des risques 		

Principes éthiques

En ce qui concerne les principes éthiques, les résultats de la présente recherche sont similaires aux constatations faites par Jobin et ses collègues (2019) dans la recherche qu'ils ont menée, c'est-à-dire qu'il est possible de constater une convergence générale autour des principes suivants : transparence et explicabilité ; responsabilité ; respect de la vie privée ; justice et équité (non-discrimination) ; sûreté et sécurité (non-nuisance).

Nous passons donc à l'analyse de chacun d'entre eux.

Transparence et explicabilité

Les principes de transparence et explicabilité exigent que les systèmes d'IA soient conçus et mis en œuvre de manière à permettre une surveillance par les parties prenantes, notamment en traduisant leurs opérations en résultats intelligibles et en fournissant des informations sur le lieu, le moment et la manière dont ils sont utilisés (Fjeld et al., 2020, p. 5). L'objectif de ces principes est « de préserver la confiance du public dans les systèmes d'IA et de fournir suffisamment d'informations pour aider à garantir la responsabilité réelle des développeurs, des sociétés déployant un système d'IA et des utilisateurs de ce système, et pour démontrer si les décisions prises par un système d'IA sont justes et impartiales » (Thomson Reuters, 2019). Les principes de transparence et d'explicabilité sont présents dans 100% des documents analysés.

Selon la Norvège, les décisions prises par les systèmes basés sur l'IA doivent être traçables, explicables et transparentes. Cela signifie que les personnes physiques ou morales doivent avoir la possibilité de comprendre comment une décision qui les affecte a été prise (2020, p. 59).

Dans le même sens, la Suède souligne que l'utilisation des algorithmes d'IA doit être transparente et compréhensible (2018, p. 8).

Responsabilité

Le principe de la responsabilité renvoie à l'importance des mécanismes visant à garantir que la responsabilité des impacts des systèmes d'IA est répartie de manière appropriée et que des recours adéquats sont prévus (Fjeld et al., 2020, p. 5). Parmi les douze stratégies analysées dans ce rapport, onze font référence à la responsabilité, soit 91,66% des documents analysés.

Compte tenu de l'importance de ce principe, le Danemark déclare qu'il œuvre activement pour que la responsabilité soit un principe directeur dans l'utilisation et le développement de l'IA (2019, p. 8).

Le Qatar rappelle que les méthodes d'IA tendent à acquérir des caractéristiques de « boîte noire » et ne se prêtent pas toujours aux principes d'équité, de responsabilité et de transparence qui sont essentiels à la viabilité à long terme des sociétés et des pays (2019, p. 15).

Respect de la vie privée

Selon le principe de respect de la vie privée, les systèmes d'IA doivent respecter la vie privée des individus, dans l'utilisation des données pour le développement de systèmes technologiques. Ces systèmes doivent aussi donner aux personnes concernées la possibilité d'agir sur leurs données et les décisions qui ont été prises sur la base de l'utilisation de ces données (Fjeld et al., 2020, p. 5). Le principe de la vie privée est présent dans 91,66% des documents.

Les Pays-Bas rappellent que les questions relatives à la vie privée touchent des domaines tels que la technologie de reconnaissance faciale, le big data et les techniques par lesquelles les données personnelles sont dérivées d'autres données (2019, p. 41).

Justice et équité (non-discrimination)

Ces principes exigent que les systèmes d'IA soient conçus et utilisés de manière à maximiser l'équité et à promouvoir l'inclusion (Fjeld et al., 2020, p. 5). « Cela signifie d'aborder les

algorithmes et les biais de données dès le début afin de garantir l'équité et la non-discrimination » (Thomson Reuters, 2019). Les principes de justice et d'équité sont présents dans 91,66% des documents.

À ce propos, le Danemark précise que l'IA ne doit pas reproduire ou, pire encore, aggraver les préjugés qui marginalisent certains groupes dans la population (2019, p. 28) tandis que Malte rappelle que le développement, le déploiement, l'utilisation et l'exploitation des systèmes d'IA doivent être équitables (2019b, p. 12).

Sûreté et sécurité (non-nuisance)

Ces principes exigent que les systèmes d'IA soient sûrs, c'est-à-dire qu'ils fonctionnent comme prévu, et qu'ils soient sécurisés, c'est-à-dire qu'ils ne puissent pas être compromis par des parties non autorisées (Fjeld et al., 2020, p. 5). Les principes de sûreté et de sécurité sont présents dans 83,33% des documents.

Le Danemark déclare que la sécurité est cruciale pour garantir le développement et l'utilisation responsables de l'IA. Il souligne que les particuliers, les entreprises et les employés doivent avoir la certitude que les données et les algorithmes ont été protégés contre les manipulations et les attaques, et que les résultats produits par l'IA sont fiables (2019, p. 26).

À son tour, Malte souligne que les systèmes d'IA ne doivent pas causer de dommages aux humains, à l'environnement naturel ou aux autres êtres vivants (2019b, p. 12).

Les mesures mises en œuvre pour assurer une IA éthique et responsable

Quant aux mesures de nature éthique mises de l'avant dans les stratégies gouvernementales en IA, nous nous limitons ici à en souligner les principaux éléments. Des informations plus détaillées sur chaque stratégie analysée sont disponibles dans les annexes de ce rapport.

L'analyse que nous avons réalisée nous a donc permis de faire un ensemble de constats.

Les actions nationales visant à répondre aux préoccupations éthiques diffèrent selon les pays en termes d'approche stratégique et de niveau d'attention en fonction des besoins et des particularités de chacun.

Comité/groupe de travail

Pour faciliter l'élaboration de lignes directrices éthiques, de nombreux gouvernements ont créé des comités et des conseils d'éthique de l'IA (Canada, Danemark, Finlande, Luxembourg et Suède). Ces organismes fournissent des recommandations sur les questions éthiques, surveillent en permanence l'utilisation et le développement des technologies d'IA et sensibilisent les chercheurs ainsi que les praticiens et gestionnaires des secteurs public et privé à une conduite éthique en IA par la diffusion des bonnes pratiques, par des formations et par des codes de conduite éthiques (OECD, 2021).

Au Canada, le Conseil consultatif a mis sur pied le « Groupe de travail sur la sensibilisation du public ». Ce groupe étudie des mécanismes visant à renforcer la sensibilisation du public et à favoriser la confiance dans l'IA. Il vise également à alimenter le débat public sur une compréhension mesurée de la technologie de l'IA, de ses utilisations potentielles et des risques qui y sont associés (Canada, s.d-a). En partenariat avec l'Institut canadien des recherches avancées (ICRA) et l'Algora Lab (Université de Montréal), ce groupe a dirigé des ateliers virtuels² dans l'ensemble du Canada pour entamer des discussions avec les membres du public au sujet de leurs perceptions de l'IA (Canada, s.d.-d).

Normes/guides/boîte à outils

Le Danemark se concentre spécifiquement sur l'utilisation responsable et durable des données par les secteurs public et privé pour assurer une IA éthique. En décembre 2019, une boîte à outils a été lancée pour aider les entreprises à adopter et à mettre en œuvre l'éthique des données dans leurs modèles économiques (Van Roy et al., 2021). La boîte à outils contient des lignes directrices sur la manière d'inclure l'éthique des données dans le code de

² « Dialogue ouvert : L'intelligence artificielle au Canada ».

pratique de leur entreprise, des études de cas sur les meilleures pratiques et des outils d'évaluation d'impact sur l'utilisation responsable des algorithmes (Larsen, 2020).

Le Danemark a élaboré aussi un outil pratique qui comprend un guide en 5 étapes qui vise à expliquer aux parties prenantes comment commencer à travailler avec l'éthique des données (*Dataetisk Råd*, 2021).

Aux Pays-Bas, une équipe de recherche a élaboré un guide sur la manière dont les organisations peuvent, étape par étape, éviter la discrimination lorsqu'elles utilisent des algorithmes (Vrije Universiteit Brussel, 2021).

Certification/sceaux/registres

Certains gouvernements mettent également en place des systèmes de contrôle et de récompense afin de faire respecter les principes de l'IA. Malte a mis au point un cadre de certification de l'IA, délivré par la *Malta Digital Innovation Authority* (MDIA) (OECD.AI, 2021). Une certification révèle qu'un système d'IA a été développé de manière éthique, transparente et socialement responsable. Ainsi, « des processus de certification par un tiers de confiance peuvent contribuer à articuler la relation entre concepteur, producteur et acquéreur » (Linden, 2020, p. 23).

Des sceaux similaires à la certification ont été considérés dans d'autres pays, comme le Danemark (Van Roy et al., 2021). L'apposition d'un sceau indiquera aux consommateurs que l'entreprise traite l'IA et ses données d'une manière digne de confiance, éthique et sûre (Larsen, 2020).

Dans la même veine, les registres d'IA créés par les villes d'Amsterdam (Pays-Bas) et d'Helsinki (Finlande) visent à garantir une utilisation sûre, responsable et transparente des algorithmes d'IA (Van Roy et al., 2021). Selon la Ville de Helsinki (s.d.),

Grâce au registre, vous pouvez vous familiariser avec les aperçus rapides des systèmes d'intelligence artificielle de la ville ou examiner leurs informations plus détaillées en fonction de vos propres intérêts. Vous pouvez également donner votre avis et participer ainsi à la construction d'une IA centrée sur l'homme.

Évaluation des risques/environnement d'expérimentation/bac à sable

La Suède a créé un centre pour la durabilité de l'IA, cofondé par des entreprises, des universités et des autorités publiques (*AI Sustainability Centre*). L'objectif de ce centre est de fournir des services de dépistage, d'évaluation, d'atténuation et d'audit ainsi que des rapports sur l'exposition d'une organisation aux risques éthiques liés à l'IA (Van Roy et al., 2021).

Certains gouvernements, comme le Danemark, la Norvège et le Portugal, envisagent également de créer des environnements contrôlés pour l'expérimentation de l'IA, par exemple en développant des outils d'évaluation des risques et des bacs à sable réglementaires (Van Roy et al., 2021). Le bac à sable est une plate-forme de test où les applications peuvent être modifiées sans interférer avec l'environnement de production. Les développeurs peuvent y effectuer toutes les opérations de modification expérimentale qui garantiront le bon fonctionnement de la solution, en évitant les dommages qui peuvent nuire au système des entreprises ou des institutions ou organisations publiques.

Le Portugal mentionne que l'initiative de développer un outil d'évaluation des risques, en plus d'un guide, constitue un moyen d'adopter des références communes pour la mise en œuvre d'une IA éthique, transparente et responsable pour le secteur public (Portugal, s.d.).

La Norvège explique qu'il existe une certaine incertitude quant au rôle de l'éthique dans l'IA et à la manière de l'interpréter dans la pratique, c'est-à-dire lorsqu'on utilise un système d'IA. Les participants ont suggéré de créer une bibliothèque de dilemmes éthiques ou une liste de contrôle éthique dans le cadre du bac à sable. L'objectif est de promouvoir le développement de solutions d'IA innovantes qui, du point de vue de la protection des données, sont à la fois éthiques et responsables (Datatilsynet, 2021).

La mise en place de bacs à sable et de bancs d'essai pour l'IA permet de tester les technologies dans un environnement réel, d'examiner l'environnement réglementaire et de procéder à des ajustements. Ces instruments permettent donc d'identifier et d'atténuer les risques.

D'autres gouvernements fournissent un questionnaire d'auto-évaluation. Le Canada, par exemple, offre l'outil d'évaluation de l'incidence algorithmique (ÉIA) pour aider les décideurs et autres responsables à déterminer comment sélectionner et utiliser de manière responsable l'IA au sein du gouvernement, en identifiant les biais éthiques des systèmes d'IA et en aidant à les atténuer (OECD.AI, 2021 ; Canada, s.d.-c). Le gouvernement estonien travaille actuellement à l'élaboration d'un questionnaire d'auto-évaluation destiné aux développeurs d'IA (Van Roy et al., 2021).

Défi éthique

La Finlande a mis en place un défi éthique de l'IA sur le site du programme d'IA pour inciter les entreprises à contribuer à la création de principes éthiques pour l'IA (Van Roy et al., 2021).

Formation

Enfin, avec une approche éducative, la Finlande a mis en place un cours en ligne sur l'éthique de l'IA, offert par l'université d'Helsinki. Conçu en partenariat avec le ministère finlandais des finances et les villes d'Helsinki, d'Amsterdam et de Londres, ce cours aide les administrations publiques, les entreprises et le grand public à comprendre ce que signifie l'utilisation éthique de l'IA (Van Roy et al., 2021).

6. Recommandations

La première recommandation proposée dans ce rapport est la création d'un comité d'experts en éthique de l'IA.

La présence d'un comité d'éthique peut favoriser la recherche, l'acquisition de connaissances et le dialogue entre les secteurs concernés, contribuer à l'établissement des normes ou des règlements sur le développement et l'utilisation responsable d'un système de l'IA et éclairer la décision publique. Il peut fournir des recommandations et renforcer la sensibilisation aux dilemmes éthiques, en partie par le moyen d'un débat public.

La mesure 9.3 de la stratégie IA (Québec, 2021b) prévoit la création d'un comité-conseil réunissant des représentants de l'écosystème québécois en IA et des représentants de l'administration publique pour fournir des orientations à prendre et les actions à mettre de l'avant pour toute la durée de la mise en œuvre de la Stratégie. De ce fait, la présence d'experts en éthique et autres experts en lien avec l'IA au sein de ce comité peut contribuer à la réalisation de l'objectif de la Stratégie IA.

La deuxième recommandation consiste à développer et mettre à disposition un instrument d'évaluation des risques éthiques pour un système d'IA, tel qu'un questionnaire comme celui mis à disposition par le gouvernement fédéral ou une grille d'auto-évaluation. Un outil d'évaluation des risques peut être utilisé à n'importe quel stade du cycle du projet et peut présenter plusieurs avantages, tels qu'éliminer l'effet boîte noire, clarifier la manière d'intégrer l'IA dans les organismes et organisations publics, réduire les biais, protéger les personnes vulnérables, garantir la non-discrimination, identifier les risques et les impacts, surveiller les résultats et améliorer continuellement la performance des systèmes. Il peut aussi contribuer à l'adoption de références communes pour la mise en œuvre d'une IA éthique, transparente et responsable par le secteur public.

Enfin, il est recommandé de mettre en place un programme de formation continue sur l'importance de l'éthique dans le développement et l'utilisation de l'IA pour les fonctionnaires et pour autres parties prenantes. En effet, un programme de formation est un mécanisme

important pour sensibiliser, informer, préparer ou améliorer les parties prenantes impliquées ou affectées par la mise en œuvre d'un système d'IA de manière responsable. L'acquisition de connaissances appropriées, non seulement dans le domaine technique, mais aussi dans le domaine éthique, peut sensibiliser et accroître la vigilance des parties prenantes à l'égard d'éventuelles conséquences éthiques et sociales indésirables dans toutes les phases d'un système d'IA.

Le développement d'un programme de formation pour le personnel de l'État est l'une des mesures envisagées dans l'incubateur de mesures de la Stratégie IA (Québec, 2021b). Plus précisément, des principes éthiques doivent être respectés depuis la conception et l'expérimentation d'un système d'IA jusqu'à son utilisation, d'où l'importance de considérer cette mesure dès le début de la mise en œuvre de la Stratégie IA.

La réalisation d'un programme de formation sur l'IA éthique peut se faire en partenariat avec l'Académie de Transformation Numérique ainsi que l'École nationale d'administration publique, institutions qui disposent d'experts compétents dans ce domaine.

Comme autres options de formation, il convient de développer un matériel éducatif, comme des vidéos tutoriels ou une boîte à outils, qui clarifient, guident et sensibilisent les parties prenantes sur le développement et l'utilisation d'une IA éthique et digne de confiance.

7. Conclusion

Le présent travail a été réalisé pour présenter les résultats obtenus à partir de l'analyse comparative des stratégies et politiques mises en place pour le développement et l'utilisation de l'IA responsables par d'autres gouvernements et organisations afin d'identifier les principes et les mesures qui pourraient inspirer l'élaboration d'un cadre éthique dans l'administration publique québécoise, tel que prévu dans la Stratégie d'intégration de l'intelligence artificielle dans l'administration publique 2021-2026.

La recherche que nous avons menée, basée sur l'analyse de plans stratégiques de l'Union européenne et de 13 pays, ainsi que d'autres documents relatifs à l'IA, nous a permis d'identifier les principes éthiques qui ont été pris en considération pour le développement et l'utilisation responsables de l'IA dans chaque stratégie analysée, ainsi que les mesures choisies pour la mise en œuvre des stratégies dans le but de garantir une IA digne de confiance.

L'IA a le pouvoir de transformer positivement la société et l'environnement, et pour exploiter ce pouvoir à son plein potentiel, l'administration publique doit s'attacher à relever les défis éthiques.

L'élaboration de principes ou lignes directrices et la mise en place d'instruments de conseil, de suivi, d'encadrement et de capacitation des individus, ainsi que d'outils d'évaluation de risques éthiques, comme cela se fait dans d'autres pays, peuvent être des leviers facilitant le développement et l'utilisation responsables de l'IA dans l'administration publique québécoise.

Nous avons présenté ici trois recommandations qui peuvent être envisagées pour guider le début de la mise en œuvre de la stratégie du Québec. Cependant, d'autres mesures également pertinentes peuvent être mises en œuvre simultanément ou progressivement pour renforcer l'objectif de la Stratégie québécoise d'assurer une IA digne de confiance, comme la certification, le bac à sable et la boîte à outils mis en œuvre dans d'autres administrations publiques comme nous l'avons vu dans les résultats présentés par la présente recherche.

8. Bibliographie

- Abid, R. (2021). Pour une éthique de l'intelligence artificielle. *Gestion*, 46(2), 100-104. doi:10.3917/riges.462.0100
- Bruno, I. (2014). Benchmarking, in *Encyclopedia of Quality of Life and Well-Being Research*, Springer Netherlands, p. 363-368.
- Canada. (s.d.-a). Conseil consultatif en matière d'intelligence artificielle. <https://ised-isde.canada.ca/site/conseil-consultatif-intelligence-artificielle/fr>
- Canada. (s.d.-b). Directive sur la prise de décisions automatisée. <https://www.tbs-sct.gc.ca/pol/doc-fra.aspx?id=32592#cha6>
- Canada. (s.d.-c). Outil d'évaluation de l'incidence algorithmique. <https://www.canada.ca/fr/gouvernement/systeme/gouvernement-numerique/innovations-gouvernementales-numeriques/utilisation-responsable-ai/evaluation-incidence-algorithmique.html>
- Canada. (s.d.-d). Groupe de travail sur la sensibilisation du public. <https://ised-isde.canada.ca/site/conseil-consultatif-intelligence-artificielle/fr/groupe-travail-sensibilisation-public>
- Canada. (s.d.-e). Utilisation responsable de l'intelligence artificielle (IA). Nos principes directeurs. <https://www.canada.ca/fr/gouvernement/systeme/gouvernement-numerique/innovations-gouvernementales-numeriques/utilisation-responsable-ai.html#toc1>
- Chauhan, L. (2019, 7 février). *QCAI releases National AI Strategy for Qatar*. Qatar Center for Artificial Intelligence. <https://qcai-blog.qcri.org/index.php/2019/02/07/qcai-releases-blueprint-of-national-ai-strategy-for-qatar/>
- CIFAR. (s.d.). Stratégie pancanadienne en matière d'IA. <https://cifar.ca/fr/ia/>
- CIFAR. (2020, octobre). *Pan-canadian AI Strategy Impact Assessment Report*. <https://cifar.ca/wp-content/uploads/2020/11/Pan-Canadian-AI-Strategy-Impact-Assessment-Report.pdf>
- Commission européenne. (2018, 7 décembre). *Communication de la Commission au Parlement européen, au Conseil européen, au Conseil, au Comité économique et social européen et au Comité des régions – Plan coordonné dans le domaine de l'intelligence artificielle* (COM(2018) 795 final). https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56032 et https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56034
- Commission européenne. (2019, avril). *Les lignes directrices en matière d'éthique pour une IA digne de confiance*. Groupe d'experts de haut niveau sur l'intelligence artificielle. Bruxelles. <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-fr/format-PDF>

- Commission européenne. (2021a). *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*. COM(2021) 206 final. 2021, 21 avril. <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:52021PC0206&from=EN>
- Commission européenne. (2021b). *Coordinated Plan on Artificial Intelligence 2021 Review*. 2021, 21 avril. <https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review>
- Danemark. (2019, mars). *National Strategy for Artificial Intelligence*. Agency for Digitisation. Ministry of Finance and Ministry of Industry, Business and Financial Affairs. <https://en.digst.dk/policy-and-strategy/denmark-s-national-strategy-for-artificial-intelligence/>
- Dataetisk Råd. (2021, 11 octobre). *Nu lanceres 'Dataetik – sådan gør du', et praktisk værktøj til når du skal træffe dataetiske beslutninger*. <https://dataetiskraad.dk/nu-lanceres-dataetik-saadan-goer-du-et-praktisk-vaerktoej-til-naar-du-skal-traeffe-dataetiske>
- Datatilsynet. (s.d.). *Sandbox for responsible artificial intelligence*. Norvège. <https://www.datatilsynet.no/en/regulations-and-tools/sandbox-for-artificial-intelligence/>
- Datatilsynet. (2021, 13 janvier). *How should we structure the sandbox?* Norvège. <https://www.datatilsynet.no/en/regulations-and-tools/sandbox-for-artificial-intelligence/how-should-we-structure-the-sandbox/>
- Déclaration de Montréal IA responsable. (s.d.). Signataires. <https://www.declarationmontreal-iaresponsable.com/signataires> (consulté le 31 janvier 2022)
- Déclaration de Montréal IA Responsable. (2018). *Rapport de la Déclaration de Montréal pour un développement responsable de l'intelligence artificielle*. https://www.declarationmontreal-iaresponsable.com/files/ugd/ebc3a3_d806f109c4104c91a2e719a7bef77ce6.pdf
- Dickson, B. (2019, 24 juillet). *How Qatar plans to become a global leader in artificial intelligence*. TechTalks. <https://bdtechtalks.com/2019/07/24/qatar-artificial-intelligence-strategy/>
- Dilhac, M.A. et al. (2021, 31 mars). *Penser l'intelligence artificielle responsable : un guide de délibération*. Algora, Université de Montréal, CIFAR et OBVIA. <https://www.docdroid.com/aeZGvtw/penser-lintelligence-artificielle-responsable-un-guide-de-deliberation-pdf>
- Éditeur officiel du Québec. (2021, 3 décembre). *Projet de loi n° 6 : loi édictant la Loi sur le ministère de la Cybersécurité et du Numérique et modifiant d'autres dispositions*.

<http://www2.publicationsduquebec.gouv.qc.ca/dynamicSearch/telecharge.php?type=5&file=2021C33F.PDF>

- Estonie. (2019a, mai). *Report of Estonia's AI Taskforce*. Government Office and Ministry of Economic Affairs and Communications. https://f98cc689-5814-47ec-86b3-db505a7c3978.filesusr.com/ugd/7df26f_486454c9f32340b28206e140350159cf.pdf
- Estonie. (2019b, juillet). *Estonia's national artificial intelligence strategy 2019-2021*. Government of the Republic of Estonia. https://f98cc689-5814-47ec-86b3-db505a7c3978.filesusr.com/ugd/7df26f_27a618cb80a648c38be427194affa2f3.pdf
- Finlande. (s.d.-a). *Ethics challenge*. AI Finland. Ministry of Economic Affairs and Employment of Finland. <https://www.tekoalyaika.fi/en/background/ethics/>
- Finlande. (s.d.-b). *Implementation of the national AuroraAI programme*. Ministry of Finance. <https://vm.fi/en/auroraai-en>
- Finlande. (2017). *Finland's Age of Artificial Intelligence. Turning Finland into a leading country in the application of artificial intelligence*. Ministry of Economic Affairs and Employment. https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf?sequence=1&isAllowed=y
- Finlande. (2018, 10 septembre). *Work in the age of artificial intelligence. Four perspectives on the economy, employment, skills and ethics*. Ministry of Economic Affairs and Employment. http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160980/TEMjul_21_2018_Work_in_the_age.pdf
- Finlande. (2018, 5 décembre). *Government report on information policy and artificial intelligence: Ethical information policy in the age of artificial intelligence*. Ministry of Finance. https://vm.fi/documents/10623/7768305/VM_Tiepo_selonteko_070219_ENG_WEB.pdf/89b99a8e-01a3-91e3-6ada-38056451ad3f/VM_Tiepo_selonteko_070219_ENG_WEB.pdf.pdf/VM_Tiepo_selonteko_070219_ENG_WEB.pdf?t=1560838655000
- Finlande. (2019, 12 juin). *Leading the way into the age of artificial intelligence. Final report of Finland's Artificial Intelligence Programme 2019*. Ministry of Economic Affairs and Employment. http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161688/41_19_Leading%20the%20way%20into%20the%20age%20of%20artificial%20intelligence.pdf
- Finlande. (2020, 13 novembre). *Artificial Intelligence 4.0 programme to speed up digitalization of business*. Ministry of Economic Affairs and Employment. <https://tem.fi/en/-/artificial-intelligence-4.0-programme-to-speed-up-digitalisation-of-business>

- Fjeld, J. et al. (2020, 15 janvier). *Principled artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI*. Berkman Klein Center for Internet & Society at Harvard University. Research Publication N° 2020-1. <https://cyber.harvard.edu/publication/2020/principled-ai>
- GlobalData Technology. (2021, 12 mars). *Qatar forms artificial intelligence committee*. Verdict. <https://www.verdict.co.uk/qatar-artificial-intelligence-committee/>
- G20. (2019). *G20 AI Principles*. <https://www.g20-insights.org/wp-content/uploads/2019/07/G20-Japan-AI-Principles.pdf>
- Helsinki (s.d.). *What is AI Register?* <https://ai.hel.fi/en/ai-register/>
- Hukoomi (2021, 23 novembre). *Qatar's National Artificial Intelligence Strategy*. Qatar e-Government. <https://hukoomi.gov.qa/en/article/qatars-national-artificial-intelligence-strategy>
- IEEE. (2017). *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems*. Institute of Electrical and Electronics Engineers (IEEE). Version 1. https://standards.ieee.org/wp-content/uploads/import/documents/other/eadi1e.pdf?utm_medium=undefined&utm_source=undefined&utm_campaign=undefined&utm_content=undefined&utm_term=undefined
- Jobin, A., Ienca, M. & Vayena, E. (2019, 2 septembre). The global landscape of AI ethics guidelines. *Nat Mach Intell* 1, 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Komet. (s.d.). *About us*. Kommittén för teknologisk innovation & etik (en ligne). Consulté en 3/11/2021. <https://www.kometinfo.se/in-english/about-us/>
- Kung, J. (2020, mai). *L'ère de l'IA : rapport sur les stratégies nationales et régionales en matière d'IA*. CIFAR. Deuxième Édition. <https://cifar.ca/wp-content/uploads/2020/11/l-ere-de-l-ia-deuxieme-edition-f.pdf>
- Larsen, F. (2020, 15 juin). *Denmark: an independent council and a labelling scheme to promote the ethical use of data*. OECD.AI. <https://oecd.ai/en/wonk/an-independent-council-and-seal-of-approval-among-denmarks-measures-to-promote-the-ethical-use-of-data>
- Le Pollotec, Y. (2018). Intelligence artificielle : le rapport de Cédric Villani. *La Pensée*, 396(4), 73-85. doi:10.3917/lp.396.0073
- Linden, I. (2020). Entre rêves et illusions... L'intelligence artificielle en question. [Between dreams and illusions... the Artificial Intelligence in question]. *Revue d'éthique et de théologie morale*, 307(3), 11-27. doi:10.3917/retm.310.0011
- Lituanie. (2019, mars). *Lithuanian Artificial Intelligence Strategy: a vision for the future*. Ministry of Economy and Innovation. [https://eimin.lrv.lt/uploads/eimin/documents/files/DI_strategija_ENG\(1\).pdf](https://eimin.lrv.lt/uploads/eimin/documents/files/DI_strategija_ENG(1).pdf)

- Luxembourg. (2019). *Artificial Intelligence: a strategic vision for Luxembourg*. The Government of the Grand Duchy of Luxembourg. https://digital-luxembourg.public.lu/sites/default/files/2020-09/AI_FR_1_0.pdf
- Luxembourg. (s.d.). *L'initiative AI4Gov*. Le gouvernement luxembourgeois. Services des médias, des communications et du numérique. Dernière mise à jour : 21.04.2021. https://smc.gouvernement.lu/fr/dossiers.gouv_digitalisation%2Bfr%2Bdossiers%2B2021%2BAI4Gov.html
- Malte. (2019a, octobre). *Malta the ultimate AI Launchpad: a strategy and vision for Artificial Intelligence in Malta 2030*. https://malta.ai/wp-content/uploads/2019/11/Malta_The_Ultimate_AI_Launchpad_vFinal.pdf
- Malte. (2019b, octobre). *Malta towards trustworthy AI: Malta's Ethical AI Framework*. https://malta.ai/wp-content/uploads/2019/10/Malta_Towards_Ethical_and_Trustworthy_AI_vFINAL.pdf
- Norvège. (2020, 14 janvier). *National Strategy for Artificial Intelligence*. Norwegian Ministry of Local Government and Modernisation. https://www.regjeringen.no/contentassets/1febbb2c4fd4b7d92c67ddd353b6ae8/en-gb/pdfs/ki-strategi_en.pdf
- OCDE. (2019). *Recommandation du Conseil sur l'intelligence artificielle*. OECD/LEGAL/0449. <https://oecd.ai/fr/ai-principles>
- OCDE. (2021). « Tools for trustworthy AI : A framework to compare implementation tools for trustworthy AI systems », *Documents de travail de l'OCDE sur l'économie numérique*, n° 312, Éditions OCDE, Paris, <https://doi.org/10.1787/008232ec-en>.
- OECD.AI. (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>
- Parlement européen. (2020a, 7 septembre). *Intelligence artificielle : définition et utilisation*. <https://www.europarl.europa.eu/news/fr/headlines/society/20200827STO85804/intelligence-artificielle-definition-et-utilisation>
- Parlement européen. (2020b, 20 octobre). *Cadre pour les aspects éthiques de l'intelligence artificielle, de la robotique et des technologies connexes*. P9_TA(2020)0275. https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_FR.pdf
- Pays-Bas. (2019). *Strategic Action Plan for Artificial Intelligence*. Ministry of Economic Affairs and Climate Policy. <https://www.government.nl/binaries/government/documents/reports/2019/10/09/strategic-action-plan-for-artificial-intelligence/Strategic+Action+Plan+for+Artificial+Intelligence.pdf>
- Portugal. (2019). *AI Portugal 2030: Portuguese national initiative on digital skills*. Coordination Office of INCoDe2030. https://www.incode2030.gov.pt/sites/default/files/julho_incode_brochura.pdf

- Portugal (s.d.). *Projeto GuIA Responsável*. Agência para a Modernização Administrativa (AMA). Tic.gov.pt. Repéré à <https://tic.gov.pt/pt/web/tic/guia>
- Portugal. (2020, décembre). *Guia para a inteligência artificial ética, transparente e responsável*. Agência para a Modernização Administrativa (AMA). Versão Beta. <https://tic.gov.pt/documents/37177/251624/GUIA+SHORT+VIII.pdf/2e8010a6-1013-fde7-7ef9-64783bcb6fdb>
- Portugal. (2021, juin). *Estratégia para a transformação digital da administração pública 2021-2026*. Conselho para as Tecnologias de Informação e Comunicação na Administração Pública. <https://bussola.gov.pt/Estratgias%20e%20Orientaes/Estrat%C3%A9gia%20de%20Transforma%C3%A7%C3%A3o%20Digital%20da%20AP%202021-2026.pdf>
- Qatar. (2019, octobre). *National artificial intelligence strategy for Qatar*. Ministry of Transport and Communications. <https://qcai-blog.qcri.org/wp-content/uploads/2020/04/QCRI-Artificial-Intelligence-Strategy-2019-ENG.pdf>
- Québec. (s.d.). *Stratégie de transformation numérique gouvernementale 2019-2023 : Contexte*. <https://www.quebec.ca/gouvernement/politiques-orientations/vitrine-numeriqc/strategie-numerique/a-propos>
- Québec. (2019). *Stratégie de transformation numérique gouvernementale 2019-2023 : Mesures clés*. Secrétariat du Conseil du Trésor. https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/secretariat-du-conseil-du-tresor/publications-adm/Transformation_numerique/2_Mesures_cles.pdf?1612901616
- Québec. (2021a). *Stratégie d'intégration de l'intelligence artificielle dans l'administration publique 2021-2026*. Secrétariat du Conseil du Trésor. https://cdn-contenu.quebec.ca/cdn-contenu/gouvernement/SCT/vitrine_numeriqc/strategie_IA/Strat_IA_2019_2023.pdf?1624995492
- Québec. (2021b). *Stratégie d'intégration de l'intelligence artificielle dans l'administration publique 2021-2026 : Mesures clés 2021-2023*. Secrétariat du Conseil du Trésor. https://cdn-contenu.quebec.ca/cdn-contenu/gouvernement/SCT/vitrine_numeriqc/strategie_IA/mesures_cles.pdf?1624995492
- Québec. (s.d.). *Ministère de la Cybersécurité et du Numérique : L'organisation et ses engagements*. <https://www.quebec.ca/gouvernement/ministere/cybersecurite-numerique>
- Secrétariat du Conseil du Trésor. (2020, 30 juin). *Le gouvernement du Québec lance sa Stratégie d'intégration de l'intelligence artificielle dans l'administration publique*. Gouvernement du Québec. <https://www.quebec.ca/nouvelles/actualites/details/le-gouvernement-du-quebec-lance-sa-strategie-dintegration-de-lintelligence-artificielle-dans-ladministration-publique-33048>

- Sept, G., & Canada. (2018). *Vision commune de Charlevoix sur l'avenir de l'intelligence artificielle* : Affaires mondiales Canada. https://www.international.gc.ca/world-monde/international_relations-reactions_internationales/g7/documents/2018-06-09-artificial-intelligence-artificielle.aspx?lang=fra
- Serbie. (2019). *Strategy for the Development of Artificial Intelligence in the Republic of Serbia for the period 2020-2025*. The Government of the Republic of Serbia. https://www.media.srbija.gov.rs/medsrp/dokumenti/strategy_artificial_intelligence.pdf
- Serbie. (2020, juin). *Action Plan of Strategy for Development of Artificial Intelligence adopted*. The Government of the Republic of Serbia. <https://www.srbija.gov.rs/vest/en/156855/action-plan-of-strategy-for-development-of-artificial-intelligence-adopted.php>
- Suède. (2018, mai). *National approach to artificial intelligence*. Government Offices of Sweden. <https://www.government.se/491fa7/contentassets/fe2ba005fb49433587574c513a837fac/national-approach-to-artificial-intelligence.pdf>
- Thomasset, A. (2020). Éditorial. Quelle éthique pour l'intelligence artificielle ? [Editorial. Which ethics for artificial intelligence?]. *Revue d'éthique et de théologie morale*, 307(3), 7-9. doi:10.3917/retn.310.0007
- Thomson Reuters. (2019, 25 octobre). Intelligence artificielle responsable – Principes 3 et 4. Éditions Yvon Blais. <https://www.editionsyvonblais.com/blogue/itech-law/intelligence-artificielle-responsable-principes-3-et-4-590/>
- UNESCO. (2021). Recommandation sur l'éthique de l'intelligence artificielle. UNESDOC Digital Library. https://unesdoc.unesco.org/ark:/48223/pf0000380455_fre
- Van Roy, V., Rossetti, F., Perset, K., Galindo-Romero, L. (2021). *AI Watch - National strategies on Artificial Intelligence: A European perspective, 2021 edition*. EUR 30745 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-39081-7, doi: 10.2760/069178, JRC122684. <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684>
- Varghese, J. (2019, octobre). *Qatar launches strategy to tap AI for future*. Gulf Times. <https://www.gulf-times.com/story/645930/Qatar-launches-strategy-to-tap-AI-for-future>
- Vrije Universiteit Brussel. (2021, 14 juillet). *New guidelines aim to correct discriminatory algorithms*. VUB Today. <https://today.vub.be/en/article/new-guidelines-aim-to-correct-discriminatory-algorithms>
- WhatIs.com. (2019, février). *Asilomar AI Principles*. <https://whatIs.techtarget.com/definition/Asilomar-AI-Principles>
- Wikipédia. (2021). *Benchmarking*. Wikipédia L'encyclopédie libre. <https://fr.wikipedia.org/wiki/Benchmarking>

Annexes

Analyse des stratégies IA - Bloc 1

Danemark et Union Européenne

Danemark

Stratégie nationale sur l'intelligence artificielle

(Mars 2019)



Faits saillants :

En mars 2019, le gouvernement danois a publié la Stratégie nationale sur l'intelligence artificielle contenant au total 24 initiatives axées sur quatre domaines : une base responsable pour l'intelligence artificielle ; plus et de meilleures données ; des compétences fortes et de nouvelles connaissances ; et une augmentation des investissements (Danemark, 2019).

La stratégie vise à garantir que le Danemark dispose d'une base éthique unifiée et centrée sur l'humain pour l'intelligence artificielle, à aider la croissance des entreprises danoises en développant et en utilisant l'IA et à faire en sorte que le secteur public danois utilise l'IA pour fournir d'excellents services au profit des citoyens et des organisations du Danemark.

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

Afin d'établir un cadre éthique et juridique pour une utilisation responsable de l'intelligence artificielle, la stratégie présente sept initiatives à mettre en œuvre, à savoir :

1. **Principes éthiques pour l'intelligence artificielle** : le gouvernement danois a préparé six principes éthiques afin d'établir un cadre commun pour le développement et l'utilisation de l'intelligence artificielle : **autodétermination, dignité, responsabilité, explicabilité, égalité et justice, développement**. Ces principes reflètent les valeurs danoises et ont été préparés sur la base des recommandations du groupe d'experts sur l'éthique des données mis en place dans le cadre du conseil danois *Disruption Council*, et du projet de principes éthiques et d'orientations de l'Union Européenne en matière d'intelligence artificielle.
2. **Création du Conseil d'éthique des données** : pour formuler des recommandations et soutenir un large débat public sur les questions relatives à l'éthique des données, y compris l'intelligence artificielle.

3. **Sécurité et intelligence artificielle** : une série d'initiatives sera lancée par le gouvernement danois pour soutenir le développement et la mise en œuvre sécurisés de l'intelligence artificielle, en veillant à ce que la confiance dans les solutions numériques soit préservée.
4. **Clarté juridique sur le développement et l'utilisation de l'intelligence artificielle** : avec la création d'un groupe de travail interministériel chargé d'examiner si les questions relatives à l'utilisation de l'intelligence artificielle peuvent être gérées dans le cadre législatif existant.
5. **Utilisation transparente des algorithmes par le secteur public** : en développant des méthodes et des lignes directrices communes au secteur public qui soutiennent les exigences légales en matière de transparence, etc. Dans le cadre de ces travaux, la stratégie prévoit le lancement d'un projet pilote visant à développer et à tester des méthodes permettant aux autorités publiques de respecter les exigences légales en matière d'utilisation raisonnable, responsable et transparente de l'IA comme base de prise de décision.
6. **Utilisation éthiquement responsable et durable des données par les entreprises** : la stratégie prévoit, entre autres, la préparation d'une boîte à outils contenant des outils spécifiques tels que des lignes directrices pour aider les entreprises dans leur travail de routine en matière de responsabilité des données, ainsi que l'introduction d'un label d'éthique des données pour les entreprises qui respectent les principes éthiques d'utilisation des données.
7. **L'empreinte danoise sur les normes relatives à l'intelligence artificielle** : le gouvernement danois veut s'assurer que les nouvelles normes internationales et nationales en matière d'intelligence artificielle sont élaborées en tenant compte des besoins spécifiques des entreprises danoises.

Initiatives mises en place :

- Un **Conseil d'éthique des données** indépendant (*Dataetisk Råd*) a été établi en mai 2019 avec l'objectif de conseiller les secteurs public et privé sur les questions d'éthique liées aux données et aux nouvelles technologies, y compris l'intelligence artificielle. Selon l'OCDE.AI (2021), le Conseil devrait définir en 2020 l'éthique des données et explorer les dimensions éthiques de la combinaison des données dans le secteur public. De manière plus générale, le Conseil prévoirait de renforcer la sensibilisation aux dilemmes éthiques, en partie par le biais d'un débat public. Pour suivre le rythme de la révolution numérique, le Conseil doit offrir une perspective continue sur la façon d'utiliser les données de manière responsable et mettre en évidence les possibilités de solutions nouvelles, innovantes et socialement

bénéfiques. Le Conseil doit soumettre un rapport annuel au gouvernement sur les recommandations et les activités³ ;

- En décembre 2019, une **boîte à outils d'éthique des données**⁴ a été lancée pour aider les entreprises à adopter et à mettre en œuvre l'éthique des données dans leurs modèles économiques (Van Roy et al., 2021). La boîte à outils contient des lignes directrices sur la manière d'inclure l'éthique des données dans le code de pratique de leur entreprise, des études de cas sur les meilleures pratiques et des outils d'évaluation d'impact sur l'utilisation responsable des algorithmes (Larsen, 2020) ;
- Le Conseil d'éthique des données a publié, le 11 octobre 2021, l'outil Ethique des Données – Comment le faire ('Dataetik – Sådan gør du!')⁵. Cet outil pratique comprend un guide en 5 étapes⁶ et devrait expliquer comment commencer à travailler avec l'éthique des données (Dataetisk Råd, 2021).

À surveiller :

Un **label commun de cybersécurité et d'éthique des données** sera lancé en 2021. Il est fondé par un consortium multipartite indépendant auquel participent les organisations industrielles nationales et le Conseil danois des consommateurs (Van Roy et al., 2021). Ce label sera un système d'étiquetage indépendant attribué aux entreprises qui répondent à ses exigences en matière de cybersécurité et de traitement responsable des données liées à l'IA. L'apposition du sceau indiquera aux consommateurs quelles entreprises traitent les données et l'IA d'une manière digne de confiance, éthique et sûre. En tant que label de qualité, il devrait inciter les acteurs du marché à faire preuve d'une plus grande éthique en matière de données (Larsen, 2020).

Liste des références et documents consultés :

Danemark (2019, mars). *National Strategy for Artificial Intelligence*. Agency for Digitisation. Ministry of Finance and Ministry of Industry, Business and Financial Affairs. Repéré à <https://en.digst.dk/policy-and-strategy/denmark-s-national-strategy-for-artificial-intelligence/>

³ Le Conseil d'éthique des données a publié son deuxième rapport annuel le 14 octobre 2021 : <https://dataetiskraad.dk/dataetisk-raad-udgiver-sin-anden-aarsrapport>

⁴ <https://virksomhedsguiden.dk/content/temaer/dataetik/>

⁵ <https://dataetiskraad.dk/nu-lanceres-dataetik-saadan-goer-du-et-praktisk-vaerktoej-til-naar-du-skal-traeffe-dataetiske>

⁶ <https://dataetiskraad.dk/sites/default/files/2021-10/Dataetisk%20Canvas.pdf>

- Dataetisk Råd (2021, 11 octobre). *Nu lanceres 'Dataetik – sådan gør du', et praktisk værktøj til når du skal træffe dataetiske beslutninger*. Repéré à <https://dataetiskraad.dk/nu-lanceres-dataetik-saadan-goer-du-et-praktisk-vaerktoej-til-naar-du-skal-traeffe-dataetiske>
- Larsen, F. (2020, 15 juin). *Denmark: an independent council and a labelling scheme to promote the ethical use of data*. OECD.AI. Repéré à <https://oecd.ai/en/wonk/an-independent-council-and-seal-of-approval-among-denmarks-measures-to-promote-the-ethical-use-of-data>
- OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>
- Van Roy, V., Rossetti, F., Perset, K., Galindo-Romero, L. (2021). *AI Watch - National strategies on Artificial Intelligence: A European perspective, 2021 edition*. EUR 30745 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-39081-7, doi: 10.2760/069178, JRC122684. Repéré à <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684>

Union Européenne

Plan coordonné pour le développement et l'utilisation de l'intelligence artificielle « made in Europe » - 2018

(Décembre 2018)



Faits saillants :

Dans le cadre de sa stratégie sur l'intelligence artificielle, adoptée en avril 2018, la Commission européenne a présenté le 7 décembre 2018 un plan coordonné élaboré avec les États membres pour promouvoir le développement et l'utilisation de l'intelligence artificielle en Europe. Le plan propose des actions communes pour une coopération plus étroite et plus efficace entre les États membres, la Norvège, la Suisse et la Commission dans quatre domaines clés : accroître les investissements, rendre davantage de données disponibles, promouvoir les talents et renforcer la confiance (Commission Européenne, 2018).

Selon OECD.AI (2021), ce plan rassemble un ensemble d'actions concrètes et complémentaires au niveau de l'UE, des pays et des régions en vue de :

- Stimuler les investissements et renforcer l'excellence dans les technologies et les applications d'IA qui sont dignes de confiance et "éthiques et sûres dès la conception". Les investissements doivent avoir lieu dans un contexte réglementaire stable qui permet l'expérimentation et soutient l'innovation de rupture dans toute l'UE, garantissant l'utilisation la plus large et la meilleure de l'IA par l'économie et la société européennes ;
- En s'appuyant sur les atouts de l'Europe, élaborer et mettre en œuvre, en partenariat avec les entreprises et les États membres, des programmes communs de recherche et développement (R&D) et d'innovation entre les entreprises et les universités ;
- Faciliter l'expérimentation de l'IA et le transfert de l'IA sur le marché. Créer un écosystème approprié pour les entreprises afin de faciliter l'adoption de l'IA ;
- Adapter les programmes et systèmes d'apprentissage et de formation pour préparer la société européenne et les générations futures à l'IA ;
- Mettre en place en Europe les capacités essentielles qui sous-tendent l'IA, telles que des espaces de données et des sites de référence de classe mondiale pour les essais et l'expérimentation ;
- Faire des administrations publiques européennes des pionniers de l'utilisation de l'IA ;

- Mettre en œuvre, sur la base de travaux d'experts, des lignes directrices claires en matière d'éthique pour le développement et l'utilisation de l'IA, dans le plein respect des droits fondamentaux, afin de fixer des normes éthiques mondiales et d'être un leader mondial de l'IA éthique et fiable ;
- Si nécessaire, revoir le cadre juridique national et européen existant pour mieux l'adapter aux défis spécifiques ;
- Favoriser la coopération internationale.

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

Alors, le plan prévoit, entre autres, qu'un environnement de confiance et de responsabilité autour du développement et de l'utilisation de l'IA soit assuré, avec le développement de lignes directrices éthiques et d'un cadre juridique approprié dans ce domaine.

Initiatives mises en place :

Les lignes directrices en matière d'éthique pour une IA digne de confiance ont été publiées en avril 2019, en définissant trois caractéristiques (licite, éthique, robuste), **quatre principes éthiques** (respect de l'autonomie humaine; prévention de toute atteinte; équité; explicabilité) et **sept éléments essentiels** pour sa mise en œuvre (**action humaine et contrôle humain; robustesse technique et sécurité; respect de la vie privée et gouvernance des données; transparence; diversité, non-discrimination et équité; bien-être sociétal et environnemental; responsabilité**).

Le document présente les orientations sous la forme de trois niveaux d'abstraction. Le chapitre I établit les fondements d'une IA digne de confiance (il recense et décrit les principes éthiques) ; le chapitre II traduit ces principes éthiques en sept exigences et propose des méthodes tant techniques (l'architecture du système d'IA, éthique et état de droit dès la conception, méthodes d'explication, essais et validations, qualité des indicateurs de service) que non techniques (comme réglementation, codes de conduite, normalisation, certification, cadres de gouvernance internes et externes, éducation et sensibilisation, participations des parties prenantes et dialogue social, diversité et équipes de conception inclusives) pour mettre en œuvre les exigences ; le chapitre III établit une liste d'évaluation préliminaire et non exhaustive pour une IA digne de confiance afin de concrétiser les exigences du chapitre II.

La dernière section expose des possibilités bénéfiques et des préoccupations importantes suscitées par les systèmes d'IA.

À surveiller :

1. En 20 octobre 2020, le Parlement européen (2020) a adopté une résolution contenant des recommandations à la Commission concernant un cadre pour les aspects éthiques de l'intelligence artificielle, de la robotique et des technologies connexes (2020/2012(INL)).

Le règlement repose sur les principes suivants : intelligence artificielle, robotique et technologies connexes axées sur l'humain, développées et contrôlées par l'homme ; évaluation obligatoire de la conformité de l'IA, de la robotique et des technologies connexes à haut risque ; sécurité, transparence et responsabilité ; garanties et solutions contre les biais et la discrimination ; droit de recours ; responsabilité sociale et égalité entre les genres en matière d'IA, de robotique et de technologies connexes ; IA, robotique et technologies connexes durables sur la plan environnemental ; respect de la vie privée et limitation de l'utilisation de la reconnaissance biométrique ; bonne gouvernance relative à l'IA, à la robotique et aux technologies connexes, y compris les données utilisées ou produites par ces technologies.

- Certificat européen de conformité éthique à être délivré par un organisme national de surveillance compétent (haut risque)
- Organismes de surveillance (contrôler l'application du règlement et réaliser les évaluations des risques et de la conformité et la certification)

2. En outre, en avril 2021, la Commission européenne a présenté à la fois une proposition de cadre réglementaire sur l'IA (2021a), avec des règles suivant une approche fondée sur les risques, et un plan coordonné révisé sur l'IA (2021b).

Les objectifs spécifiques du cadre réglementaire sont les suivants : veiller à ce que les systèmes d'IA mis sur le marché de l'Union et utilisés soient sûrs et respectent la législation en vigueur en matière de droits fondamentaux et les valeurs de l'Union ; garantir la sécurité juridique pour faciliter les investissements et l'innovation dans le domaine de l'IA ; renforcer la gouvernance et l'application effective de la législation existante en matière de droits fondamentaux et des exigences de sécurité applicables aux systèmes d'IA ; faciliter le développement d'un marché unique pour des applications d'IA légales, sûres et dignes de confiance, et empêcher la fragmentation du marché.

Liste des références et documents consultés :

Commission européenne (2018, 7 décembre). *Communication de la Commission au Parlement européen, au Conseil européen, au Conseil, au Comité économique et social européen et au Comité des régions – Plan coordonné dans le domaine de l'intelligence artificielle* (COM(2018) 795 final). Repéré à

https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56032 et à
https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56034

Commission européenne (2019, avril). *Les lignes directrices en matière d'éthique pour une IA digne de confiance*. Groupe d'experts de haut niveau sur l'intelligence artificielle. Repéré à <https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-fr/format-PDF>

Commission européenne (2021a). *Proposition de règlement du Parlement européen et du Conseil établissant des règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union*. COM(2021) 206 final. 2021, 21 avril. Repéré à <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CELEX:52021PC0206&from=EN>

Commission européenne (2021b). *Coordinated Plan on Artificial Intelligence 2021 Review*. 2021, 21 avril. Repéré à <https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review>

OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>

Parlement européen (2020). *Cadre pour les aspects éthiques de l'intelligence artificielle, de la robotique et des technologies connexes*. P9_TA(2020)0275. Repéré à https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_FR.pdf

Analyse des stratégies IA - Bloc 2

Finlande, Lituanie et Suède

Finlande

Ouvrir la voie à l'ère de l'intelligence artificielle (2019-2025)

(Juin 2019)



Faits saillants :

En octobre 2017, le ministère finlandais des affaires économiques et de l'emploi a publié sa stratégie nationale en matière d'intelligence artificielle (Finlande, 2017), qui fait partie d'un programme plus large sur l'intelligence artificielle en Finlande (AI Finland). Cette stratégie présentait les objectifs suivants : accroître la compétitivité des entreprises et de l'industrie ; fournir des services publics de haute qualité et améliorer l'efficacité du secteur public ; et assurer le bon fonctionnement de la société et le bien-être de ses citoyens.

En septembre 2018, la stratégie a été complétée par un rapport politique sur le travail à l'ère de l'intelligence artificielle, intitulé *Work in the age of artificial intelligence: Four perspectives on the economy, employment, skills and ethics*, traitant de l'impact de l'IA sur la dynamique du marché du travail et les besoins en compétences (Finlande, 2018, septembre).

En décembre 2018, le gouvernement finlandais a soumis au Parlement le rapport *Government report on information policy and artificial intelligence: Ethical information policy in the age of artificial intelligence*, proposant une nouvelle politique d'information pour promouvoir la bonne gestion et l'utilisation efficace de l'information (Finlande, 2018, décembre). Le rapport expose les principes d'une gouvernance équitable des données, y compris les lignes directrices pour l'utilisation des informations et les valeurs éthiques.

En février 2019, le bureau du Premier ministre a publié un rapport politique commandé par le gouvernement finlandais sur l'intelligence artificielle dans l'utilisation de l'autorité - questions d'acceptabilité éthique et sociale. Il étudie les concepts d'éthique et d'acceptabilité dans le contexte du développement et de l'application des technologies (Van Roy et al., 2021).

En juin 2019, le ministère finlandais des Affaires économiques a publié le rapport final du programme finlandais sur l'intelligence artificielle, intitulé *Leading the way into the age of artificial intelligence* (Finlande, 2019).

En novembre 2020, la Finlande a annoncé le programme *Artificial Intelligence 4.0*, fondé sur les objectifs de l'UE et la stratégie finlandaise en matière d'IA. Ce programme a l'objectif d'encourager le développement et l'introduction de l'IA et d'autres technologies numériques dans les entreprises, en mettant l'accent sur les PME de divers secteurs industriels et de services (Finlande, 2020).

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

La stratégie finlandaise en matière d'intelligence artificielle présente les objectifs suivants dans le domaine de l'éthique (Kung, 2020) :

- Équilibrer la protection de la vie privée et d'autres droits avec les avantages de l'IA ;
- Faire de la Finlande un banc d'essai international pour la mise en œuvre d'une IA éthique ;
- Établir un conseil national d'éthique pour la technologie ;
- Inviter les citoyens à des discussions sur l'éthique et les répercussions sociétales de l'IA ;
- Assurer l'introduction centrée sur l'humain de l'intelligence artificielle et la mise en œuvre de principes éthiques dans le secteur public à travers le projet *AuroraA⁷I* ;
- Encourager les entreprises et les acteurs du secteur public à s'auto-réglementer sur le plan éthique et à partager les meilleures pratiques ;
- Augmenter la recherche interdisciplinaire à long terme sur les impacts sociétaux de l'IA.

Initiatives mises en place :

En Finlande, nous pouvons voir les initiatives suivantes qui ont été prises pour le développement et l'utilisation responsable de l'IA (Van Roy et al., 2021) :

⁷ "The aim of the Aurora national artificial intelligence (AI) programme is to implement an operations model based on people's needs, where artificial intelligence helps citizens and companies to utilise services in a timely and ethically sustainable manner. The activities of the relevant authorities will be organised in such a way that they support people's life-events and companies' business-related events, facilitating seamless, effective and smoothly functioning service paths consisting of the services of several service providers. This will provide people with access to a new way of taking care of their overall well-being and, at the same time, will promote service providers' ability to form customer-oriented and dynamic service chains in collaboration with other operators and to manage their activities based on up-to-date information". <https://vm.fi/en/auroraai-en>

- **Comité d'éthique de l'IA** : le gouvernement finlandais a mis en place un comité d'éthique de l'IA visant à mieux comprendre les principes éthiques et à s'assurer que le développement de l'IA en Finlande est axé sur l'homme et fondé sur la confiance ;
- Mise en place d'un **défi éthique** de l'IA sur le site du programme d'intelligence artificielle pour inciter les entreprises à contribuer à la création de principes éthiques pour l'IA ;
- **Projet de recherche** *Ethical AI for the Governance of the Society* (ETAIIROS) : un projet financé par l'Académie de Finlande et auquel collaborent plusieurs universités finlandaises, dont l'objectif est d'étudier et de co-créer des cadres et des outils de gouvernance qui répondent au défi d'une utilisation socialement durable de l'IA ;
- **Registre d'IA** d'Helsinki (en partenariat avec la ville d'Amsterdam) : Helsinki et Amsterdam ont lancé des registres d'IA ouverts qui permettent de suivre la manière dont les algorithmes sont utilisés dans les municipalités. Un livre blanc sur le registre de l'IA a été publié en septembre 2020 pour inspirer d'autres gouvernements/organisations qui veulent être transparents sur leur utilisation des algorithmes et de l'IA ;
- **Cours en ligne sur l'éthique de l'IA**, offert par l'université d'Helsinki : ce cours, conçu en partenariat avec le ministère finlandais des finances et les villes d'Helsinki, d'Amsterdam et de Londres, aide les administrations publiques, les entreprises et le grand public à comprendre ce que signifie l'utilisation éthique de l'IA et ce qu'elle exige de la société et des individus.

À surveiller :

Une autre politique envisagée par le gouvernement finlandais consiste à préparer la base des directives éthiques du programme *AuroraAI* en se fondant sur l'écosystème de l'administration publique (Van Roy et al., 2021).

Liste des références et documents consultés :

Finlande (s.d.-a). *Ethics challenge*. AI Finland. Ministry of Economic Affairs and Employment of Finland. Repéré à <https://www.tekoalyaika.fi/en/background/ethics/>

Finlande (s.d.-b). *Implementation of the national AuroraAI programme*. Ministry of Finance. Repéré à <https://vm.fi/en/auroraai-en>

Finlande (2017). *Finland's Age of Artificial Intelligence. Turning Finland into a leading country in the application of artificial intelligence*. Ministry of Economic Affairs and Employment. Repéré à

https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_v_erkkojulkaisu.pdf?sequence=1&isAllowed=y

Finlande (2018, 10 septembre). *Work in the age of artificial intelligence. Four perspectives on the economy, employment, skills and ethics*. Ministry of Economic Affairs and Employment. Repéré à

http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160980/TEMjul_21_2018_Work_in_the_age.pdf

Finlande (2018, 5 décembre). *Government report on information policy and artificial intelligence: Ethical information policy in the age of artificial intelligence*. Ministry of Finance. Repéré à

https://vm.fi/documents/10623/7768305/VM_Tiepo_selonteko_070219_ENG_WEB.pdf/89b99a8e-01a3-91e3-6ada-38056451ad3f/VM_Tiepo_selonteko_070219_ENG_WEB.pdf.pdf/VM_Tiepo_selonteko_070219_ENG_WEB.pdf?t=1560838655000

Finlande (2019, 12 juin). *Leading the way into the age of artificial intelligence. Final report of Finland's Artificial Intelligence Programme 2019*. Ministry of Economic Affairs and Employment. Repéré à

http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161688/41_19_Leading%20the%20way%20into%20the%20age%20of%20artificial%20intelligence.pdf

Finlande (2020, 13 novembre). *Artificial Intelligence 4.0 programme to speed up digitalization of business*. Ministry of Economic Affairs and Employment. Repéré à

<https://tem.fi/en/-/artificial-intelligence-4.0-programme-to-speed-up-digitalisation-of-business>

Kung, J. (2020, mai). *L'ère de l'IA : rapport sur les stratégies nationales et régionales en matière d'IA*. CIFAR. Deuxième Édition. Repéré à <https://cifar.ca/wp-content/uploads/2020/11/l-ere-de-l-ia-deuxieme-edition-f.pdf>

OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>

Van Roy, V., Rossetti, F., Perset, K., Galindo-Romero, L. (2021). *AI Watch - National strategies on Artificial Intelligence: A European perspective, 2021 edition*. EUR 30745 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-39081-7, doi: 10.2760/069178, JRC122684. Repéré à <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684>

Lituanie

Stratégie lituanienne de l'intelligence artificiel : Une vision pour l'avenir (Mars 2019)



Faits saillants :

En mars 2019, le ministère de l'Économie et de l'Innovation a présenté la stratégie lituanienne de l'intelligence artificielle : une vision pour l'avenir (Lituanie, 2019), avec l'objectif de moderniser et développer l'écosystème actuel de l'IA en Lituanie.

La stratégie présente une série de recommandations politiques dans des domaines clés dans le but de (Van Roy et al., 2021) :

- Améliorer les compétences et l'éducation en matière d'IA pour tous les citoyens ;
- Renforcer l'écosystème national de recherche et d'innovation dans le domaine de l'IA ;
- Accroître le déploiement, le développement et l'utilisation de l'IA dans toutes les activités économiques, y compris dans les secteurs privé et public ;
- Promouvoir les collaborations nationales et internationales dans le domaine de l'IA et améliorer les opportunités de réseau ;
- Développer un cadre éthique et juridique pour un développement durable et transparent des applications de l'IA ;
- Mettre en place un écosystème de données responsable et efficace pour l'IA.

Comme le notent Van Roy et ses collègues (2021), la stratégie lituanienne ne comprend pas d'initiatives politiques concrètes, mais sert simplement de document d'orientation pour tous les acteurs du pays, avec des recommandations politiques.

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

Les principaux mécanismes énoncés dans la stratégie de la Lituanie visant à une réglementation éthique et juridique accrue pour le développement et l'utilisation de l'IA de confiance dans le pays sont les suivants (Lituanie, 2019) :

- Établir un comité d'éthique multidisciplinaire sur l'IA pour examiner l'impact de la technologie sur les droits fondamentaux. Les recommandations du comité devraient

être utilisées pour créer et mettre à jour les normes éthiques existantes en Lituanie. En outre, le comité devrait encourager le développement de compétences liées à l'IA dans l'administration publique ;

- Développer des mécanismes de rétroaction permettant à la société de s'exprimer sur les réglementations créées par l'administration publique ;
- Élaborer des règles, des normes, des lignes directrices, des standards et des principes éthiques afin de guider le développement éthique et durable de l'IA et de son utilisation ;
- Œuvrer en faveur des normes et réglementations lituaniennes, européennes et internationales qui favorisent l'utilisation de l'IA et préviennent les risques ;
- Établir des principes directeurs et des politiques pour l'étude des meilleures pratiques en matière de conception, de construction et d'utilisation, y compris la formation adéquate des opérateurs pour un fonctionnement sûr ;
- Faire des investissements supplémentaires pour faire progresser la sûreté et la sécurité de l'IA, notamment l'explicabilité et la transparence, la confiance, la vérification et la validation, la sécurité contre les attaques, ainsi que la sûreté de l'IA à long terme et l'alignement des valeurs ;
- Soutenir la recherche pour minimiser les biais dans les systèmes d'IA ;
- Créer des labels de qualité pour les entreprises concernées qui respectent les normes définies par le comité d'éthique de l'IA. Rendre les résultats publics ;
- Faciliter la création d'un centre national interdisciplinaire sur l'IA pour promouvoir les discussions sur l'éthique de l'IA ;
- Mettre en place un mécanisme de sauvegarde pour que les chercheurs développent des systèmes transparents et intrinsèquement capables d'expliquer aux utilisateurs les raisons de leurs résultats ;
- Encourager les écoles secondaires à discuter des implications éthiques de la technologie et les universités à développer des cours sur ces implications pour le plus grand nombre de disciplines possible. L'éducation éthique doit faire partie intégrante de l'enseignement supérieur technique et non technique de l'IA et des programmes connexes des universités et des collèges ;
- Demander aux programmes universitaires d'informatique et d'ingénierie de suivre un cours sur l'éthique de la technologie ;
- Faciliter le débat sur l'IA et ses implications éthiques dans l'ensemble de la société ;
- Développer des capacités de médiation pour soutenir les populations vulnérables ;
- Travailler activement avec les principes relatifs à l'infrastructure des données afin de garantir la protection de la vie privée et la sécurité.

À surveiller :

- Selon OECD.AI (2021), les lignes directrices étaient en cours d'élaboration et devaient être approuvées au quatrième trimestre 2020. Elles contiendront des guides éthiques et des recommandations à l'intention des développeurs et des responsables de niveau C, et définiront l'approche générale en matière de sélection et d'annotation des données, etc.
- La Lituanie envisage un réexamen et, si nécessaire, une mise à jour de sa stratégie nationale en matière d'IA en 2021 (Van Roy et al., 2021).

Liste des références et documents consultés :

Lituanie (2019, mars). *Lithuanian Artificial Intelligence Strategy: a vision for the future*. Ministry of Economy and Innovation. Repéré à [https://eimin.lrv.lt/uploads/eimin/documents/files/DI_strategija_ENG\(1\).pdf](https://eimin.lrv.lt/uploads/eimin/documents/files/DI_strategija_ENG(1).pdf)

OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>

Van Roy, V., Rossetti, F., Perset, K., Galindo-Romero, L. (2021). *AI Watch - National strategies on Artificial Intelligence: A European perspective, 2021 edition*. EUR 30745 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-39081-7, doi: 10.2760/069178, JRC122684. Repéré à <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684>

Suède

Approche nationale de l'intelligence artificielle

(Mai 2018)



Faits saillants :

En mai 2018, le gouvernement suédois a lancé sa stratégie nationale de l'intelligence artificielle qui vise à définir une orientation générale pour les travaux liés à l'IA en Suède et à définir les bases des futures actions et priorités politiques, en renforçant le bien-être et la compétitivité du pays grâce à l'IA. À cette fin, la stratégie suédoise se concentre sur les domaines prioritaires suivants : éducation et formation ; recherche ; innovation et utilisation ; et structure et infrastructure (Suède, 2018).

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

Pour favoriser la réglementation de l'IA, le gouvernement a identifié la nécessité de développer des règles, des normes, des standards et des principes éthiques pour une IA éthique et durable et son utilisation, et de faire pression pour des normes et des réglementations suédoises et internationales qui favorisent une utilisation sans risque de l'IA.

En outre, pour promouvoir la création d'une IA éthique et durable la stratégie prévoit les initiatives suivantes :

- La création d'un centre de durabilité de l'IA qui se concentrera sur les aspects sociaux et éthiques de l'IA ;
- L'offre de séminaires académiques sur les défis éthiques de l'IA dans les affaires, l'administration et divers domaines sectoriels.

La stratégie suédoise recommande également l'élaboration de standards (internationaux) appropriés.

Initiatives mises en place :

- Le centre de durabilité de l'IA (*AI Sustainability Center*) a été créé (Van Roy et al., 2021). Il s'agit d'un centre cofondé par des entreprises, des universités et des autorités

publiques, qui se concentre spécifiquement sur les aspects sociaux et éthiques de l'IA. Le centre a développé un cadre, des outils et des services en collaboration avec certains des principaux instituts de recherche mondiaux et des organisations de premier plan dans le domaine de l'IA afin de guider les parties prenantes vers une IA éthique et durable ;

- Le Comité pour l'innovation technologique et l'éthique (*Komet*) a été établi le 14 août 2018 avec la mission d'aider le gouvernement suédois à identifier les défis politiques, de contribuer à réduire l'incertitude entourant les réglementations existantes et d'accélérer le développement des politiques liées aux technologies de la quatrième révolution industrielle (OECD.AI, 2021). Le rapport final est attendu pour le 31 décembre 2022. Le comité doit soumettre des rapports intermédiaires annuels (*Komet*, s.d.).

À surveiller :

- Le rapport final du Comité pour l'innovation technologique et l'éthique est attendu pour le 31 décembre 2022 (*Komet*, s.d.).
- Une nouvelle législation suédoise visant à promouvoir l'utilisation responsable de l'IA et à prévenir les risques est en préparation (*Van Roy et al.*, 2021).

Liste des références et documents consultés :

Komet (s.d.). *About us*. Kommittén för teknologisk innovation & etik (en ligne). Consulté en 3/11/2021. Repéré à <https://www.kometinfo.se/in-english/about-us/>

OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>

Suède (2018, mai). *National approach to artificial intelligence*. Government Offices of Sweden. Repéré à <https://www.government.se/491fa7/contentassets/fe2ba005fb49433587574c513a837fac/national-approach-to-artificial-intelligence.pdf>

Van Roy, V., Rossetti, F., Perset, K., Galindo-Romero, L. (2021). *AI Watch - National strategies on Artificial Intelligence: A European perspective, 2021 edition*. EUR 30745 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-39081-7, doi: 10.2760/069178, JRC122684. Repéré à <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684>

Analyse des stratégies IA - Bloc 3

Estonie, Malte, Norvège et Qatar

Estonie

Stratégie nationale sur l'intelligence artificielle 2019-2021

(Juillet 2019)



Faits saillants :

En mai 2019, un groupe d'experts dirigé par le ministère des Affaires économiques et des Communications et le bureau du gouvernement (*Estonia's AI Taskforce*) a présenté un rapport politique ainsi que des propositions pour faire progresser l'adoption de l'IA en Estonie (Estonie, 2019a).

Sur la base de ces propositions, le ministère des Affaires économiques et des Communications a présenté en juillet de la même année la stratégie nationale de l'Estonie en matière d'IA (Estonie, 2019b). Cette stratégie vise à tirer le meilleur parti du potentiel de l'IA en élaborant et en mettant en œuvre des mesures politiques dans le but de (Van Roy et al., 2021) :

- Encourager l'utilisation et le développement des applications de l'IA dans les secteurs public et privé ;
- Fournir un soutien direct à la recherche sur l'IA et renforcer les aptitudes et les compétences pertinentes à cet effet ;
- Développer l'environnement juridique pour faciliter le déploiement de l'IA.

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

Le rapport préparé par le groupe d'experts *Estonia's AI Taskforce* a présenté les principes éthiques suivants liés à l'IA (Estonie, 2019a) :

- Le principe d'utilité : faire le bien !
- Le principe de non-nuisance : ne pas nuire !

- Le principe d'autonomie : respecter le droit des personnes à prendre leurs propres décisions !
- Le principe de justice : être libre de tout préjugé !
- Le principe de clarté : agir avec transparence !

Ainsi, les éléments suivants ont été mis en avant dans le rapport pour atteindre l'objectif éthique de l'IA :

- Veiller à ce que l'IA soit centrée sur l'homme ;
- Découler des droits fondamentaux, des principes éthiques et des valeurs. Il est particulièrement important de se tourner vers les groupes les plus vulnérables de la société (enfants, personnes handicapées, minorités, employés) ;
- Être conscient et prudent du fait qu'en plus des avantages, l'IA peut également entraîner des conséquences inattendues. Il est particulièrement important de rester vigilant dans les domaines d'importance critique.

Compte tenu de ces propositions, la stratégie nationale de l'Estonie présente les initiatives suivantes pour le développement et l'utilisation responsables de l'intelligence artificielle (Estonie, 2019b) :

- Rédiger une proposition de principes d'utilisation responsable des données ;
- Créer et publier des documents d'orientation pour le lancement et l'évaluation de projets d'IA, y compris le développement responsable et la gestion durable de solutions d'IA, la méthodologie d'évaluation d'impact, etc.

Initiatives mises en place :

Le gouvernement estonien travaille actuellement à l'élaboration d'un questionnaire d'auto-évaluation destiné aux développeurs d'IA et basé sur la liste d'évaluation accompagnant les Lignes directrices éthiques pour une IA digne de confiance (Van Roy et al., 2021).

À surveiller :

Un groupe de travail sera créé pour suivre la mise en œuvre de la stratégie estonienne, pour lancer des initiatives politiques supplémentaires si nécessaire et pour commencer la préparation de la stratégie à long terme de l'Estonie en matière d'IA pour 2022 (Van Roy et al., 2021).

Liste des références et documents consultés :

- Estonie (2019a, mai). *Report of Estonia's AI Taskforce*. Government Office and Ministry of Economic Affairs and Communications. Repéré à https://f98cc689-5814-47ec-86b3-db505a7c3978.filesusr.com/ugd/7df26f_486454c9f32340b28206e140350159cf.pdf
- Estonie (2019b, juillet). *Estonia's national artificial intelligence strategy 2019-2021*. Government of the Republic of Estonia. Repéré à https://f98cc689-5814-47ec-86b3-db505a7c3978.filesusr.com/ugd/7df26f_27a618cb80a648c38be427194affa2f3.pdf
- OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>
- Van Roy, V., Rossetti, F., Perset, K., Galindo-Romero, L. (2021). *AI Watch - National strategies on Artificial Intelligence: A European perspective, 2021 edition*. EUR 30745 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-39081-7, doi: 10.2760/069178, JRC122684. Repéré à <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684>

Malte

Malte : plateforme par excellence de lancement de l'intelligence artificielle — Stratégie et vision de l'intelligence artificielle à Malte 2030

(Octobre 2019)



Faits saillants :

En octobre 2019, le gouvernement maltais a publié la stratégie nationale de Malte en matière d'IA (Malte, 2019a). Cette stratégie a été élaborée en s'appuyant sur trois piliers (Van Roy et al., 2021) :

- La création d'un écosystème d'IA robuste basé sur l'investissement, le soutien aux start-up et l'innovation ;
- Soutien à une augmentation de l'adoption de l'IA dans le secteur public ;
- Mesures de soutien à l'adoption de l'IA dans le secteur privé.

Chaque pilier repose sur trois éléments stratégiques : l'éducation et la main-d'œuvre, le cadre juridique et éthique et l'infrastructure.

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

Pour encourager l'adoption de l'IA, un accent particulier est mis sur le développement de la réglementation et de la législation, notamment avec la nomination d'un comité consultatif sur la réglementation des technologies chargé de conseiller sur les questions juridiques et la création d'un bac à sable réglementaire pour l'IA, prévoyant des exceptions réglementaires pour les entreprises afin de promouvoir les tests de solutions d'IA. En outre, un bac à sable pour les données sera mis au point pour fournir des orientations sur la manière dont les règles de protection des données s'appliquent.

En termes de normalisation, la stratégie maltaise en matière d'IA souligne l'importance de collaborer avec des organisations internationales sur les normes et standards émergents en matière d'IA. Cela va de pair avec l'élaboration d'une définition commune de l'IA, alignée sur celle du groupe d'experts de haut niveau de l'UE sur l'IA.

Et pour faire en sorte que le développement de l'IA soit conforme à l'éthique, transparent et socialement responsable, Malte a prévu les initiatives suivantes :

- Finaliser et mettre en œuvre le cadre éthique de l'IA ;
- Mettre en place un comité national d'éthique des technologies sous l'égide de la *Malta Digital Innovation Authority* (MDIA)⁸ pour superviser le cadre éthique de l'IA et son intersection avec diverses initiatives politiques, notamment les investissements dans les outils et les mécanismes de surveillance continue, les compétences et les capacités, l'écosystème et les mécanismes de réglementation ;
- Lancer un programme national de certification de l'IA pour les systèmes développés de manière éthique, transparente et socialement responsable.

Initiatives mises en place :

- Une première version du **cadre éthique de l'IA** a été publiée en 09 août 2019 pour une consultation publique. En octobre 2019 la version révisée a été publiée (*Malta Towards Trustworthy AI*), avec **quatre principes éthiques IA (autonomie humaine, prévention des dommages, équité et explicabilité)** et **sept éléments essentiels**, non exhaustives, pour sa mise en œuvre (**agence humaine, vie privée et gouvernance des données, explicabilité et transparence, bien-être, responsabilité, équité et impartialité, performance et sécurité**). Documents qui ont été considérés lors de l'élaboration du cadre : Lignes directrices éthiques pour une IA digne de confiance publiées en 8 avril 2019 par le groupe d'experts de haut niveau sur l'intelligence artificielle mis en place par la Commission européenne ; Recommandations du Conseil sur l'intelligence artificielle adoptées le 21 mai 2019 par les pays de l'OCDE et un certain nombre d'adhérents non-membres ; les principes d'IA d'Asilomar ; la conception alignée sur l'éthique de l'*Institute of Electrical and Electronics Engineers* (IEEE) ; la déclaration de Montréal ; et les cadres éthiques de l'IA publiés par d'autres gouvernements et entreprises technologiques (Malte, 2019b).
- **Programme de certification en IA** (octobre 2019) - MDIA (*Malta Digital Innovation Authority*) est la principale autorité responsable de la promotion de toutes les politiques gouvernementales qui favorisent l'innovation technologique à Malte. Elle a maintenant étendu son programme de certification à l'IA. La certification se fait sur une base volontaire ; toutefois, elle peut être imposée par les autorités qui souhaitent mettre en œuvre des solutions technologiques innovantes dans le cadre de leurs activités. En outre, le programme de certification est destiné aux solutions technologiques innovantes considérés comme présentant un risque moyen ou élevé (OECD.AI, 2021).

⁸ <https://mdia.gov.mt/>

À surveiller :

Comme indiqué plus haut, la stratégie de Malte prévoit la création d'un comité national d'éthique technologique sous l'égide de l'autorité maltaise de l'innovation numérique (MDIA), chargé de superviser le cadre éthique de l'IA (OECD.AI, 2021).

Liste des références et documents consultés :

Malte (2019a, octobre). *Malta the ultimate AI Launchpad: a strategy and vision for Artificial Intelligence in Malta 2030*. Repéré à https://malta.ai/wp-content/uploads/2019/11/Malta_The_Ultimate_AI_Launchpad_vFinal.pdf

Malte (2019b, octobre). *Malta towards trustworthy AI: Malta's Ethical AI Framework*. Repéré à https://malta.ai/wp-content/uploads/2019/10/Malta_Towards_Ethical_and_Trustworthy_AI_vFINAL.pdf

OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>

Van Roy, V., Rossetti, F., Perset, K., Galindo-Romero, L. (2021). *AI Watch - National strategies on Artificial Intelligence: A European perspective, 2021 edition*. EUR 30745 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-39081-7, doi: 10.2760/069178, JRC122684. Repéré à <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684>

Norvège

La stratégie nationale pour l'intelligence artificielle

(Janvier 2020)



Faits saillants :

Le gouvernement de la Norvège a présenté sa stratégie nationale en matière d'intelligence artificielle en 14 janvier 2020, décrivant les actions politiques pour les années à venir afin de maximiser les opportunités que l'IA peut apporter aux individus, aux entreprises et à l'industrie norvégiens, ainsi qu'au secteur public (Norvège, 2020a).

Pour atteindre ce résultat, la stratégie nationale en matière d'intelligence artificielle met en avant les initiatives politiques suivantes (Van Roy et al., 2021) :

- Développer l'offre de programmes éducatifs et de formations sur le lieu de travail dans le domaine de l'IA afin de créer une base solide de compétences et de capacités numériques ;
- Renforcer la recherche norvégienne en matière d'IA ;
- Accroître la capacité d'innovation en matière d'IA dans les secteurs privé et public ;
- Définir des principes éthiques pour l'IA afin de permettre des développements équitables, fiables et dignes de confiance dans ce domaine ;
- Établir des réglementations favorables à la numérisation afin de définir le cadre législatif dans lequel s'inscrivent les développements de l'IA ;
- Construire une infrastructure de données solide garantissant l'ouverture et le partage des données dans tous les secteurs et domaines d'activité ;
- Déployer une infrastructure de télécommunications offrant une connectivité et une puissance de calcul de grande capacité, et garantissant la sécurité des systèmes basés sur l'IA.

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

Pour le développement et l'utilisation responsable de l'IA, le gouvernement norvégien soutient dans sa stratégie que (Norvège, 2020a) :

- L'intelligence artificielle qui est développée et utilisée en Norvège devrait être fondée sur des principes éthiques et respecter les droits de l'homme et la démocratie ;
- La recherche, le développement et l'utilisation de l'intelligence artificielle en Norvège doivent promouvoir une IA responsable et digne de confiance ;
- Le développement et l'utilisation de l'IA en Norvège doivent préserver l'intégrité et la vie privée de l'individu ;
- La cybersécurité doit être intégrée dans le développement, le fonctionnement et l'administration des systèmes qui utilisent l'IA ;
- Les autorités de surveillance doivent veiller à ce que les systèmes d'IA dans leurs domaines de surveillance soient exploités conformément aux principes d'une utilisation responsable et digne de confiance de l'IA.

Dans cette optique, les initiatives suivantes, entre autres, ont été prévues dans la stratégie de la Norvège pour promouvoir une IA responsable, redevable, transparente et digne de confiance :

- Fixer des exigences en matière de transparence et de responsabilité dans les systèmes publics utilisant l'IA ;
- Créer un organe consultatif et un bac à sable (*sandbox*) réglementaire sur la protection de la vie privée et des données ;
- Encourager le développement et l'utilisation de l'IA conformément aux lignes directrices éthiques européennes, en adoptant les sept principes proposés par le Groupe d'experts de haut niveau de la Commission européenne (**action humaine et contrôle humain ; robustesse technique et sécurité ; respect de la vie privée et gouvernance des données ; transparence ; diversité, non-discrimination et équité ; bien-être sociétal et environnemental ; responsabilité**) ;
- Encourager l'industrie et les organisations à établir leurs propres normes industrielles ou leurs propres systèmes d'étiquetage ou de certification basés sur les principes d'une utilisation responsable de l'IA ;
- Encourager les établissements d'enseignement à réfléchir à la manière dont la vie privée et l'éthique peuvent occuper une place centrale dans leurs programmes d'intelligence artificielle ;
- Mettre en place un forum de coopérations pour les organismes chargés de l'application des lois sur la consommation, la concurrence et la protection des données : *Digital Clearinghouse Norway* ;
- Continuer à participer aux forums européens et internationaux, notamment aux travaux de l'UE visant à créer un cadre réglementaire pour promouvoir une utilisation

responsable et fiable de l'IA et à moderniser les droits des consommateurs à la lumière des développements numériques ;

- Stimuler le débat public sur l'utilisation éthique de l'IA.

Initiatives mises en place :

En 2020, l'autorité norvégienne de protection des données (DPA) a créé un bac à sable (*sandbox*) réglementaire pour l'IA afin de promouvoir le développement de solutions d'IA éthiques et responsables (Van Roy et al., 2021).

Selon OCDE.AI (2021), cela permettra de remplir plusieurs objectifs :

- Les entreprises peuvent mieux comprendre les exigences réglementaires en matière de protection des données et réduire le délai entre le développement et les tests et le déploiement effectif des solutions d'IA sur le marché. Les systèmes qui sont déployés après avoir été développés dans le bac à sable peuvent servir d'exemples à suivre et aider d'autres entreprises qui souhaitent développer des systèmes similaires ;
- Les autorités peuvent mieux comprendre les nouvelles solutions technologiques et identifier plus facilement les risques et les problèmes potentiels à un stade précoce, de sorte que des documents d'orientation peuvent être produits pour clarifier la manière dont les règlements doivent être appliqués ;
- Les autorités et les industries peuvent identifier les secteurs qui ont besoin de leurs propres normes industrielles ;
- Les individus et la société dans son ensemble bénéficieront des solutions nouvelles et innovantes développées dans le cadre de paramètres responsables.

Liste des références et documents consultés :

Datatilsynet (s.d.). *Sandbox for responsible artificial intelligence*. Repéré à <https://www.datatilsynet.no/en/regulations-and-tools/sandbox-for-artificial-intelligence/>

Norvège (2020, 14 janvier). *National Strategy for Artificial Intelligence*. Norwegian Ministry of Local Government and Modernisation. Repéré à https://www.regjeringen.no/contentassets/1febbb2c4fd4b7d92c67ddd353b6ae8/en-gb/pdfs/ki-strategi_en.pdf

OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>

Van Roy, V., Rossetti, F., Perset, K., Galindo-Romero, L. (2021). *AI Watch - National strategies on Artificial Intelligence: A European perspective, 2021 edition*. EUR 30745 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-39081-7, doi: 10.2760/069178, JRC122684. Repéré à <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684>

Qatar

Plan d'action : stratégie nationale du Qatar sur l'intelligence artificielle (Octobre 2019)



Faits saillants :

Début 2019, le *Qatar Center for Artificial Intelligence* (QCAI), qui fait partie du *Qatar Computing Research Institute* - QCRI, a élaboré et publié un plan pour une stratégie nationale d'intelligence artificielle pour le Qatar, afin d'informer et de conseiller les dirigeants du pays sur le potentiel de l'IA et sur une voie à suivre pour le Qatar (Dickson, 2019). Dans ce document, le QCAI a proposé six piliers pour cette stratégie et les recommandations clés pour chacun de ces piliers sont les suivantes (Chauhan, 2019) :

1. **Talent** : Construire un écosystème stimulant et attractif pour développer et mettre à niveau les talents locaux dans le domaine de l'IA et attirer des chercheurs et des entrepreneurs internationaux pour venir au Qatar et se développer.
2. **Accès aux données** : Les données sont la ressource stratégique clé qui sous-tend la technologie de l'IA. Le Qatar devrait élaborer sa politique en matière de production et d'accès aux données afin d'équilibrer le compromis entre la garantie de la vie privée et le respect des normes culturelles tout en permettant leur utilisation pour créer des solutions innovantes et personnalisées pour ses citoyens. Il devrait également initier et mener des efforts diplomatiques multilatéraux pour le partage des données en vue d'une croissance inclusive de l'IA au niveau mondial.
3. **L'emploi** : L'IA va transformer l'état actuel de l'emploi dans le monde. Les emplois de cols bleus et de cols blancs seront touchés. Le gouvernement du Qatar doit nommer un groupe de travail chargé d'étudier l'impact de l'émergence de la technologie de l'IA sur la main-d'œuvre du Qatar et d'élaborer un plan pour tirer parti de ces changements afin de réaliser le QNV 2030.
4. **Création de richesses** : Le Qatar devrait se positionner comme la destination la plus attrayante pour créer et développer des entreprises d'IA. Faire des investissements dans les personnes, l'infrastructure cloud et prendre la tête des questions internationales entourant l'IA comme la création de politiques, l'établissement de normes et de protocoles pour assurer un terrain de jeu équitable pour toutes les parties prenantes. Utiliser l'IA pour augmenter les rendements

grâce à des gains d'efficacité et à de nouvelles opportunités génératrices de revenus au Qatar et dans le monde.

5. **Transformer le Qatar en une nation IA + X** : L'avenir activé par l'IA conduira à un paradigme du "gagnant prend tout". Le Qatar doit réaliser des investissements de recherche dans des domaines d'importance stratégique et dans lesquels le pays bénéficie d'un avantage naturel, notamment : Les entreprises de contenu arabe, le pétrole et le gaz, les soins de santé, la sécurité nationale et les transports.
6. **Leadership éclairé en matière d'éthique de l'IA** : Élaborer un cadre d'éthique et de gouvernance de l'IA ancré dans le contexte local et aligné sur les normes internationales. Alors que l'IA s'étend à tous les aspects de la vie et de la société, son impact sur les questions politiques et éthiques liées à la gouvernance, à l'ordre public, à la santé et à la guerre nécessite un débat et un leadership réfléchis.

Le 29 octobre 2019, ce document a été adopté et annoncé officiellement par le gouvernement qatari comme la stratégie nationale en matière d'intelligence artificielle pour le pays (Hukoomi, 2021).

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

Pour un fondement responsable pour l'IA, la stratégie qatarie prévoit les initiatives suivantes (Qatar, 2019) :

- Introduire des lignes directrices concernant le niveau d'explicabilité et d'interprétabilité requis pour les différents types de décisions prises par les algorithmes d'IA ;
- Renforcer les lignes directrices existantes sur la confidentialité et le partage des données ;
- Développer un cadre d'éthique et de gouvernance de l'IA cohérent avec les normes du Qatar et les directives internationales.

À surveiller :

Selon GlobalData Technology (2021), le Qatar prévoit de créer un comité d'intelligence artificielle (IA) sous l'égide du ministère des transports et des communications (MoTC) du pays. Le cabinet du pays a approuvé le projet de décision établissant le comité le 3 mars 2021. Ce comité aidera à établir des mécanismes de suivi et à mettre en œuvre la stratégie nationale du Qatar en matière d'IA.

Liste des références et documents consultés :

- Chauhan, L. (2019, 7 février). *QCAI releases National AI Strategy for Qatar*. Qatar Center for Artificial Intelligence. Repéré à <https://qcai-blog.qcri.org/index.php/2019/02/07/qcai-releases-blueprint-of-national-ai-strategy-for-qatar/>
- Dickson, B. (2019, 24 juillet). *How Qatar plans to become a global leader in artificial intelligence*. TechTalks. Repéré à <https://bdtechtalks.com/2019/07/24/qatar-artificial-intelligence-strategy/>
- GlobalData Technology (2021, 12 mars). *Qatar forms artificial intelligence committee*. Verdict. Repéré à <https://www.verdict.co.uk/qatar-artificial-intelligence-committee/>
- Hukoomi (2021, 23 novembre). *Qatar's National Artificial Intelligence Strategy*. Qatar e-Government. Repéré à <https://hukoomi.gov.qa/en/article/qatars-national-artificial-intelligence-strategy>
- Qatar (2019, octobre). *National artificial intelligence strategy for Qatar*. Ministry of Transport and Communications. Repéré à <https://qcai-blog.qcri.org/wp-content/uploads/2020/04/QCRI-Artificial-Intelligence-Strategy-2019-ENG.pdf>
- Varghese, J. (2019, octobre). *Qatar launches strategy to tap AI for future*. Gulf Times. Repéré à <https://www.gulf-times.com/story/645930/Qatar-launches-strategy-to-tap-AI-for-future>

Analyse des stratégies IA - Bloc 4

Luxembourg, Pays-Bas, Portugal et Serbie

Luxembourg

Intelligence artificielle : une vision stratégique pour le Luxembourg (Mai 2019)



Faits saillants :

En mai 2019, le gouvernement du Luxembourg a publié sa stratégie nationale en matière d'IA, intitulée Intelligence artificielle : une vision stratégique pour le Luxembourg (Luxembourg, 2019), présentant les recommandations suivantes (Van Roy et al., 2021) :

- Accroître les aptitudes et les compétences dans le domaine de l'IA et offrir des possibilités d'apprentissage tout au long de la vie ;
- Soutenir la recherche et le développement de l'IA, en faisant du Luxembourg un laboratoire vivant pour l'IA appliquée ;
- Accroître les investissements publics et privés dans l'IA et les technologies connexes ;
- Favoriser l'adoption et l'utilisation de l'IA dans le secteur public ;
- Renforcer les possibilités de réseaux nationaux et internationaux et de collaborations avec des partenaires stratégiques dans le domaine de l'IA ;
- Élaborer un cadre éthique et réglementaire, en accordant une attention particulière à la réglementation en matière de vie privée et de sécurité, afin de garantir un développement transparent et fiable de l'IA ;
- Libérer le potentiel de l'économie des données comme pierre angulaire du développement de l'IA.

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

Outre les objectifs d'investissement dans un cadre amélioré propice à l'IA et de promotion du document *Ethics Guidelines for Trustworthy AI* (Code éthique pour une IA fiable) publié par le Groupe d'experts de haut niveau de la Commission Européenne sur l'intelligence

artificielle, pour garantir l'application d'orientations juridiques et éthiques visant à protéger les libertés et droits fondamentaux la stratégie prévoit la concentration sur les actions clés suivantes (Luxembourg, 2019) :

- Collaborer avec l'autorité nationale de protection des données et tirer parti de son expertise pour traiter des questions liées à l'IA ;
- Créer un comité consultatif gouvernemental spécialisé en technologies et en éthique pour débattre des conséquences éthiques des technologies et prodiguer des conseils au gouvernement sur les risques potentiels et les impacts sur la société ;
- Collaborer avec des organes clés qui travaillent au développement et à la protection de la « corporate gouvernance » au Luxembourg, pour accélérer l'adoption d'une gouvernance d'entreprise appropriée en matière d'IA ;
- Suivre activement, sous la direction de l'Institut Luxembourgeois de la Normalisation, de l'Accréditation, de la Sécurité et qualité des produits et services (ILNAS)⁹, les processus de normalisation internationaux dans le domaine de l'IA, par exemple dans le cadre du système ISO ;
- Développer et mettre en œuvre des technologies innovantes visant à renforcer la confidentialité, en vue de l'utilisation de vastes jeux de données dans le cadre de l'apprentissage en IA.

Initiatives mises en place :

Au sein de l'État, le comité interministériel AI4Gov a été créé, composé de représentants du ministère de la Digitalisation, du Service des médias et des communications (SMC), et du Service information et presse du gouvernement (SIP). L'objectif de ce comité est d'encourager les ministères et les administrations à faire usage de l'IA et de la science des données pour transformer leurs actions et leurs métiers. En d'autres termes, il s'agit d'inciter les administrations à faire usage de ces technologies innovantes pour développer de nouvelles méthodes et analyses aux problématiques existantes, d'augmenter l'efficacité du traitement des données et, ainsi, améliorer les services et les procédures de l'État (Luxembourg, s.d.).

L'initiative AI4Gov permet aussi de doter les agents de l'État de l'expérience pratique nécessaire pour maîtriser le processus d'acquisition de technologies de pointe d'IA. Le but est d'établir à terme une expertise au sein de l'État et de positionner l'Administration publique comme un moteur d'innovation (Luxembourg, s.d.).

Le comité AI4Gov assure également un accompagnement juridique et éthique à travers la création au début de 2020 d'un groupe de travail interministériel « AI Legal & Ethics Working Group ». Ce groupe est composé du ministère de la Digitalisation, du Service des

⁹ <https://www.iso.org/member/1776.html>

médias et des communications, du Service information et presse du gouvernement et du Commissariat du gouvernement à la protection des données auprès de l'Etat. À travers l'accompagnement des projets AI4Gov et le dialogue avec les porteurs de projets, ce groupe de travail réalise une évaluation préliminaire des risques juridiques et éthiques liés aux projets, posant ainsi les premiers jalons pour une approche de gouvernance de l'IA du secteur public basée sur les risques (Luxembourg, s.d.).

À surveiller :

Un groupe de coordination interministériel, sous la direction du Premier ministre, procédera à une évaluation régulière de la vision stratégique.

Liste des références et documents consultés :

Luxembourg (2019). *Artificial Intelligence: a strategic vision for Luxembourg*. The Government of the Grand Duchy of Luxembourg. Repéré à https://digital-luxembourg.public.lu/sites/default/files/2020-09/AI_FR_1_0.pdf

Luxembourg (s.d.). *L'initiative AI4Gov*. Le gouvernement luxembourgeois. Services des médias, des communications et du numérique. Dernière mise à jour : 21.04.2021. Repéré à https://smc.gouvernement.lu/fr/dossiers.gouv_digitalisation%2Bfr%2Bdossiers%2B2021%2BAI4Gov.html

OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>

Van Roy, V., Rossetti, F., Perset, K., Galindo-Romero, L. (2021). *AI Watch - National strategies on Artificial Intelligence: A European perspective, 2021 edition*. EUR 30745 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-39081-7, doi: 10.2760/069178, JRC122684. Repéré à <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684>

Pays-Bas

Plan d'action stratégique pour l'intelligence artificielle

(Octobre 2019)



Faits saillants :

En octobre 2019, le gouvernement néerlandais a lancé son plan d'action stratégique en matière d'intelligence artificielle (Pays-Bas, 2019), présentant une série d'initiatives politiques visant à renforcer la compétitivité des Pays-Bas en matière d'IA sur le marché mondial. La vision de cette stratégie repose sur trois piliers stratégiques, dont les objectifs sont les suivants (Van Roy et al., 2021) :

- Tirer parti des opportunités sociales et économiques : politiques qui encouragent l'adoption, l'utilisation et le développement de l'IA dans le secteur privé et public et qui favorisent l'utilisation de l'IA pour relever les défis sociétaux ;
- Créer les bonnes conditions : politiques soutenant l'éducation et le développement des compétences en IA ; encourager la recherche et l'innovation en matière d'IA, faciliter l'accès aux données qualitatives et améliorer l'infrastructure numérique ;
- Renforcer les fondements : y compris les actions politiques liées aux questions éthiques telles que la confiance, les droits de l'homme, la protection des consommateurs et la sécurité des citoyens.

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

Le plan d'action stratégique néerlandais présente les initiatives suivantes pour le développement et l'utilisation responsables pour l'IA (Kung, 2020) :

- Financer la recherche sur les conséquences de l'IA sur les valeurs publiques, le droit, la vie privée ;
- Mettre en place un « laboratoire de la transparence » pour les organismes gouvernementaux afin d'échanger des connaissances sur la transparence, l'explicabilité et la responsabilité ;
- Participer à des conversations entre les gouvernements sur les considérations relatives à la publication des algorithmes ;

- Travailler avec la coalition néerlandaise sur l'IA pour s'assurer que les entreprises participent à la mise en œuvre des lignes directrices européennes en matière d'éthique ;
- Financer la recherche sur l'utilisation responsable de l'IA et la transparence des algorithmes ;
- Contribuer à l'élaboration de normes sur l'IA ;
- Évaluer si les organismes de régulation sont équipés pour surveiller les algorithmes ;
- Moderniser et mieux appliquer les règles sur la protection des consommateurs et la concurrence.

Initiatives mises en place :

Aux Pays-Bas, les initiatives suivantes ont été prises pour le développement et l'utilisation responsable de l'IA :

- Plusieurs réformes législatives sont en cours pour soutenir la protection des valeurs publiques et encourager l'utilisation de l'IA dans un environnement de confiance (Van Roy et al., 2021) ;
- **Registre d'IA** d'Amsterdam (en partenariat avec la ville d'Helsinki) : Amsterdam et Helsinki ont lancé des registres d'IA ouverts qui permettent de suivre la manière dont les algorithmes sont utilisés dans les municipalités. Un livre blanc sur le registre de l'IA a été publié en septembre 2020 pour inspirer d'autres gouvernements/organisations qui veulent être transparents sur leur utilisation des algorithmes et de l'IA (Van Roy et al., 2021) ;
- Le gouvernement néerlandais a élaboré des lignes directrices pour l'application de l'analyse algorithmique des données, avec l'intention de les traduire en législation (Van Roy et al., 2021) ;
- En juillet 2021, une équipe de recherche - composée d'experts de la *Vrije Universiteit Brussel* - VUB, de l'université de Tilburg, de l'université de technologie d'Eindhoven et de l'Institut national néerlandais des droits de l'homme - a élaboré un **guide étape par étape** (Non-discrimination by design) sur la manière dont les organisations peuvent éviter la discrimination lorsqu'elles utilisent des algorithmes. L'équipe de recherche a été chargée par le ministère néerlandais de l'intérieur d'étudier les conditions techniques, juridiques et organisationnelles à prendre en considération lorsque les organisations utilisent l'intelligence artificielle dans leurs opérations. Les lignes directrices pour le développement et l'application d'algorithmes s'appliquent aux secteurs public et privé et sont disponibles en néerlandais et en anglais (Vrije Universiteit Brussel, 2021).

Liste des références et documents consultés :

- Kung, J. (2020, mai). L'ère de l'IA : rapport sur les stratégies nationales et régionales en matière d'IA. CIFAR. Deuxième Édition. Repéré à <https://cifar.ca/wp-content/uploads/2020/11/l-ere-de-l-ia-deuxieme-edition-f.pdf>
- OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>
- Pays-Bas (2019). *Strategic Action Plan for Artificial Intelligence*. Ministry of Economic Affairs and Climate Policy. Repéré à <https://www.government.nl/binaries/government/documents/reports/2019/10/09/strategic-action-plan-for-artificial-intelligence/Strategic+Action+Plan+for+Artificial+Intelligence.pdf>
- Van Roy, V., Rossetti, F., Perset, K., Galindo-Romero, L. (2021). *AI Watch - National strategies on Artificial Intelligence: A European perspective, 2021 edition*. EUR 30745 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-39081-7, doi: 10.2760/069178, JRC122684. Repéré à <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684>
- Vrije Universiteit Brussel (2021, 14 juillet). *New guidelines aim to correct discriminatory algorithms*. VUB Today. Repéré à <https://today.vub.be/en/article/new-guidelines-aim-to-correct-discriminatory-algorithms>

Portugal

IA Portugal 2030

(Juin 2019)



Faits saillants :

En juin 2019, le gouvernement portugais a présenté la stratégie nationale d'IA Portugal 2030 (Portugal, 2019) afin de définir les défis et les opportunités de l'écosystème croissant de l'IA dans le pays. Cette stratégie présente le plan visant à favoriser l'utilisation de l'IA dans le secteur public et privé au cours des prochaines années, y compris sept lignes d'action (OECD.AI, 2021) :

1. Inclusion et éducation : diffuser des connaissances généralistes sur l'IA ;
2. Qualification et spécialisation ;
3. Domaines thématiques de recherche et d'innovation dans les réseaux européens et internationaux ;
4. L'administration publique et sa modernisation ;
5. Domaines d'expertise spécifiques au Portugal ayant un impact international ;
6. Nouveaux développements et domaines de soutien dans les réseaux européens et internationaux ;
7. Relever les défis sociaux posés par l'IA : éthique et sécurité.

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

Entre autres objectifs, la stratégie portugaise vise à garantir que l'intelligence artificielle est appliquée de manière sûre et éthique dans divers domaines, y compris l'administration publique ; aider les entreprises et les régulateurs à trouver des cadres juridiques appropriés ; et renforcer les compétences et les capacités du secteur public en ce qui concerne l'IA et la science des données (Portugal, 2019). Pour atteindre ces objectifs, la stratégie prévoit un certain nombre d'initiatives, dont les suivantes (Van Roy et al., 2021) :

- Création d'un comité d'éthique pour définir des lignes directrices pour l'IA et l'automatisation ;
- Soutenir le développement d'un cadre juridique et réglementaire pour identifier les responsabilités lors de la prise de décision en matière d'IA ;

- Soutenir les entreprises et les régulateurs pour trouver des cadres juridiques appropriés.

Initiatives mises en place :

- En décembre 2020, l'Agence pour la modernisation administrative (AMA) a lancé un guide (en version bêta) pour une IA éthique, transparente et responsable (Van Roy et al., 2021). Le Guide de l'IA dans l'administration publique présente des principes, des lignes directrices et un modèle pour la préparation de projets d'intelligence artificielle éthique, transparente et responsable dans l'administration publique. Ce guide a été préparé sur la base de cinq principes de l'IA : responsabilité, transparence, explicabilité, justice et éthique. Le modèle pour l'élaboration de projets d'IA aborde les étapes, les ressources et les méthodologies pour des processus d'idéation participatifs, responsables et durables (Portugal, 2020).
- Un outil d'évaluation des risques a également été mis à disposition par AMA, permettant de faire évaluer les projets d'intelligence artificielle par toutes les parties prenantes des secteurs public et privé, quel que soit le stade du projet dans lequel il se trouve et la personne qui souhaite l'évaluer. À la fin de l'évaluation, la partie intéressée reçoit un rapport d'évaluation avec un score et des recommandations (Portugal, s.d.).

À surveiller :

- La stratégie portugaise en matière d'IA fera l'objet d'examens et de mises à jour annuels par le ministère des sciences, des technologies et de l'enseignement supérieur et le ministère de l'économie et de la transition numérique.
- Le gouvernement du Portugal prévoit de mettre à disposition une nouvelle version du GuIA - guide et outil d'évaluation de l'intelligence artificielle et de l'éthique dans l'administration publique (Portugal, 2021).

Liste des références et documents consultés :

Portugal (2019). *AI Portugal 2030: Portuguese national initiative on digital skills*. Coordination Office of INCoDe2030. Repéré à https://www.incode2030.gov.pt/sites/default/files/julho_incode_brochura.pdf

Portugal (s.d.). *Projeto GuIA Responsável*. Agência para a Modernização Administrativa (AMA). Tic.gov.pt. Repéré à <https://tic.gov.pt/pt/web/tic/guia>

Portugal (2020, décembre). *Guia para a inteligência artificial ética, transparente e responsável*. Agência para a Modernização Administrativa (AMA). Versão Beta. Repéré

à <https://tic.gov.pt/documents/37177/251624/GUIA+SHORT+VIII.pdf/2e8010a6-1013-fde7-7ef9-64783bcb6fdb>

Portugal (2021, juin). Estratégias para a transformação digital da administração pública 2021-2026. Conselho para as Tecnologias de Informação e Comunicação na Administração Pública. Repéré à

<https://bussola.gov.pt/Estratgias%20e%20Orientaes/Estrat%C3%A9gia%20de%20Transforma%C3%A7%C3%A3o%20Digital%20da%20AP%202021-2026.pdf>

OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>

Van Roy, V., Rossetti, F., Perset, K., Galindo-Romero, L. (2021). *AI Watch - National strategies on Artificial Intelligence: A European perspective, 2021 edition*. EUR 30745 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-39081-7, doi: 10.2760/069178, JRC122684. Repéré à <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684>

Serbie

Stratégie pour le développement de l'intelligence artificielle en République de Serbie pour la période 2020-2025

(Décembre 2019)



Faits saillants :

En décembre 2019, la Stratégie pour le développement de l'intelligence artificielle en République de Serbie pour la période 2020-2025 a été annoncée, définissant les objectifs et les mesures à mettre en œuvre visant la croissance économique, l'amélioration des services publics, l'avancement du personnel scientifique et le développement des compétences pour les emplois du futur (Serbie, 2019).

La stratégie présente les objectifs suivants (OECD.AI, 2021) :

- Développer une éducation adaptée aux besoins de la société moderne et de l'économie conditionnée par les progrès de l'IA ;
- Développer la science et l'innovation dans le domaine de l'IA et de sa mise en œuvre ;
- Développer l'économie basée sur l'IA ;
- Améliorer les locaux pour le développement de l'IA et des services du secteur public par la mise en œuvre de l'IA ;
- Prendre conscience de l'application éthique et sûre de l'IA.

En 4 juin 2020, le gouvernement serbe a adopté le plan d'action de la stratégie de développement de l'intelligence artificielle en République de Serbie pour la période 2020-2022, qui comprend des projets visant à soutenir le développement de l'intelligence artificielle avec de nouveaux programmes dans les domaines de l'éducation, de l'économie, des sciences et de l'administration publique (Serbie, 2020).

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

Afin de garantir que l'IA en République de Serbie soit développée et mise en œuvre en toute sécurité et conformément aux principes éthiques internationalement reconnus, les initiatives suivantes ont été envisagées dans la stratégie (Kung, 2020) :

- Élaborer et mettre en œuvre un plan pour la protection des données personnelles dans les applications d'IA, en offrant une certification, par exemple ;
- Publier des directives éthiques nationales basées sur les directives européennes ;
- Assurer la transparence, la durabilité et la responsabilité écologique du développement de l'IA ;
- Encourager le dialogue public sur les défis du développement et de la mise en œuvre de l'IA.

À surveiller :

La stratégie prévoit la publication des directives éthiques nationales basées sur les directives européennes (Serbie, 2019).

Liste des références et documents consultés :

Kung, J. (2020, mai). L'ère de l'IA : rapport sur les stratégies nationales et régionales en matière d'IA. CIFAR. Deuxième Édition. Repéré à <https://cifar.ca/wp-content/uploads/2020/11/1-ere-de-l-ia-deuxieme-edition-f.pdf>

Serbie (2019). *Strategy for the Development of Artificial Intelligence in the Republic of Serbia for the period 2020-2025*. The Government of the Republic of Serbia. Repéré à https://www.media.srbija.gov.rs/medsrp/dokumenti/strategy_artificial_intelligence.pdf

Serbie (2020, juin). *Action Plan of Strategy for Development of Artificial Intelligence adopted*. The Government of the Republic of Serbia. Repéré à <https://www.srbija.gov.rs/vest/en/156855/action-plan-of-strategy-for-development-of-artificial-intelligence-adopted.php>

OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, <https://oecd.ai>

Analyse des stratégies IA - Bloc 5

Canada

Stratégie pancanadienne en matière d'IA (2017)



Faits saillants :

En 2017, le gouvernement du Canada a chargé le CIFAR d'élaborer et de diriger la stratégie pancanadienne sur l'intelligence artificielle (CIFAR, 2020). Cette stratégie présente les objectifs suivants (CIFAR, s.d.) :

- **Attirer et retenir d'éminents chercheurs en IA** en augmentant le nombre de chercheurs d'exception et de diplômés qualifiés au Canada ;
- **Favoriser l'émergence d'un écosystème collaboratif de l'IA** en établissant des centres d'excellence scientifiques interconnectés dans les trois grands pôles canadiens de l'intelligence artificielle situés à Edmonton, à Montréal et à Toronto ;
- **Faire progresser les initiatives nationales en IA** en soutenant une communauté de recherche nationale en IA au moyen de programmes de formation, d'ateliers et d'autres possibilités de collaboration ;
- **Comprendre les incidences sociétales de l'IA** en établissant un leadership éclairé d'envergure mondiale quant aux répercussions économiques, éthiques, politiques et juridiques des percées en intelligence artificielle.

Initiatives prévues pour un fondement responsable pour l'intelligence artificielle :

La stratégie pancanadienne vise favoriser un leadership éclairé sur les répercussions éthiques de l'IA par l'entremise du programme IA et société du CIFAR (Kung, 2020).

Pour assurer une utilisation efficace et éthique de l'IA, le gouvernement présente les principes directeurs suivants (Canada, s.d.-e) :

1. **Comprendre et mesurer** l'incidence de l'utilisation de l'IA en concevant et en diffusant des outils et des approches ;
2. **Faire preuve de transparence** quant à la façon et au moment d'utiliser l'IA, en se fondant sur un besoin clair des utilisateurs et l'intérêt du public ;
3. **Fournir des explications claires** sur le processus décisionnel en matière d'IA tout en offrant des occasions d'examiner les résultats et de remettre en question les décisions ;

4. **Être le plus ouvert possible** en communiquant le code source, les données sur la formation et d'autres renseignements pertinents et ce, en protégeant les renseignements personnels, l'intégrité du système, ainsi que la sécurité et la défense nationales ;
5. **Offrir une formation adéquate** pour que les employés du gouvernement qui conçoivent et utilisent des solutions liées à l'IA aient les compétences nécessaires en matière de conception, de fonctionnement et de mise en œuvre responsables pour améliorer les services gouvernementaux fondés sur l'IA.

Initiatives mises en place :

- En 2019, le gouvernement du Canada a fait appel à la recherche participative de centaines d'experts en informatique et de fonctionnaires pour élaborer le livre blanc « L'intelligence artificielle responsable au sein du gouvernement du Canada », et a publié le 1er avril 2019 sa « Directive sur la prise de décision automatisée » (OECD.AI, 2021). Cette directive a pour objet de veiller à ce que les systèmes décisionnels automatisés soient déployés d'une manière qui permet de réduire les risques pour les Canadiens et les institutions fédérales, et qui donne lieu à une prise de décisions plus efficace, exacte et conforme, qui peut être interprétée en vertu du droit canadien (Canada, s.d.-b). Elle prévoit une approche fondée sur le risque pour assurer la transparence, la responsabilité, la légalité et l'équité des décisions automatisées qui touchent les Canadiens. La directive ne s'applique qu'aux systèmes automatisés de prise de décision destinés au public, tels que les programmes de prestations qui déterminent si les demandeurs remplissent les conditions requises. Elle ne couvre pas encore les services gouvernementaux internes ou les questions de sécurité nationale (OECD.AI, 2021).
- En 2019, le gouvernement du Canada a mis au point l'outil d'évaluation de l'incidence algorithmique (ÉIA) pour aider les décideurs et autres responsables à déterminer comment sélectionner et utiliser de manière responsable l'IA au sein du gouvernement, en identifiant les biais éthiques des moteurs d'IA et en aidant à les atténuer (OECD.AI, 2021). L'ÉIA est un outil d'évaluation des risques obligatoire destiné à soutenir la *Directive sur la prise de décisions automatisée* du Conseil du Trésor. L'outil est un questionnaire qui détermine le niveau d'incidence d'un système de décision automatisé. Il est composé de 48 questions sur les risques et de 33 questions sur l'atténuation. Les notes d'évaluation sont fondées sur de nombreux facteurs, y compris la conception des systèmes, l'algorithme, le type de décision, l'incidence et les données (Canada, s.d.-c).
- En 2019, le Conseil consultatif en matière d'intelligence artificielle a été créé avec l'objectif de conseiller le gouvernement canadien sur la façon de renforcer les forces et le leadership mondial du Canada en matière d'IA, de déterminer les possibilités de

créer une croissance économique qui profite à tous les Canadiens et de veiller à ce que les progrès de l'IA reflètent les valeurs canadiennes (OECD.AI, 2021).

- En 2020, le Conseil consultatif a mis sur pied le « Groupe de travail sur la sensibilisation du public ». Ce groupe étudie des mécanismes visant à renforcer la sensibilisation du public et à favoriser la confiance dans l'IA. Il vise également à fonder le débat canadien sur une compréhension mesurée de la technologie de l'IA, de ses utilisations potentielles et des risques qui y sont associés (Canada, s.d.-a). En partenariat avec l'Institut canadien des recherches avancées (ICRA) et Algora Lab (Université de Montréal), ce groupe a dirigé des ateliers virtuels (« Dialogue ouvert : L'intelligence artificielle au Canada ») dans l'ensemble du Canada pour entamer des discussions avec les membres du public au sujet de leurs perceptions de l'intelligence artificielle (Canada, s.d.-d).

Liste des références et documents consultés :

Canada (s.d.-a). Conseil consultatif en matière d'intelligence artificielle. Repéré à <https://ised-isde.canada.ca/site/conseil-consultatif-intelligence-artificielle/fr>

Canada (s.d.-b). Directive sur la prise de décisions automatisée. Repéré à <https://www.tbs-sct.gc.ca/pol/doc-fra.aspx?id=32592#cha6>

Canada (s.d.-c). Outil d'évaluation de l'incidence algorithmique. Repéré à <https://www.canada.ca/fr/gouvernement/systeme/gouvernement-numerique/innovations-gouvernementales-numeriques/utilisation-responsable-ai/evaluation-incidence-algorithmique.html>

Canada (s.d.-d). Groupe de travail sur la sensibilisation du public. Repéré à <https://ised-isde.canada.ca/site/conseil-consultatif-intelligence-artificielle/fr/groupe-travail-sensibilisation-public>

Canada (s.d.-e). Utilisation responsable de l'intelligence artificielle (IA). Nos principes directeurs. Repéré à <https://www.canada.ca/fr/gouvernement/systeme/gouvernement-numerique/innovations-gouvernementales-numeriques/utilisation-responsable-ai.html#toc1>

CIFAR (s.d.). Stratégie pancanadienne en matière d'IA. Repéré à <https://cifar.ca/fr/ia/>

CIFAR (2020, octobre). *Pan-canadian AI Strategy Impact Assessment Report*. Repéré à <https://cifar.ca/wp-content/uploads/2020/11/Pan-Canadian-AI-Strategy-Impact-Assessment-Report.pdf>

Kung, J. (2020, mai). L'ère de l'IA : rapport sur les stratégies nationales et régionales en matière d'IA. CIFAR. Deuxième Édition. Repéré à <https://cifar.ca/wp-content/uploads/2020/11/l-ere-de-l-ia-deuxieme-edition-f.pdf>

OECD.AI (2021), alimenté par EC/OECD (2021), Base de données STIP Compass, consultée le 2/11/2021, Repéré à <https://oecd.ai>

Fiche technique

Titre : La Déclaration de Montréal pour un développement responsable de l'intelligence artificielle

Sous-titre : non

Date de publication : 2018

Pays (ou Province) : Montréal

Langue : Français et traductions disponibles en 6 autres langues

Organisation ou signataires : 200 organismes et 2386 citoyens (jusqu'à 31 janvier 2022)

Nombre de pages : 321 (rapport)

Résumé : oui (21)

Principes généraux bien identifiés : oui (bien-être, respect de l'autonomie, protection de l'intimité et de la vie privée, solidarité, participation démocratique, équité, inclusion de la diversité, prudence, responsabilité e développement soutenable)

Recommandations bien identifiées : oui (8) – Organisme indépendant de veille et de consultation citoyenne ; Politique d'audit et de certification des SIA ; Encapacitation et autonomisation ; Formations en éthique ; Favoriser un développement inclusif de l'IA ; Protéger la démocratie des manipulations politiques de l'information ; Développement international de l'IA ; Empreinte environnementale directe et induite des SIA.

Thèmes principaux : les enjeux éthiques de l'IA, les applications par secteur (éducation, santé, ville intelligente, justice et marché du travail)

Notes : Une initiative de l'Université de Montréal. La rédaction de la Déclaration de Montréal est le fruit du travail d'une équipe scientifique pluridisciplinaires et interuniversitaire qui s'appuie sur un processus de consultation citoyenne et sur la concertation avec des experts et des parties prenantes du développement de l'IA.

Lien : <https://www.declarationmontreal-iaresponsable.com/>

Organismes interpellés :

Ministères : Culture et Communications ; Éducation ; Enseignement supérieur ; Immigration, francisation et intégration ; Travail, emploi et solidarité sociale ; Économie et Innovation ; Justice ; Relations internationales et Francophonie ; Énergie et Ressources naturelles ; Environnement et Lutte contre les changements climatiques.

Organismes : Commission de l'éthique en science et en technologie ; Commission d'accès à l'information du Québec.

Fiche technique

Titre : Recommandation sur l'éthique de l'intelligence artificielle

Sous-titre : non

Date de publication : 24 novembre 2021 (date de l'adoption par la Conférence générale de l'UNESCO lors de sa 41^e session)

Pays (ou Province) : Paris

Langue : Français et autres 5 langues (original anglais)

Organisation ou signataires : UNESCO - Groupe d'experts ad hoc (GEAH)

Nombre de pages : 39

Résumé : non

Principes généraux bien identifiés : oui (proportionnalité et innocuité ; sûreté et sécurité ; équité et non-discrimination ; durabilité ; droit au respect de la vie privée et protection des données ; surveillance et décision humaines ; transparence et explicabilité ; responsabilité et redevabilité ; sensibilisation et éducation ; gouvernance et collaboration multipartites et adaptatives)

Recommandations bien identifiées : oui (orientations éthiques axées sur onze domaines stratégiques : évaluations de l'impact éthique ; gouvernance et gestion éthiques ; politiques en matière de données ; développement et coopération internationale ; environnement et écosystèmes ; égalité des genres ; culture ; éducation et recherche ; communication et information ; économie et travail ; santé et bien-être social).

Thèmes principaux : éducation, science, culture et communication et information.

Notes : La Recommandation cherche à « offrir un cadre universel de valeurs, de principes et d'actions pour guider les États dans la formulation de leur législation, de leurs politiques ou d'autres instruments concernant l'IA ».

Lien : https://unesdoc.unesco.org/ark:/48223/pf0000379920_fre.page=15

Fiche technique

Titre : Lignes directrices en matière d'éthique pour une IA digne de confiance

Sous-titre : non

Date de publication : 8 avril 2019

Pays (ou Province) : Commission européenne

Langue : Français et autres 22 langues

Organisation ou signataires : Groupe d'experts de haut niveau sur l'intelligence artificielle (Commission européenne)

Nombre de pages : 56

Résumé : oui (3)

Principes généraux bien identifiés : oui (respect de l'autonomie humaine, prévention de toute atteinte, équité et explicabilité)

Recommandations bien identifiées : oui (7 orientations liées à 7 exigences : action humaine et contrôle humain ; robustesse technique et sécurité ; respect de la vie privée et gouvernance des données ; transparence ; diversité, non-discrimination et équité ; bien-être sociétal et environnemental ; responsabilité)

Thèmes principaux : Présentation d'une liste de principes éthiques et de lignes directrices qui établissent un cadre pour parvenir à la réalisation d'une IA digne de confiance.

Notes : Le document présente les orientations sous la forme de trois niveaux d'abstraction. Le chapitre I établit les fondements d'une IA digne de confiance (il recense et décrit les principes éthiques) ; le chapitre II traduit ces principes éthiques en sept exigences et propose des méthodes tant techniques (l'architecture du système d'IA, éthique et état de droit dès la conception, méthodes d'explication, essais et validations, qualité des indicateurs de service) que non techniques (comme réglementation, codes de conduite, normalisation, certification, cadres de gouvernance internes et externes, éducation et sensibilisation, participations des parties prenantes et dialogue social, diversité et équipes de conception inclusives) pour mettre en œuvre les exigences ; le chapitre III établit une liste d'évaluation préliminaire et non exhaustive pour une IA digne de confiance afin de concrétiser les exigences du chapitre II. La dernière section expose des possibilités bénéfiques et des préoccupations importantes suscitées par les systèmes d'IA.

Lien : <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>