



JUNE 2022



ENVIRONMENTAL
LAW • INSTITUTE

Artificial Intelligence Advancing Environmental Compliance, Enforcement & Follow-Up Programs



Impact Assessment
Agency of Canada

Agence d'évaluation
d'impact du Canada

Canada

REMERCIEMENTS

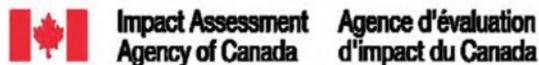
Le présent document a été préparé par le laboratoire d'innovation de l'Environmental Law Institute (ELI). Le laboratoire d'innovation de l'ELI se concentre sur les effets sur l'environnement et les occasions découlant des technologies émergentes et de leurs systèmes d'innovation sous-jacents. Le présent rapport a été rédigé par Heather Luedke, avec la contribution de Dave Rejeski, Margaret Badding, Zhuoshi Liu, Kasantha Moodley, Simone Baldwin et Crystal Zhao. L'équipe du projet tient à remercier les responsables gouvernementaux et les concepteurs de technologies pour le temps qu'ils ont consacré à ce rapport ainsi que leur contribution. Elle remercie également l'Agence d'évaluation d'impact du Canada (AEIC) pour son soutien financier à ce travail et pour ses commentaires au cours du projet. Le contenu de ce rapport, y compris toute erreur ou omission, relève de la seule responsabilité de l'ELI. Les auteurs vous invitent à apporter des corrections et des ajouts au besoin.

À PROPOS DE L'ENVIRONMENTAL LAW INSTITUTE



L'Environmental Law Institute met le droit au service des personnes, des endroits et de la planète. Depuis 1969, l'ELI a joué un rôle central pour façonner les domaines du droit, de la politique et de la gestion de l'environnement, à l'échelle nationale et internationale. Aujourd'hui, à l'aube de son sixième anniversaire, l'ELI est un centre d'édition, de recherche et d'éducation non partisan et reconnu sur la scène internationale, qui cherche à renforcer la protection de l'environnement en améliorant le droit et la gouvernance dans le monde entier. Consultez le site www.eli.org.

À PROPOS DE L'AGENCE D'ÉVALUATION D'IMPACT DU CANADA



L'Agence est un organisme fédéral relevant du ministre de l'Environnement et du Changement climatique qui réalise des évaluations d'impact de grande qualité pour examiner les répercussions positives et négatives de projets potentiels sur l'environnement, l'économie, la société et la santé. L'Agence contribue à une prise de décisions éclairées sur les grands projets en faveur du développement durable au Canada. Consultez le site <https://www.canada.ca/fr/agence-evaluation-impact.html>.

Photo de couverture sous licence CC0 1.0 universel.

Table des matières

REMERCIEMENTS	2
À PROPOS DE L'ENVIRONMENTAL LAW INSTITUTE	2
À PROPOS DE L'AGENCE D'ÉVALUATION D'IMPACT DU CANADA	2
1 INTRODUCTION	5
2 CONTEXTE	6
3 PORTÉE DU PROJET DE RECHERCHE	9
4 APPLICATIONS DE L'IA	10
4.1 IA de détection des changements par satellite	12
Défi	12
Modèle d'IA pour la détection des changements	12
Résultats	14
Ressources mobilisées	14
Facteurs de réussite	14
Obstacles	15
Possibilités pour l'Agence	15
4.2 IA prédictive pour le ciblage des inspections	16
Défi	16
Modèle d'IA pour l'analyse prédictive	17
Résultats	17
Ressources mobilisées	18
Facteurs de réussite	18
Obstacles	18
Possibilités pour l'Agence	19
4.3 IA de classification supervisée pour le traitement des plaintes	20
Défi	20
Modèle de classification supervisée	20
Ressources mobilisées	23
Facteurs de réussite	23
Obstacles	23
Possibilités pour l'Agence	23

4.4 IA prédictive pour le ciblage des inspections	24
Défi	25
Modèle d'IA pour l'analyse prédictive	25
Résultats	25
Ressources mobilisées	26
Facteurs de réussite	26
Obstacles	26
Possibilités pour l'Agence	26
4.5 Traitement intelligent des documents pour l'évaluation de la conformité	28
Défi	28
Traitement intelligent des documents pour l'évaluation de la conformité	28
Résultats	29
Ressources mobilisées	29
Facteurs de réussite	29
Obstacles	29
Possibilités pour l'Agence	30
4.6 Résumé des cas d'utilisation	30
4.7 EnviroVerse – L'avenir de la surveillance réglementaire	33
5 CRITÈRES D'ADOPTION	34
6 PROCHAINES ÉTAPES	38
7 RÉFÉRENCES	40

1 INTRODUCTION

L'intelligence artificielle (IA) désigne de manière générale les systèmes informatiques qui perçoivent, apprennent, abstraient et raisonnent pour traiter l'information. Alors que l'IA continue de progresser et de s'étendre à de nouveaux domaines, l'intérêt et les préoccupations à l'égard des applications de cette technologie ne cessent de croître. Le secteur public tient compte de ces tendances. Une enquête menée en 2021 auprès de 500 dirigeants gouvernementaux dans le monde entier révèle qu'une écrasante majorité d'organismes fédéraux, étatiques et locaux considèrent l'IA comme un facteur qui contribuera grandement à l'atteinte des résultats des missions dans les cinq prochaines années [Van Buren, 2021]. La même étude rapporte un écart généralisé entre l'état actuel et l'état souhaité des capacités en matière d'IA parmi les organismes observés. Cet écart est dû à une variété de préoccupations, notamment le manque de compétences, l'attitude des établissements envers l'IA, la capacité technologique et les considérations éthiques. Même en reconnaissant ces limites, l'IA offre des possibilités de résoudre d'importants problèmes de gouvernance et de gestion pour le secteur public.

Il existe une vaste gamme d'applications de l'IA dans le secteur public, et ces applications continuent de s'étendre. Les organismes gouvernementaux du monde entier utilisent l'IA pour détecter les fraudes, répondre aux demandes du public, verser les prestations d'aide sociale, statuer sur les audiences sur la libération sous caution, trier les cas de soins de santé, et plus encore [Deloitte AI Institute, 2021]. Dans le domaine de la gouvernance environnementale, l'IA améliore les prévisions climatiques, aide les villes à renforcer leur résilience et facilite la planification urbaine durable. Les gouvernements et les entreprises du monde entier s'appuient sur l'IA pour atteindre leurs objectifs en matière d'atténuation des changements climatiques et de protection de l'environnement.

Les activités de conformité, d'application de la loi et de suivi en matière d'environnement font partie de ce large éventail des applications environnementales possibles de l'IA. La conformité environnementale fait référence à la promotion et à la vérification des mesures d'atténuation, telles qu'elles sont dictées par la réglementation et les organismes environnementaux. Parmi ces mesures d'atténuation peuvent figurer des séances d'information, des inspections sur place et hors site, des enquêtes et la coordination avec d'autres autorités. L'application de la loi englobe toutes les mesures prises pour rétablir la conformité après qu'un cas de non-conformité a été relevé. Enfin, les activités de suivi font référence aux processus permettant de vérifier l'exactitude de l'évaluation d'impact d'un projet désigné et de déterminer l'efficacité des mesures d'atténuation pour réduire les effets négatifs potentiels d'un projet. Les organismes environnementaux chargés de prendre ces mesures ont souvent des ressources financières, humaines et techniques très limitées. Une grande quantité de rapports et un nombre insuffisant d'inspecteurs font partie des problèmes courants auxquels se heurtent ces organisations. Nombreux sont ceux qui cherchent maintenant à savoir comment la technologie de l'IA peut optimiser les ressources disponibles et améliorer leur travail.

L'ELI a été mandaté par l'Agence d'évaluation d'impact du Canada (AEIC) pour définir les utilisations à forte incidence de l'IA qui renforceront les activités postérieures aux décisions. Ces utilisations devraient guider l'AEIC et la préparer à ces activités. Les applications de l'IA abordées plus bas ont été jugées pertinentes pour l'AEIC. Cette recherche exploratoire a pour but d'orienter l'AEIC en ce qui a trait à une éventuelle adoption des technologies de l'IA et de l'apprentissage machine.

Bien que le présent rapport ait été élaboré pour les activités postérieures aux décisions de l'AEIC, les conclusions qui s'y trouvent peuvent s'appliquer à d'autres fonctions de l'AEIC et à d'autres organismes se trouvant à un stade semblable d'étude et de développement de l'IA. Nous espérons que ces travaux de recherche aideront les gestionnaires de l'environnement dans leurs efforts visant à intégrer l'innovation de l'IA dans les activités essentielles de conformité, d'application de la loi et de suivi en matière d'environnement.

2 CONTEXTE

2.1 AEIC

L'Agence d'évaluation d'impact du Canada (AEIC), connue sous le nom d'Agence canadienne d'évaluation environnementale avant 2019, est un organisme fédéral qui relève du ministre de l'Environnement et du Changement climatique. Son mandat principal est de mener le processus d'évaluation d'impact. En plus de diriger et de gérer le processus d'évaluation d'impact de tous les grands projets désignés par le gouvernement fédéral, l'AEIC sert de principal point de consultation et fait participer les peuples autochtones et d'autres intervenants aux étapes appropriées du processus. L'AEIC est également responsable des activités postérieures aux décisions, notamment le respect des conditions énoncées dans les déclarations de décision des projets, la surveillance des programmes de suivi et la production de rapports sur ceux-ci. Le présent rapport constitue une étape exploratoire dans les plans de l'AEIC visant à déterminer quelles activités postérieures aux décisions pourraient être automatisées.

Avant d'aborder d'éventuels domaines d'intervention de l'IA, il est nécessaire de définir l'IA en contexte avec quelques autres termes (autonome, apprentissage machine et algorithme). Il n'y a pas de consensus parmi les experts sur les limites de l'intelligence artificielle. Aux fins du présent rapport, l'« intelligence artificielle » est reconnue comme étant « des machines qui répondent à une stimulation compatible avec les réponses types des humains, compte tenu de la capacité humaine de contemplation, de jugement et d'intention », comme le soutiennent les chercheurs Shubhendu et Vijay [Shubhendu, 2013]. En d'autres termes, l'IA est un système informatique capable de prendre des décisions qui nécessiteraient normalement une expertise humaine. Les systèmes d'IA possèdent généralement la capacité de percevoir, d'apprendre, d'abstraire et de raisonner. Le terme « autonome » fait référence à la capacité de prendre des décisions sans intervention humaine. L'« apprentissage machine » est une application de l'IA qui permet aux systèmes d'apprendre et de s'améliorer à partir de données sans supervision humaine. Enfin, le terme « algorithme » désigne l'ensemble des règles ou instructions qui dictent le comportement de fonctionnement d'une IA.

2.2 DOMAINES D'ÉTUDE

L'AEIC commence maintenant à explorer les possibilités d'intégrer l'IA à ses activités postérieures aux décisions. Le cadre des activités postérieures aux décisions de l'AEIC révèle de nombreux domaines où l'IA peut intervenir (figure 1). Plus particulièrement, la vérification de la conformité, les mesures d'application de la loi et les activités de suivi offrent un potentiel d'amélioration par l'intégration de l'IA. La vérification de la conformité relève habituellement d'agents d'application de la loi et d'analystes qui effectuent des inspections sur place et hors site, enquêtent sur les contraventions présumées et assurent la coordination avec les autorités gouvernementales. Des études de cas révèlent qu'une combinaison d'imagerie satellitaire, d'analyse prédictive, d'évaluation des risques et d'algorithmes de détection des changements pourrait être efficace pour optimiser l'efficacité de la détermination des infractions.

Les mesures d'application de la loi consistent en l'émission, par les agents d'application de la loi et les analystes, d'avertissements, d'avis et d'ordonnances visant la prise de mesures correctives. Les demandes d'injonctions, de poursuites et de sanctions entrent également dans cette catégorie. Les activités de suivi, menées par l'équipe de suivi, comprennent généralement l'examen des plans et des rapports de suivi, l'analyse de l'efficacité des mesures d'atténuation, l'évaluation de l'exactitude des prédictions, l'élaboration des rapports de suivi de l'Agence, la mobilisation des intervenants et la détermination des améliorations à apporter au processus. Les mesures d'application de la loi et les activités de suivi génèrent beaucoup de documents; il faut donc de nombreuses heures de travail humain pour traiter les documents nécessaires. Des études de cas révèlent que le traitement intelligent des documents pourrait accroître l'efficacité des agents d'application de la loi et de l'équipe de suivi dans le cadre de ces tâches.

La promotion de la conformité par l'éducation et la formation constitue la dernière grande catégorie d'activités postérieures aux décisions. Le présent rapport se concentre sur les cas d'utilisation relatifs aux trois autres domaines, car l'AEIC juge ces domaines prioritaires.

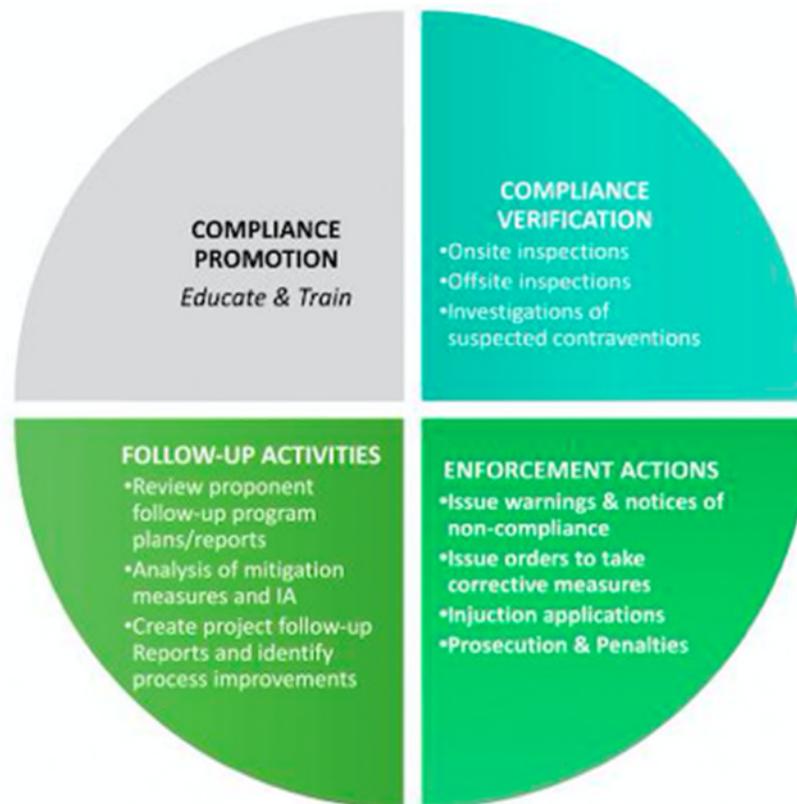


Figure 1 : Activités pos de l'AEIC

2.3 OBSTACLES ET DÉFIS

Même si l'on reconnaît le potentiel que présente l'AEIC pour une intégration efficace de l'IA, il est important de tenir compte des défis et des obstacles qui nuiront probablement à son adoption. D'abord, les systèmes d'IA ont besoin de grands volumes de données pour fonctionner efficacement. Comme l'AEIC n'a commencé à participer aux activités postérieures aux décisions qu'en 2012 et qu'elle ne dispose que de quantités limitées de données historiques, la disponibilité des données peut entraver ses efforts d'automatisation. Le manque d'expérience organisationnelle de l'AEIC en matière d'IA peut également poser problème, car elle ne possède probablement pas les compétences et les capacités technologiques requises pour exploiter des systèmes d'IA, et devra embaucher et former un grand nombre de nouveaux employés ou consulter des experts externes. Le financement peut aussi constituer un défi pour l'AEIC, qui devra s'efforcer d'obtenir un financement continu pour l'IA afin d'entretenir et d'exploiter les systèmes d'IA et de garder en poste les experts requis.

L'AEIC peut également se heurter à des problèmes d'interopérabilité des systèmes. L'intérêt de l'AEIC pour l'IA arrive à un moment où le gouvernement du Canada encourage l'adoption et l'exploration de l'IA dans de nombreux organismes. À mesure que d'autres organismes adoptent des solutions d'IA, il devient de plus en plus nécessaire que les applications soient conçues pour fonctionner en conjonction avec les systèmes existants et futurs. Cela est particulièrement vrai pour l'AEIC, dont le mandat requiert qu'elle collabore avec d'autres organismes à de nombreux moments du processus d'évaluation d'impact. Enfin,

l'AEIC devrait sérieusement tenir compte de l'éthique lorsqu'il s'agit de créer et de mettre en œuvre des solutions d'IA. Le gouvernement du Canada a élaboré une évaluation de l'incidence algorithmique (EIA) pour déterminer le caractère éthique des applications de l'IA utilisées en gouvernance, mais l'expérience passée révèle que cette mesure ne protège pas totalement les technologies de l'IA contre la controverse publique. La tentative d'Immigration, Réfugiés et Citoyenneté Canada d'avoir recours à des algorithmes de contrôle de l'immigration a fait l'objet d'une couverture médiatique critique, même si ces algorithmes avaient réussi l'EIA [Nalbandian, 2021]. L'AEIC devra tenir compte de ces problèmes et les résoudre afin de maximiser l'incidence de ses éventuelles solutions d'IA.

3 PORTÉE DU PROJET DE RECHERCHE

Le présent rapport est le fruit d'un processus de recherche de six mois. L'ELI a commencé par communiquer avec l'équipe de l'AEIC dans le but de confirmer les objectifs de recherche et d'acquérir des connaissances sur le contexte organisationnel en ce qui concerne l'adéquation de l'IA. Les mois suivants ont été consacrés à la collecte de renseignements et à la recherche, principalement par l'intermédiaire d'entrevues et de l'examen de la documentation existante.

3.1 ENTREVUES

L'ELI a mené des entrevues avec des membres de la communauté de pratique de l'International Network of Environmental Compliance and Enforcement (INECE) sur l'utilisation de l'analyse des données et de l'intelligence artificielle pour la conformité et l'application de la loi en environnement, des membres de l'INECE qui ont répondu à une demande de renseignements dans l'infolettre de l'INECE et d'autres personnes utilisant l'IA dans leurs organismes respectifs. Au total, sept entrevues ont été réalisées avec des professionnels de la gouvernance environnementale du monde entier : Jed Anderson d'EnviroAI, Michael Enns d'Environnement et Changement climatique Canada, Justin Budgell de Santé Canada, Cristobal de la Maza Guzmán de la Surintendance de l'environnement (Chili), Randy Hill de l'Office of Enforcement and Compliance Assurance de l'Environmental Protection Agency (EPA) des États-Unis, Ray Purdy d'Air & Space Evidence au nom du ministre de l'Environnement (Nouvelle-Zélande) et Paul Stevens de la Game Management Authority de Victoria (Australie). Les personnes interrogées ont été invitées à s'exprimer sur l'état d'avancement de l'adoption de l'IA au sein de leur organisme, la manière dont elle contribue à la réalisation de leurs objectifs, les responsables de l'exploitation du système, les résultats, les facteurs essentiels de réussite, les obstacles et le coût de l'adoption. Ces entrevues ont été déterminantes pour éclairer les cas d'utilisation présentés dans le rapport.

3.2 ANALYSE DOCUMENTAIRE

Une analyse documentaire a également été réalisée dans le cadre du processus de recherche. Les principaux sujets d'intérêt comprenaient l'IA au sein du gouvernement, les cas d'utilisation dans un contexte environnemental et hors de celui-ci, les considérations en matière d'adoption, la responsabilité et le traitement des documents. Cette étape a permis de générer les connaissances de base de l'ELI sur

le sujet. Plus précisément, l'analyse documentaire a porté sur l'IA dans le secteur public et sur l'adoption de l'IA par les administrations publiques. Les publications s'accordent largement sur l'utilité de l'IA pour accroître l'efficacité des organismes et libérer du temps pour les employés afin qu'ils travaillent sur des tâches qui ne peuvent être automatisées [Martinho-Truswell, 2018]. Cependant, les inquiétudes des employés concernant la perte d'emplois liée à l'adoption de l'IA constituent un obstacle à sa mise en œuvre. Ainsi, des mesures comme la mobilisation des intervenants et la sensibilisation sont essentielles pour aider les organismes à intégrer l'IA dans leur travail d'une manière qui suscite l'adhésion de tous et mène au succès [Chenok, 2018]. La documentation se concentre également sur l'importance d'embaucher des spécialistes internes pour les projets d'IA, relevant que les scientifiques des données sont très demandés et que les organismes doivent trouver comment offrir des occasions d'emploi attrayantes et retenir le personnel [GoDataDriven, 2019].

Par ailleurs, la documentation met en évidence un autre défi lié à la conception et à la mise en œuvre de l'IA : les questions de gouvernance, de légalité et d'éthique. Il est impératif que les organismes soient responsables de leurs systèmes d'IA, qu'ils soient transparents, qu'ils sensibilisent les intervenants concernés et intègrent leurs commentaires, et qu'ils utilisent un plan de gestion des risques pour atténuer les risques et les biais associés aux systèmes d'IA. Ils doivent également s'assurer qu'ils agissent dans les limites de la loi [Government Accountability Office des États-Unis, 2021]. D'autres documents indiquent qu'une stratégie ou une feuille de route répondant à ces préoccupations, et à d'autres, peut être utile pour un organisme qui commence à s'intéresser à l'IA [Van Buren, 2021; GoDataDriven, 2019].

Malgré une documentation imposante sur l'IA dans le secteur public et dans d'autres industries, comme les soins de santé, l'utilisation de l'IA pour faciliter les activités de conformité, d'application de la loi et de suivi en matière d'environnement est relativement nouvelle. De nombreux projets sont encore en phase pilote ou d'essai. Ainsi, on observe un manque de résultats fiables et de rapports sur l'intégration de l'IA aux activités de conformité et d'application de la loi en environnement et de suivi. Le présent rapport vise à fournir plus de renseignements sur les projets d'IA existants qui sont pertinents aux activités de conformité et d'application de la loi en environnement et de suivi. Après avoir rassemblé des cas d'utilisation potentiels à partir des entrevues et de la documentation existante, l'ELI a choisi les cas les plus pertinents pour les activités de l'AEIC, qui sont également pertinents pour d'autres organismes environnementaux. Ces applications sont abordées ci-dessous.

4 APPLICATIONS DE L'IA

Cette section présente cinq cas d'utilisation de l'IA dans les contextes canadien et international. Ces études de cas ont été choisies pour leur pertinence pour l'AEIC et tiennent compte des défis qui lui sont propres. L'AEIC reçoit un nombre croissant de projets à évaluer en vertu de la *Loi canadienne sur l'évaluation environnementale* (LEI), ce qui entraînera une augmentation régulière du nombre de projets qu'elle supervisera dans la phase postérieure aux décisions. Par conséquent, les modèles d'IA qui peuvent contribuer aux processus de conformité, d'application de la loi et de suivi présentent un intérêt

particulier. Il s'agit notamment de modèles destinés à faciliter la détection des cas de non-conformité, l'évaluation des risques, le traitement et la rédaction des rapports réglementaires, le traitement des plaintes et la détermination des tendances en vue d'un suivi. Une liste complète des études de cas est présentée ci-dessous.

Type de modèle d'IA	Intention	Autorités gouvernementales
IA de détection des changements par satellite	Déterminer les sites de déchets inconnus et non autorisés	Ministère de l'Environnement, Nouvelle-Zélande
IA prédictive pour le ciblage des inspections	Déterminer la non-conformité des installations de traitement des déchets dangereux	Environmental Protection Agency des États-Unis
IA de classification supervisée pour le traitement des plaintes	Prévoir la pertinence et la gravité des plaintes déposées	Surintendance de l'environnement, Chili
IA prédictive pour le ciblage des inspections	Prévoir les cas de non-conformité et les risques dans les installations réglementées	Environnement et Changement climatique Canada
Traitement intelligent des documents pour l'évaluation de la conformité	Attribuer des cotes de risque aux cas de non-conformité et aider à la rédaction des rapports	Santé Canada

Tous ces cas en sont aux premiers stades de conception ou de mise en œuvre, et tous ont des projets pilotes en cours. Bien que ces technologies d'IA puissent faire l'objet d'applications plus avancées dans d'autres secteurs, ces cas particuliers ont été choisis en raison de leur pertinence directe pour l'AEIC en tant qu'organisme participant aux activités d'évaluation d'impact postérieures aux décisions.

Bien que ces cas aient été choisis en fonction de l'AEIC, ils fournissent également un échantillon des possibilités d'IA qui pourraient être mises en œuvre dans les organismes de conformité et d'application de la loi en environnement à l'échelle mondiale. Les types de projets, ainsi que les renseignements sur leurs facteurs de réussite, les obstacles et les ressources mobilisées, pourraient aider de nombreux organismes environnementaux à élaborer et à mettre en œuvre des projets d'IA.

Le format de chaque étude de cas est le suivant : l'étude de cas commence par une description du défi auquel l'organisme doit faire face, puis décrit le modèle d'IA particulier qui a été élaboré pour relever ce défi. La description du modèle comprend des renseignements sur les données utilisées, le processus de conception, le statut actuel et l'utilisation. Les résultats, les ressources mobilisées, les facteurs de réussite et les obstacles sont ensuite présentés. Chaque étude de cas se termine par une section sur la pertinence du modèle pour l'AEIC.

4.1 IA de détection des changements par satellite

Autorité :	Ministère de l'Environnement, Nouvelle-Zélande
Collaborateur :	Air & Space Evidence
Intention :	Déterminer les sites de déchets inconnus et non autorisés
Personne-ressource :	Shaun Lewis, Waste and Resource Efficiency Division (division de la gestion efficace des déchets et des ressources), ministère de l'Environnement, Nouvelle-Zélande

Défi

En 2018, la Banque mondiale a classé la Nouvelle-Zélande au dixième rang des pays les plus gaspilleurs en matière de déchets produits par habitant. Environ 3,6 kg de déchets étaient produits par personne chaque jour à l'époque. Ce chiffre est cinq fois supérieur à la moyenne mondiale quotidienne de 0,65 kg par personne. Au cours des dix dernières années, la Nouvelle-Zélande a envoyé plus de 30 millions de tonnes de déchets dans les sites d'enfouissement. Ce volume devrait augmenter, notamment dans le secteur de la construction et de la démolition. Pour freiner cette tendance, le ministère de l'Environnement de la Nouvelle-Zélande (« le Ministère ») propose d'augmenter la taxe nationale d'élimination des déchets pour la faire passer de 10 \$ NZ à 60 \$ NZ par tonne (de 6 \$ US à 37 \$ US) sur une période de quatre ans, et d'étendre son application. Pour mettre en œuvre la taxe, le Ministère crée actuellement un registre de tous les sites d'enfouissement opérationnels, historiques et proposés, avec l'intention d'enregistrer les sites d'enfouissement autorisés et non autorisés. Le Ministère a eu de la difficulté à déterminer les sites inconnus ou non autorisés.

Modèle d'IA pour la détection des changements

En 2021, le Ministère a chargé Air & Space Evidence (ASE), une entreprise dérivée de l'University College de Londres, d'aider à repérer les sites de déchets pour le registre. ASE est spécialisée dans la détection de la non-conformité réglementaire au moyen de données satellitaires. Plus précisément, ASE se concentre sur les infractions liées aux déchets et utilise l'intelligence artificielle dans un modèle de détection des changements par satellite pour repérer les sites de déchets illégaux et ceux qui s'étendent au-delà des limites autorisées.

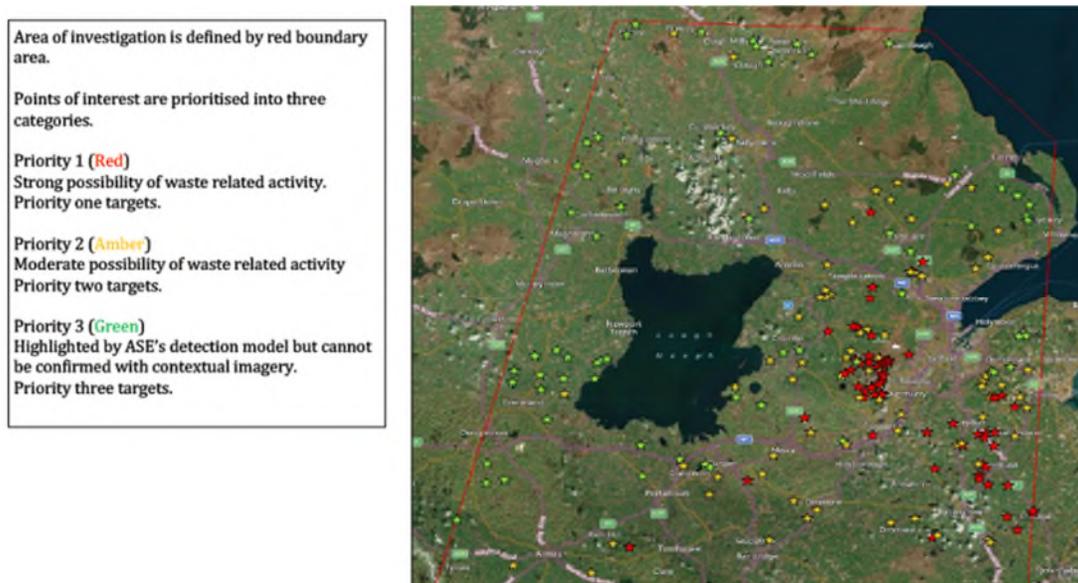
Le modèle de détection des changements par satellite d'ASE comprend un processus en deux étapes pour repérer et caractériser les sites de déchets. La première étape utilise les données à moyenne résolution spatiale des satellites Sentinel de l'Agence spatiale européenne. Les données du satellite Sentinel-2 sont filtrées pour repérer les sites d'enfouissement potentiels. Le modèle est étalonné au moyen de données d'observation de la Terre (OT) des sites d'enfouissement connus. Ces sites se trouvent dans la zone à examiner par le modèle et sont fournis par le Ministère. La méthodologie d'analyse utilisée dans cette étape fait appel à des techniques d'IA mises au point par ASE. La sortie du modèle est un fichier qui révèle les emplacements des sites d'enfouissement

potentiels, y compris les emplacements géographiques. Les faux positifs comme les carrières peuvent être signalés au cours de cette étape.

La deuxième étape utilise des données à haute résolution spatiale et des images aériennes pour examiner en détail les sites qui ont été repérés. Cette analyse est effectuée manuellement et permet de caractériser la structure interne des sites d'enfouissement, y compris par l'estimation du type de déchets. Les données utilisées dans cette étape proviennent du satellite DigitalGlobe de Maxar, de données d'OT à haute résolution et d'archives de photographies aériennes. ASE étudie actuellement l'automatisation de cette étape.

Une fois la liste des sites potentiels établie, des résumés d'une page sont générés pour tous les sites qui semblent être des sites d'enfouissement. Chaque résumé comprend une image satellite, l'emplacement du site (coordonnées et nom de l'emplacement ou de la rue, selon le cas), une précision quant au fait que le site est connu ou non, l'état du site (actif, historique, indéterminé, ainsi qu'une cote de confiance correspondante), la catégorie du site d'enfouissement selon la classification gouvernementale et une cote de confiance correspondante. Le résumé comprend également des commentaires sur les changements apportés au site au fil du temps, la taille approximative du site, la présence de véhicules, les bassins de décantation et le nom de l'organisation exploitante. Le catalogage et la caractérisation des risques de ces sites d'enfouissement fournissent au Ministère les renseignements nécessaires pour réduire l'évasion fiscale des sites d'enfouissement et réorienter le temps et les ressources financières vers les zones présentant un risque plus élevé de décharges illégales (voir figure 2).

Figure 2 : Profil de risque des sites de déchets repérés (Air & Space Evidence)



Résultats

ASE a mené deux projets pilotes pour le Ministère qui couvraient quatre zones distinctes en Nouvelle-Zélande. Environ 55 % des sites potentiels repérés par le modèle de détection ont été classés comme des sites d'enfouissement, tandis que les 45 % restants étaient des faux positifs. Parmi les sites classés comme étant des sites d'enfouissement, environ 60 % étaient connus du Ministère. Ils comprennent à la fois des sites autorisés et des sites illégaux qui font ou ont fait l'objet d'un processus d'application de la loi. Les 40 % de sites restants étaient « nouveaux », c'est-à-dire inconnus du Ministère. Le Ministère a décrit environ un tiers des « nouveaux » sites comme étant des découvertes « vraiment utiles ». Les deux autres tiers des « nouveaux » sites étaient des sites plus petits ou inactifs. Ces sites plus petits ou inactifs peuvent recevoir des lettres d'avertissement ou être surveillés de près pour détecter de futurs changements. Les projets pilotes ont démontré que la technologie de télédétection peut corroborer les informations existantes et permettre d'ajouter de nouveaux renseignements au registre des sites d'enfouissement. Le Ministère envisage actuellement de déployer le programme de repérage par surveillance satellite à l'échelle nationale.

Ressources mobilisées

Les projets pilotes ont nécessité la participation du personnel clé de la Waste and Resource Efficiency Division du Ministère. Une fois que le modèle a repéré les sites potentiels, des analystes spatiaux, des spécialistes des systèmes d'information géographique (SIG) et l'équipe chargée des autorisations d'enfouissement ont collaboré pour recouper les sites avec le registre national et d'autres registres des autorités locales. Le soutien des autorités locales dans le processus était également nécessaire pour vérifier et confirmer l'existence du site de déchets non autorisé. Le coût actuel de fonctionnement du système est d'environ de 31 275 \$ US à 37 530 \$ US par cellule de grille (zone de 110 × 110 km). Le coût comprend le temps du personnel du Ministère, d'ASE et des gouvernements locaux. Le coût total dépend de la disponibilité des images ainsi que du nombre de sites de déchets potentiels détectés par le modèle.

Facteurs de réussite

Expertise – Outre un spécialiste de l'IA, il était essentiel de solliciter des personnes ayant une expertise en matière d'environnement et de données satellitaires. Les connaissances des spécialistes de l'environnement ont été essentielles à l'entraînement du modèle d'IA (c.-à-d. que les connaissances environnementales ont été intégrées au modèle pour garantir un bon repérage des sites). Le spécialiste des satellites a également joué un rôle important et permis d'accéder à de vastes données satellitaires pouvant être utilisées pour l'application de techniques d'IA.

Partenariats – Bien qu'il y ait une certaine difficulté à établir des collaborations entre le secteur public et le secteur privé, notamment en ce qui concerne la technologie, la réussite du projet pilote peut être attribuée à une étroite collaboration entre le Ministère et ASE. La relation de travail a été couronnée de succès en raison des besoins distincts des deux parties. Le Ministère avait besoin d'une solution précise

pour repérer les sites de déchets inconnus, tandis qu'ASE avait besoin de données pour mettre à l'essai et améliorer le modèle. La collaboration est essentielle pour réaliser des progrès en matière d'IA et, à terme, pour faire avancer la gouvernance environnementale.

Obstacles

Adhésion de la haute direction – Dans le cas du ministère de l'Environnement de Nouvelle-Zélande, des renseignements étaient nécessaires. Cependant, certaines équipes de direction d'organismes environnementaux hésitaient à obtenir les renseignements générés par le modèle d'ASE, en raison des ressources et du financement limités. Le fait que les renseignements soient désormais révélés signifie que les organismes sont tenus d'agir, par l'intermédiaire d'inspections de sites et de mesures d'application de la loi. En réponse à cette « crainte liée aux connaissances obtenues », ASE a travaillé à l'élaboration d'un processus de profilage des risques pour les données générées, qui permettrait aux responsables de cibler les sites présentant le risque le plus élevé et possiblement les effets les plus importants et d'ainsi réduire le fardeau pesant sur les ressources et le financement d'un organisme.

Financement – Un financement adéquat permettant de mettre à l'essai de nouvelles technologies est un obstacle courant, en particulier pour les organismes gouvernementaux. Les organismes qui ont élaboré et réalisé une validation de principe, en particulier pour des modèles d'IA fondés sur des satellites, ont reconnu la valeur d'utiliser la technologie pour faciliter la surveillance et la vérification de la conformité. Toutefois, les organismes sont souvent limités par des budgets restreints et l'obtention de fonds pour étendre ces essais. La difficulté d'obtenir un financement est notamment liée à l'isolement des projets pilotes. Les organismes ne disposent pas d'une approche globale pour mettre à l'essai et adopter de nouvelles technologies comme l'IA. Pour aider à renforcer cette approche, le financement devrait également tenir compte de l'élargissement des équipes de données pour qu'elles renferment des capacités d'analyse avancées.

Possibilités pour l'Agence

L'AEIC pourrait bénéficier de la mise en œuvre d'un modèle de détection par satellite pendant le processus de vérification de la conformité. Les modèles de détection par satellite, comme le modèle d'ASE, pourraient être utilisés pour surveiller les limites et les activités des sites autorisés pendant la construction et l'exploitation, ainsi que pour repérer les activités de construction de projets non autorisés. Cette surveillance pourrait permettre de repérer des cas potentiels de non-conformité sans qu'il soit nécessaire d'être physiquement présent sur le site. Un modèle de détection pourrait en outre fournir un profil de risque pour une région particulière. Le signalement des cas de non-conformité et l'établissement d'un profil de risque aideraient l'AEIC à orienter ses ressources limitées vers les sites non conformes ou à risque élevé, ce qui pourrait s'avérer particulièrement utile dans la mesure où le nombre de déclarations de décision et de projets nécessitant une surveillance de la conformité augmente et où les ressources de l'Agence sont de plus en plus sollicitées. Un modèle de détection par satellite pourrait également servir de base pour les mesures d'application de la loi, dissuader les cas de non-conformité futurs et prévenir ou limiter les dommages environnementaux localisés en détectant les

cas de non-conformité plus tôt. L'AEIC doit tenir compte de nombreux facteurs pour déterminer si elle doit tenter d'acquérir un modèle de détection des changements par satellite.

Adhésion de la haute direction – Étant donné la « crainte liée aux connaissances obtenues » des autres organismes, l'AEIC devra se demander si la direction de l'Agence souhaite obtenir les données supplémentaires provenant d'un modèle de détection par satellite ou les utiliser. Le fait de concevoir le modèle comme un moyen de diriger les ressources existantes vers les sites à risque élevé pourrait atténuer d'éventuelles préoccupations liées à la découverte de cas de non-conformité supplémentaires.

Financement – Une source de financement continue peut être nécessaire pour qu'un modèle de détection par satellite puisse couvrir de grandes zones et être régulièrement mis à jour. Comme le financement peut être particulièrement difficile à obtenir pour les autorités environnementales, l'AEIC devra continuer à réfléchir à la façon d'intégrer le financement de l'IA à son budget. Toutefois, le financement peut également provenir d'autres sources. En Europe, il existe des programmes de financement qui permettent de financer des solutions technologiques lorsqu'un organisme gouvernemental intéressé s'associe à l'entreprise qui met au point la solution. Dans ces cas, l'organisme ne paie rien ou environ 20 % du coût. La recherche d'autres solutions comme celle-ci au Canada peut être une option pour l'AEIC. Par exemple, le partenariat pour l'IA, les données et la robotique, qui est l'un des partenariats européens dans le domaine du numérique, de l'industrie et de l'espace dans le programme Horizon Europe, fournit des fonds aux organismes, aux entreprises, aux universités et à d'autres organisations pour les innovations en matière d'IA.

Partenariats – Des experts en IA, en satellites et en environnement ont joué un rôle clé dans l'élaboration du modèle d'ASE. Pour tenter d'acquérir un modèle de détection par satellite, l'AEIC devrait envisager d'engager une société externe en vue d'exécuter le modèle pour elle ou d'embaucher des experts internes pour qu'ils mettent au point un modèle.

4.2 IA prédictive pour le ciblage des inspections

Autorité :	Environmental Protection Agency des États-Unis
Collaborateur :	Université de Chicago
Intention :	Déterminer la non-conformité des installations de traitement des déchets dangereux
Personne-ressource :	Randy Hill, Office of Enforcement and Compliance Assurance

Défi

L'Environmental Protection Agency (EPA) des États-Unis est responsable de la réglementation d'environ 1,2 million d'installations dans tout le pays en vertu des dix principales lois environnementales des États-Unis. Cependant, les ressources de l'EPA pour surveiller activement les installations réglementées

sont limitées. À l'heure actuelle, l'EPA compte environ 2800 employés travaillant dans le domaine de la conformité et de l'application de la loi, et avant la pandémie de COVID-19, ils ne réalisaient qu'environ 10 000 inspections sur place des installations réglementées chaque année. Bien que les États effectuent également des inspections, de nombreuses installations ne sont pas inspectées chaque année, de sorte que les cas de non-conformité pourraient continuer à ne pas être relevés. Afin de remédier à ce manque de ressources, l'EPA s'efforce d'améliorer son ciblage et de détecter davantage de cas de non-conformité.

Modèle d'IA pour l'analyse prédictive

En 2015, l'EPA a établi un partenariat avec l'Energy and Environment Lab de l'Université de Chicago (« le Laboratoire »), un laboratoire universitaire qui s'associe aux décideurs politiques afin de déterminer, de mettre à l'essai et de développer des solutions, des politiques et des programmes visant à relever les défis environnementaux. Le partenariat entre l'EPA et le Laboratoire vise à améliorer le ciblage des inspections et à accroître l'efficacité de l'EPA. Dans le cadre de cet objectif général, un projet se concentre sur la recherche des contrevenants actifs sur les sites de déchets dangereux. Pour trouver ces contrevenants, le Laboratoire utilise l'analyse prédictive afin de signaler les installations les plus susceptibles de contrevenir à la *Resources Conservation and Recovery Act* (RCRA).

Le modèle d'analyse prédictive de la RCRA prévoit des infractions « graves », définies comme étant le stockage sans permis, le traitement et l'élimination illégaux, ainsi que la détermination des déchets. Le modèle génère une cote de risque pour chaque installation qui représente la probabilité qu'une inspection de l'installation révèle une infraction grave à la RCRA. Avant le déploiement du modèle à l'échelle nationale, le Laboratoire a travaillé avec les bureaux régionaux pour effectuer des essais supplémentaires sur le terrain qui visaient à démontrer le rendement du modèle et à augmenter la confiance des inspecteurs dans sa valeur.

Le modèle est un algorithme d'apprentissage machine appelé Random Forest (« forêt aléatoire »), soit un modèle d'arbres de classification et de régression à étiquettes multiples écrit dans les logiciels R et Python. Il a conçu à partir de 15 ans de données historiques, y compris des dizaines de milliers de variables provenant de rapports réglementaires. Ces variables comprenaient les caractéristiques des installations, comme l'emplacement et l'industrie, et les données historiques de conformité et d'application de la loi pour la RCRA et d'autres lois comme la *Clean Air Act*. Une fois les variables et les prédicteurs générés, le modèle a été entraîné sur des données historiques des années 2000 à 2013. Le modèle a ensuite prédit le niveau de risque des installations en 2014, qui a ensuite été comparé aux infractions réelles.

Résultats

Le Laboratoire et l'EPA ont constaté que le modèle permettait d'augmenter de 50 % la détection des infractions graves par rapport au ciblage inchangé. D'après les déploiements régionaux en 2017, ils ont

constaté que, si l'EPA avait utilisé le modèle à l'échelle nationale, ils auraient pu trouver 214 contrevenants graves de plus avec les mêmes ressources d'inspection. Cependant, le modèle n'a été largement diffusé auprès de l'EPA et des partenaires étatiques, locaux et tribaux qu'en août 2020, et le temps écoulé depuis n'a pas permis d'évaluer son incidence réelle.

L'EPA et d'autres organismes partenaires étudient la possibilité d'étendre le modèle d'analyse prédictive existant à d'autres lois, notamment au ciblage des inspections effectuées en vertu de la *Clean Water Act* et de la *Clean Air Act*.

Ressources mobilisées

Le modèle a nécessité la participation d'employés clés de l'EPA, notamment Mike Barrette, spécialiste de la surveillance et de la protection de l'environnement, Rusty Wasem, spécialiste de la protection, et John Veresh, spécialiste de la gestion de l'information. M. Barrette est le responsable de la gestion des projets d'apprentissage machine à l'Université de Chicago. Le soutien des bureaux régionaux était également nécessaire pour les projets pilotes, et ils ont apporté leur contribution tout au long du processus de conception. Outre le temps de travail des membres de l'équipe de l'EPA qui ne sont pas des scientifiques des données et d'autres dépenses, le modèle a coûté environ 200 heures de travail d'un développeur chevronné en équivalent temps plein, soit quelque 20 000 \$ US sur une période de six mois, en plus des coûts d'hébergement annuels des Amazon Web Services (AWS), qui sont de 2400 \$ US.

Facteurs de réussite

Participation des intervenants – Les futurs utilisateurs du modèle ont eu la possibilité de faire part de leurs commentaires tout au long du processus de conception. Cette contribution a permis d'accroître la crédibilité du processus et la confiance des utilisateurs. L'EPA et le Laboratoire ont également travaillé avec les organismes et les associations d'organismes d'État, l'Environmental Council of the States et l'Association of State and Territorial Solid Waste Management Officials, tout au long du processus pour s'assurer que les résultats étaient accessibles et communiqués aux bons intervenants.

Disponibilité des données – Le modèle a été construit à partir de 15 ans de données historiques internes de l'EPA.

Obstacles

Compréhension des données – Au début du projet, la courbe d'apprentissage des spécialistes des données était abrupte. Le Laboratoire a dû investir beaucoup de ressources et de temps pour comprendre les données sous-jacentes et s'assurer qu'elles étaient exhaustives et de bonne qualité. Cependant, en raison du financement de sources multiples à l'université et des diverses contributions

des étudiants, le coût et la durée de cette aide sont impossibles à déterminer.

Plateforme de données – Après sa conception, le modèle a été transféré du Laboratoire à l'EPA. Il a fallu l'adapter pour qu'il fonctionne dans l'environnement AWS, qui était une technologie relativement nouvelle à l'EPA à l'époque. Il y a donc eu une courbe d'apprentissage pour faire fonctionner le modèle correctement sur les systèmes de l'EPA. Un remaniement important a également été nécessaire pour mettre à l'échelle le modèle afin de traiter cinq fois plus d'installations que celles incluses dans le premier essai sur le terrain.

Possibilités pour l'Agence

Les modèles d'analyse prédictive, comme celui de l'EPA, pourraient être utilisés pour le ciblage des inspections et l'application des lois, ce qui permettrait d'accroître l'efficacité de l'AEIC et la détection des cas de non-conformité. La prédiction des cas de non-conformité aiderait l'AEIC à diriger ses ressources limitées vers les installations les plus susceptibles d'enfreindre les règlements. Elle serait d'autant plus utile que le nombre de projets supervisés par l'AEIC va probablement fortement augmenter dans les années à venir, le nombre de projets approuvés augmentant considérablement. L'AEIC doit tenir compte de nombreux facteurs pour déterminer si elle doit tenter d'acquérir un modèle d'analyse prédictive.

Disponibilité des données – Le modèle a été construit à partir de 15 ans de données historiques et de dizaines de milliers de variables, ce qui constitue un obstacle important à la conception d'un modèle d'analyse prédictive semblable à court terme, étant donné que le programme de conformité, d'application de la loi et de suivi de l'AEIC est relativement nouveau et dispose de peu de données historiques. Par conséquent, cette technologie d'IA particulière pourrait être réexaminée dans plusieurs années, une fois que l'on disposera de plus de données historiques. L'AEIC pourrait également envisager la possibilité d'obtenir des données d'autres organismes canadiens. Les données historiques d'autres organismes peuvent faire du modèle d'analyse prédictive une option réalisable dans un avenir proche pour l'AEIC.

Partenariats – L'EPA s'est associée à un laboratoire universitaire pour fournir l'expertise en IA nécessaire à l'élaboration du modèle, qui n'a été intégré à la plateforme AWS de l'EPA qu'après la phase de conception. Comme l'AEIC ne dispose pas actuellement de l'expertise interne nécessaire en matière d'IA, elle pourrait chercher à établir des partenariats avec des groupes universitaires, des entreprises et des experts en IA ou engager des experts internes pour l'aider dans la recherche et le développement d'un tel modèle.

Participation des intervenants – Les bureaux régionaux ont participé à la phase d'essai pour s'assurer que le modèle fournissait des résultats facilement accessibles et pour obtenir l'appui des intervenants. L'AEIC devrait étudier comment mobiliser les intervenants tout au long des processus de recherche et

développement et de la phase pilote afin de garantir l'efficacité de la conception et de la mise en œuvre du modèle.

Financement – Il a fallu beaucoup de temps et de fonds pour concevoir et mettre à l'essai ce modèle. Le coût combiné de l'hébergement AWS et de la conception du modèle par un scientifique des données chevronné à l'EPA s'élève à environ 22 400 \$ US. Cependant, le projet a également bénéficié d'un soutien important de l'équipe de l'Energy and Environment Lab de l'Université de Chicago, pour un coût indéterminé. En outre, la sollicitation, la réception et l'intégration des commentaires des intervenants ont nécessité du temps et des fonds supplémentaires. Enfin, l'EPA s'attend à ce que la tenue à jour du projet comprenne le coût de l'hébergement AWS de 2400 \$ US et 20 heures par an de travail d'un scientifique des données chevronné à l'EPA. Ainsi, l'AEIC devra évaluer le rapport coûts-avantages d'un modèle d'analyse prédictive, envisager d'établir des relations avec des partenaires universitaires et obtenir un financement continu pour le projet, à même son budget ou d'une autre source.

4.3 IA de classification supervisée pour le traitement des plaintes

Autorité :	Superintendencia del Medio Ambiente (Surintendance de l'environnement), Chili
Collaborateur :	Université de Warwick
Intention :	Prévoir la pertinence et la gravité des plaintes déposées
Personne-ressource :	Cristóbal De La Maza Guzmán, surintendant de l'environnement

Défi

Au Chili, la Surintendance de l'environnement (SMA) est chargée de veiller au respect des règlements environnementaux et mène des activités de surveillance, de suivi et d'inspection. Elle reçoit également des milliers de plaintes relatives à l'environnement de la part de citoyens chaque année. Cependant, la SMA dispose de ressources limitées pour prendre connaissance de chaque plainte et orienter sa réponse en conséquence. En outre, certaines des plaintes déposées ne concernent pas la SMA ou ne contiennent pas suffisamment de renseignements pour être utiles. Puisque la SMA dispose de ressources limitées pour lire toutes les plaintes et déterminer leur urgence et leur pertinence, d'importants dommages environnementaux pourraient rester sans solution pendant de longues périodes.

Modèle de classification supervisée

En 2021, la SMA s'est associée à l'initiative Data Science for Social Good de l'Université de Warwick pour mettre au point un modèle d'apprentissage machine permettant de hiérarchiser les ressources au moment d'examiner les plaintes de citoyens relatives à l'environnement et de leur répondre, afin de pouvoir traiter les importants dommages environnementaux dans les meilleurs délais. Le modèle de

pertinence, un modèle de classification supervisée, classe les plaintes de citoyens relatives à l'environnement aux fins d'un examen plus approfondi par le personnel. Le modèle attribue aux plaintes les étiquettes « pertinentes », « dérivation » ou « archive I ». Les plaintes « dérivation » et « archive I » ne contiennent pas suffisamment de renseignements pour que des mesures soient prises. Un deuxième modèle, le modèle de gravité des sanctions, prédit la gravité des sanctions pour les plaintes indiquées comme « pertinentes » pour la SMA, ce qui facilite davantage la hiérarchisation des inspections.

Les deux modèles génèrent actuellement des prévisions pour les nouvelles plaintes reçues par la SMA, et les résultats sont recueillis dans une base de données Azure SQL, une base de données infonuagique fournie par Microsoft, pour effectuer des requêtes. À terme, ces prévisions seront utilisées pour déterminer les plaintes qui feront l'objet d'un suivi et l'ordre de ce suivi, mais pour l'instant, ces prévisions sont comparées aux décisions prises par le personnel de la SMA.

Le modèle a été conçu à partir des données du registre des plaintes antérieures de la SMA. Certaines de ces données étaient structurées, d'autres étaient sous forme de texte libre provenant directement des citoyens. Le modèle a également été alimenté de données supplémentaires sur les facteurs sociaux et géographiques dans l'ensemble Chili. Un modèle de forêts aléatoires est utilisé pour analyser les données.

Comme le montre la figure 3, le modèle est capable de mieux prédire les plaintes « pertinentes » et « dérivation » que les plaintes « archive I », lorsqu'on compare ses résultats aux étiquettes réelles attribuées par le personnel de la SMA. Cette situation est en grande partie attribuable à la structure du modèle, car les plaintes de type « archive I » ne sont pas l'objet d'une inspection, et la SMA a déterminé qu'il est préférable d'étudier, à terme, les plaintes sans importance plutôt que de ne jamais les étudier. En outre, la baisse des plaintes « pertinentes » est peut-être due à la transition vers un nouveau système de plaintes en ligne au début de 2021. Comme le modèle a été entraîné sur l'ancien système, les plaintes qu'il tente maintenant de prévoir peuvent être très différentes de celles avec lesquelles il a été entraîné. C'est un domaine que la SMA prévoit d'approfondir en continuant à perfectionner le modèle.

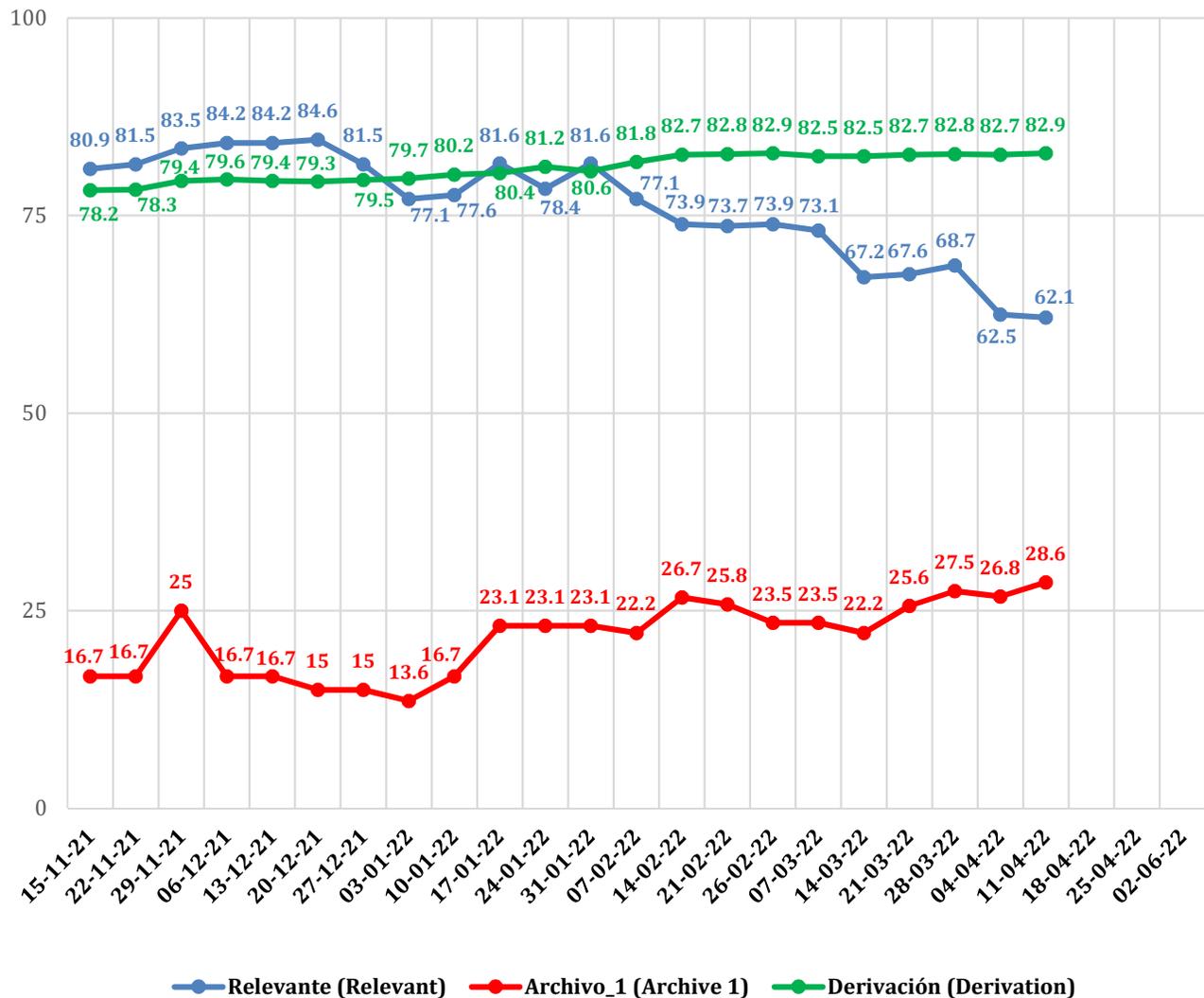


Figure X :

Résultats

La SMA estime que la combinaison des modèles de pertinence et de gravité des sanctions réduira de 80 % le temps de réacheminement des plaintes vers l'organisme approprié, de 85 % le temps d'archivage des plaintes et de 65 % le temps d'inspection des plaintes pouvant mener à des sanctions. Cette estimation est basée sur une comparaison des délais moyens de traitement des plaintes avant et après les contributions du modèle. Par exemple, le traitement des plaintes de type « archive » prend en moyenne cinq jours, et on estime que le modèle permet de réduire ce délai à un jour, ce qui représente une réduction de 80 % du temps de traitement. À terme, les modèles permettraient d'accroître l'efficacité de la SMA.

Ressources mobilisées

Les modèles sont actuellement exploités par un ingénieur des données qui gère l'infrastructure infonuagique du système. Toutefois, l'initiative Data Science for Social Good de l'Université de Warwick a conçu le modèle avant de passer à la plateforme infonuagique de la SMA. En outre, au moins deux scientifiques des données participeront à une prochaine analyse du modèle afin d'étudier davantage la baisse de précision illustrée à la figure 3.

La SMA a reçu une bourse pour que l'équipe de l'initiative Data Science for Social Good appuie la conception du modèle. Le principal coût pour la SMA a donc été de payer le salaire de son personnel interne pendant les phases de conception, de maintenance et d'exploitation du projet. La SMA devra également payer le salaire des nouveaux scientifiques des données pour l'analyse des modèles à venir. Puisque les données utilisées dans le cadre de cette utilisation de l'IA sont déjà recueillies et stockées dans les bases de données du SMA, les coûts d'expansion du projet devraient être relativement minimes à l'avenir.

Facteurs de réussite

Disponibilité des données – Les données nécessaires à la conception du modèle étaient facilement accessibles dans les bases de données historiques des organismes, et elles étaient directement liées au problème que la SMA essayait de résoudre.

Expertise interne – Si les partenaires externes ont soutenu le projet et joué un rôle crucial, disposer d'une expertise interne en apprentissage machine, science des données et ingénierie des données était impératif pour la mise en œuvre du projet.

Participation des intervenants – Bien que la SMA n'ait pas encore fait participer les intervenants au processus parce que le modèle n'est pas encore prêt pour cette phase, elle sait qu'il sera essentiel de mobiliser tous les intervenants pour que le modèle soit un succès.

Obstacles

Partenariats – Communiquer efficacement avec les partenaires externes au sujet du problème et des objectifs de l'Agence, en plus d'intégrer le travail des partenaires externes à la culture organisationnelle, a été un défi. Les réunions d'équipe hebdomadaires avec les partenaires externes ont amélioré la communication et ont permis de faire participer l'équipe externe à la culture de la SMA, ce qui a ainsi amélioré ainsi les partenariats et les résultats.

Possibilités pour l'Agence

Un modèle de classification supervisé comme celui que la SMA conçoit actuellement pourrait être utile à l'AEIC afin de trier et de classer les plaintes du public en vue d'une inspection plus approfondie. Le tri des commentaires en fonction de leur pertinence et de leur gravité pourrait contribuer à accroître

l'efficacité et l'incidence au moment de répondre aux plaintes et de prendre des mesures de conformité et d'application de la loi. L'AEIC devrait tenir compte de nombreux facteurs pour déterminer si elle doit tenter d'acquiescer un modèle de classification supervisée pour les plaintes du public.

Disponibilité des données – La SMA a construit son modèle à partir d'un registre interne de plaintes antérieures. Puisque le programme de conformité, d'application de la loi et de suivi de l'AEIC est relativement nouveau et que relativement peu d'installations sont actuellement surveillées, il pourrait ne pas y avoir de données historiques permettant de construire un modèle de classification supervisée pour le moment. L'AEIC pourrait envisager un tel modèle dans plusieurs années, lorsqu'elle disposera de plus de données historiques. À titre de référence, la SMA a mis en œuvre et mis son modèle à l'essai en utilisant neuf ans de données et plusieurs milliers de plaintes. Toutefois, comme les paramètres de sa collecte de données ont changé au fil du temps (c.-à-d. un nouveau formulaire de plainte), elle s'attend à devoir entraîner de nouveau son modèle.

Partenariats – La SMA s'est associée à l'équipe de l'initiative Data Science for Social Good pour concevoir son modèle. L'AEIC devra déterminer s'il convient de mettre au point un modèle semblable à l'interne, auquel cas elle devra engager des experts compétents pour le faire ou conclure un marché avec une partie externe. Si l'AEIC devait passer solliciter les services d'une partie externe, elle devrait également envisager la meilleure façon de communiquer avec elle afin d'harmoniser les objectifs de tous et de concevoir le modèle en tenant compte du contexte particulier de l'AEIC.

Expertise – La SMA a relevé l'importance d'avoir à la fois des partenaires externes pour aider à la conception du modèle et une expertise interne pour le mettre en œuvre. L'AEIC devra envisager de recourir à des sous-traitants ou d'engager des experts internes pour exécuter et entretenir le modèle après sa conception si elle ne dispose pas déjà des experts internes nécessaires.

Financement – La SMA a reçu une bourse pour son partenariat avec l'équipe de l'initiative Data Science for Social Good pour la conception du modèle, mais a également payé le salaire des scientifiques des données internes. L'AEIC devra chercher des sources de financement pour la conception du modèle ainsi que pour sa maintenance régulière et sa mise en œuvre. Bien qu'une partie du financement puisse provenir de sources ou de bassins externes, l'AEIC devra peut-être aussi envisager d'intégrer dans son budget un financement continu d'environ 5900 \$ US par an.

4.4 IA prédictive pour le ciblage des inspections

Autorité :	Environnement et Changement climatique Canada
Collaborateur :	S.O., conçu à l'interne
Intention :	Prévoir les cas de non-conformité et les risques dans les installations réglementées
Personne-ressource :	Michael Enns, directeur général de l'évaluation des risques à Environnement et Changement climatique Canada

Défi

La Direction générale de l'application de la loi (DGAL) d'Environnement et Changement climatique Canada (ECCC) est chargée d'enquêter sur les installations réglementées pour s'assurer de leur conformité réglementaire. Cependant, la DGAL n'a la capacité d'inspecter qu'une petite fraction des installations chaque année. Le Bureau du vérificateur général a relevé que la DGAL devait faire meilleur usage d'une approche fondée sur le risque pour ses activités, en tenant compte des installations qui présentent la plus forte probabilité de non-conformité et des activités qui posent le plus grand risque pour l'environnement, afin de réduire plus efficacement les cas de non-conformité.

Modèle d'IA pour l'analyse prédictive

Le Bureau du dirigeant principal des données et la DGAL d'ECCC se sont associés à des scientifiques des données externes pour mettre au point le Micro Enforcement Targeting Algorithm (META). Le modèle META est un modèle d'apprentissage machine conçu pour aider la DGAL à adopter une approche fondée sur le risque pour les inspections de sites. Il permet de prévoir les cas de non-conformité à l'échelle des installations afin d'aider ECCC à orienter ses ressources vers les zones soupçonnées d'être les plus préoccupantes.

Le modèle META est un modèle de classification supervisée construit à partir de plus de 25 ans de données historiques d'application de la loi et de milliers d'inspections antérieures. Ces données ont été recueillies auprès d'entités cibles par l'intermédiaire de diverses exigences de déclaration réglementaires et dans des ensembles de données financières accessibles au public. Les données ont été extraites, transformées et téléchargées dans un nouveau modèle de données, puis les paramètres opérationnels du Bureau du dirigeant principal des données ont été intégrés. Plus précisément, le modèle est basé sur des données relatives aux activités d'application de la loi et à la promotion de la conformité, des données scientifiques (eau, air, sol), des données industrielles, des données financières, et plus encore. Le modèle utilise la régression logistique, XGBoost (un algorithme d'apprentissage machine d'ensemble à base d'arbres de décision) et un réseau neuronal profond pour prédire la non-conformité et les risques en trouvant des relations complexes entre les caractéristiques d'entrée et le résultat de non-conformité.

À l'heure actuelle, de nombreuses caractéristiques du modèle META reposent sur des données historiques et des visites de sites. ECCC continue d'ajouter de nouvelles données historiques au modèle afin d'améliorer sa précision globale et de calculer et mettre à l'essai de nouvelles hypothèses. À terme, ECCC vise à concevoir un modèle qui peut être extrapolé aux sites qui n'ont pas fait l'objet de visites par la DGAL.

Résultats

La première génération du modèle META est terminée. Lors des essais, elle a permis de prédire les cas de non-conformité à l'échelle des installations avec trois fois plus de précision que les méthodes de

ciblages antérieures utilisées par la DGAL. Cependant, en raison de la pandémie de COVID-19, les inspections sur place ont été suspendues, de sorte qu'ECCC n'a pas encore reçu un ensemble complet de résultats de sa première expérience de « vérification sur le terrain ».

Ressources mobilisées

Le Bureau du dirigeant principal des données et la DGAL d'ECCC ont tous deux participé à la mise au point du modèle. Cependant, le modèle META a principalement été conçu par un petit groupe de scientifiques des données qui font maintenant tous partie de la DGAL.

Le paiement du salaire d'un à cinq employés à plein temps spécialisés dans la science des données constitue le principal coût du modèle META. En outre, au cours de la mise au point, deux scientifiques des données chevronnés ont dû travailler à temps plein, ce qui représente un salaire et des coûts de fonctionnement d'environ 130 000 \$ CA par an. Ont également contribué aux travaux trois analystes subalternes, pour environ 90 000 \$ CA par an, et plusieurs étudiants à 20 \$ CA par heure pendant des périodes de quatre mois. Enfin, ECCC a par la suite investi environ 30 000 \$ CA dans la puissance de calcul.

Facteurs de réussite

Adhésion de la haute direction – Comme le modèle META n'a pas une incidence immédiate, l'adhésion de la direction à l'AEIC est essentielle. Elle accorde aux développeurs le temps et les ressources nécessaires pour concevoir et mettre à l'essai le modèle.

Évolution du modèle – Le processus cyclique comprenant la génération de prévisions basées sur un ensemble d'indicateurs, la validation de ces prévisions par des inspections sur place et l'utilisation des résultats de ces inspections pour améliorer le modèle META est important. Ce processus itératif permet une mise au point et une amélioration continues de la précision et de la pertinence du modèle.

Obstacles

Incohérences dans les données – Il a fallu beaucoup de temps pour normaliser, nettoyer et formater correctement les données historiques d'application de la loi. Les scientifiques des données ont dû élaborer des méthodes permettant de dédupliquer les données, de corriger les erreurs et de relier les données pour jeter les bases du modèle META. Plus particulièrement, ils ont consacré beaucoup de temps à la correction des noms et des adresses des installations. Cependant, bien que la préparation des données ait nécessité beaucoup de ressources, elle a également permis à l'organisme d'améliorer la précision de ses rapports et de ses produits de renseignement au-delà du modèle META.

Possibilités pour l'Agence

Les modèles d'analyse prédictive, comme celui d'ECCC, pourraient servir à cibler les inspections en fonction des risques à l'AEIC, ce qui pourrait permettre d'accroître la détection des cas de non-

conformité et d'optimiser l'utilisation des ressources. Alors que le nombre de projets supervisés par l'AEIC augmentera dans les prochaines années, le modèle pourrait s'avérer particulièrement utile pour repérer les projets non conformes. L'AEIC devrait tenir compte de nombreux facteurs pour déterminer si elle doit tenter d'acquérir un modèle d'analyse prédictive.

Disponibilité des données – Le modèle a été construit à partir de 25 ans de données historiques et d'inspections de site, ce qui constitue un obstacle important à la conception d'un modèle d'analyse prédictive semblable, étant donné que le programme de conformité, d'application de la loi et de suivi de l'AEIC est relativement nouveau et dispose de peu de données historiques. Par conséquent, cette technologie d'IA particulière pourrait être réexaminée dans plusieurs années, une fois que l'on disposera de plus de données historiques. L'AEIC pourrait également envisager la possibilité d'obtenir des données d'autres organismes canadiens. Les données historiques d'autres organismes peuvent faire du modèle d'analyse prédictive une option réalisable dans un avenir proche pour l'AEIC.

Adhésion de la haute direction – Puisqu'il faut du temps pour concevoir, mettre à l'essai et mettre en œuvre un modèle d'analyse prédictive, l'adhésion de la haute direction a été essentielle pour ECCC lors de la réalisation du modèle META. L'adhésion a permis aux scientifiques des données de disposer du temps et des ressources nécessaires pour concevoir le modèle. L'AEIC devrait se demander si les cadres concernés seraient d'accord avec un modèle d'analyse prédictive.

Expertise – ECCC comptait entre un et cinq scientifiques des données travaillant à temps plein pour concevoir et mettre à l'essai le modèle META, et ces scientifiques font désormais partie de la DGAL. L'AEIC devrait déterminer qui a l'expertise nécessaire pour mettre au point un modèle d'analyse prédictive semblable, et si cette expertise est déjà présente à l'interne, s'il faut embaucher de nouveaux employés ou si des experts externes peuvent aider à la conception du modèle. L'AEIC doit également prendre en compte la normalisation et le formatage des données que les experts devront effectuer lorsqu'elle désignera des experts potentiels. L'adhésion de la haute direction est également essentielle pour recruter les experts nécessaires.

Financement – Il a fallu beaucoup de temps et de fonds pour concevoir ce modèle. Plus précisément, l'équipe fonctionnait à un coût d'environ 555 000 \$ CA par an pour le personnel et le financement. L'AEIC devra évaluer le rapport coûts-avantages d'un modèle d'analyse prédictive et obtenir un financement continu pour le projet, à même son budget ou d'une autre source.

4.5 Traitement intelligent des documents pour l'évaluation de la conformité

Autorité :	Santé Canada
Collaborateur :	S.O., conçu à l'interne
Intention :	Attribuer des cotes de risque aux cas de non-conformité et aider à la rédaction des rapports
Personne-ressource :	Justin Budgell, agent principal de la conformité

Défi

Une partie des responsabilités de Santé Canada comprend les activités de surveillance de la conformité et d'application de la loi relatives aux produits de santé. Dans le cadre de ces responsabilités, les inspecteurs de Santé Canada de la Direction générale des opérations réglementaires et de l'application de la loi effectuent des inspections de sites pour s'assurer que les installations respectent la *Loi sur les aliments et drogues*. Ils rédigent et publient également un rapport d'inspection final pour chaque installation. La Direction générale inspecte un grand nombre de sites chaque année et rédige des rapports sur ceux-ci. Historiquement, la rédaction du rapport d'inspection final pour chaque site était effectuée manuellement, y compris l'attribution d'une cote de risque et la mise en correspondance d'une partie précise de la *Loi sur les aliments et drogues* avec chaque cas de non-conformité sur un site. Ce processus manuel laissait la place à des écarts et à des erreurs humaines et d'interprétation, en plus de demander beaucoup de temps aux employés.

Traitement intelligent des documents pour l'évaluation de la conformité

Pour normaliser les rapports d'inspection et faire gagner du temps aux inspecteurs, Santé Canada a commencé en 2019 à mettre au point un algorithme d'apprentissage machine, Cipher. Cet algorithme automatise l'attribution d'une cote de risque à chaque cas de non-conformité dans une installation et la mise en correspondance de chaque cas à un règlement précis de la *Loi sur les aliments et drogues*. Il attribue également un texte type – une explication générique d'un type particulier de non-conformité qui peut être rendue publique dans le rapport d'inspection final – à chaque cas de non-conformité. Santé Canada travaille également à l'expansion de Cipher pour y inclure une composante d'analyse prédictive permettant de prédire si un site respectera la réglementation.

Cipher a été construit à partir de 10 ans de données historiques internes. Au cours de la première phase, les données ont été extraites des rapports d'inspection et analysées à la recherche de tendances en vue de déterminer leur utilité pour automatiser les processus et pour prévoir les cas de non-conformité. Un outil de validation de principe a permis de déterminer que l'IA et l'apprentissage machine pouvaient attribuer une cote de risque et faire correspondre un règlement de la *Loi sur les aliments et drogues* et une seule observation de non-conformité. Au cours de la deuxième phase, Cipher a été étendu pour

permettre aux inspecteurs de saisir en une seule fois toutes les observations d'une visite de site et d'ajouter le texte type pour chaque cas de non-conformité.

Les inspecteurs ont procédé à des essais et fourni des commentaires sur cette deuxième phase, et Cipher entame désormais son deuxième cycle d'essais.

Résultats

Pour l'instant, les résultats du projet CIPHER sont inconnus, car il est encore en phase d'essai.

Ressources mobilisées

Le projet a commencé par un partenariat interne entre les directions générales de Santé Canada. Le personnel clé comprend Justin Budgell, chef de projet; Peter Yoon, conseiller de projet; Betty Palma, inspectrice; Valerie Bergeron, championne de projet à la Division de l'inspection des produits de santé et de l'octroi des permis (DIPSOP); Cecilia Bong, gestionnaire de projet à la DIPSOP, Bryan Paget, scientifique des données/développeur; Mithu Selvakumar, inspecteur; et Sherry Bahaw, inspectrice.

Santé Canada s'est occupé de l'extraction des données en interne, mais a fait appel à un entrepreneur tiers pour certains aspects de la phase d'analyse des données et pour déplacer CIPHER vers un environnement infonuagique protégé B, ce qui était nécessaire pour protéger les renseignements classés. Santé Canada est en processus d'embauche d'un autre scientifique des données afin de continuer à mettre au point l'outil. Santé Canada a également signé un protocole d'entente avec l'École de la fonction publique du Canada (EFPC). Cette dernière est invitée à participer aux réunions du projet, et Santé Canada est informé des autres projets d'IA dont l'EFPC a pris connaissance.

Le coût de la mise au point, des essais et de la mise en œuvre du modèle CIPHER est estimé à environ 500 000 \$ CA.

Facteurs de réussite

La précision des outils, la confiance des utilisateurs, la disponibilité des ressources et la disponibilité de la technologie sont des facteurs que Santé Canada juge importants pour la réussite du projet CIPHER. Comme le projet CIPHER est encore en phase d'essai et que les résultats sont inconnus, l'incidence de chaque facteur demeure également inconnue.

Obstacles

Plateforme de données – La nécessité de stocker les données dans un environnement infonuagique protégé B tout en permettant à plusieurs utilisateurs d'y accéder en même temps a retardé la phase pilote du projet et les étapes suivantes. Pour surmonter cet obstacle, Santé Canada s'est associé à Environnement et Changement climatique Canada (ECCC) pour accéder à son environnement

informatique afin de mettre le modèle à l'essai. L'environnement informatique d'ECCC n'a pas les mêmes restrictions de sécurité que celui de Santé Canada.

Expertise – Santé Canada avait besoin d'embaucher des scientifiques de données qualifiés pour concevoir et mettre à l'essai l'outil. L'embauche d'employés possédant l'expertise nécessaire en temps opportun a constitué un défi pour Santé Canada; toutefois, le ministère a finalement pu embaucher des experts internes et concevoir l'outil essentiellement en interne.

Possibilités pour l'Agence

Les modèles de traitement intelligent des documents, comme celui de Santé Canada, pourraient être utilisés pour les activités de suivi, comme le traitement des rapports des programmes de suivi des promoteurs de projets à l'AEIC. Comme l'Agence est chargée de rédiger des rapports de suivi à partir des résultats du programme de suivi d'un promoteur, un modèle comme Cipher pourrait être particulièrement utile pour normaliser et simplifier le traitement des documents et la rédaction des rapports. Lorsqu'il s'agit de savoir s'il faut mettre en œuvre un tel modèle, l'AEIC doit tenir compte de quelques facteurs.

Disponibilité des données – Cipher a été construit à partir de 10 ans de données historiques et de rapports. Comme le programme de conformité, d'application de la loi et de suivi de l'AEIC est relativement nouveau, il se peut qu'elle ne dispose pas des données historiques et des rapports nécessaires pour concevoir un modèle d'apprentissage machine précis qui l'aiderait à rédiger les rapports et à évaluer les risques. Par conséquent, cette technologie pourrait devenir accessible à l'AEIC plus tard, quand davantage de données historiques auront été accumulées.

Expertise – Santé Canada a embauché plusieurs scientifiques des données pour créer Cipher, et a collaboré avec un tiers pour l'analyse des données et le travail sur la plateforme. L'AEIC devrait envisager d'embaucher des scientifiques des données ou de solliciter les services d'experts tiers pour concevoir un modèle comme Cipher, car l'Agence ne dispose pas à l'heure actuelle de l'expertise ou des capacités nécessaires à l'interne.

Financement – Plutôt que de faire appel à des partenaires externes, Santé Canada se fie à des experts internes en ce qui a trait à la science des données, à l'inspection et à d'autres questions. Cette façon de faire pourrait permettre de réduire certains coûts de conception.

4.6 Résumé des cas d'utilisation

Les cas d'utilisation ont mis en évidence plusieurs facteurs qui ont eu une incidence positive ou négative sur la conception et la mise en œuvre du modèle d'IA d'un organisme. Bien que la situation de chaque organisme soit unique et nécessite donc des considérations différentes, de nombreux facteurs

s'avéraient pertinents pour plus d'un organisme. Les cas d'utilisation ont révélé que le financement, l'adhésion de la haute direction, l'expertise, la plateforme de données, les partenariats et la compréhension des données ainsi que les incohérences dans ces dernières constituaient des obstacles pour certains organismes travaillant à la conception des modèles d'IA (tableau 1).

<i>Cas d'utilisation</i>	<i>Obstacles</i>					
	<i>Financement</i>	<i>Adhésion de la haute direction</i>	<i>Compétences</i>	<i>Plateforme de données</i>	<i>Partenariats</i>	<i>Compréhension des données et incohérences dans ces dernières</i>
3.1 Détection des changements par ASE	Nécessaire pour les phases pilote et d'essai	Les organismes ne veulent pas de renseignements de la part d'ASE.				
3.2 Modèle de la RCRA de l'EPA				Courbe d'apprentissage pour faire fonctionner le système sur la plateforme de l'EPA et le développer		Importantes ressources investies pour comprendre les données historiques
3.3 Classification supervisée de la SMA					Communication efficace et intégration de la culture de l'organisme	
3.4 Modèle META d'ECCC						Importantes ressources investies pour normaliser, nettoyer et formater les données historiques
3.5 Cipher de Santé Canada			Nécessité de recruter des scientifiques des données en interne	La sécurité requise sur la plateforme a retardé la phase pilote.		

Tableau 1 : Obstacles à la conception et à la mise en œuvre de l'IA à partir des cas d'utilisation des organismes.

Les cas d'utilisation ont également révélé que l'adhésion de la haute direction, l'évolution du modèle, la disponibilité des données, les partenariats, l'expertise et la participation des intervenants étaient des facteurs qui ont favorisé la conception de modèles d'IA dans certains organismes (tableau 2).

<i>Cas d'utilisation</i>	<i>Facteurs de réussite</i>					
	<i>Adhésion de la haute direction</i>	<i>Évolution du modèle</i>	<i>Disponibilité des données</i>	<i>Partenariats</i>	<i>Compétences</i>	<i>Participation des intervenants</i>
3.1 Détection des changements par ASE				Collaboration étroite entre l'organisme et l'équipe d'ASE	Participation d'experts en IA, en environnement et en données satellitaires	
3.2 Modèle de la RCRA de l'EPA			15 ans de données historiques pour mettre au point le modèle			Les futurs utilisateurs fournissent des commentaires tout au long du processus de conception.
3.3 Classification supervisée de la SMA			Données provenant d'une base de données historiques interne		L'expertise interne est essentielle à la mise en œuvre.	Sera impérative à l'avenir
3.4 Modèle META d'ECCC	Adhésion de la haute direction pour fournir les ressources nécessaires du projet	Processus itératif d'inspections sur place et d'amélioration du modèle				
3.5 Cipher de Santé Canada			Disponibilité des ressources et des outils			Renforcement de la confiance des utilisateurs

Tableau 2 : Facteurs de réussite de la conception de l'IA à partir des cas d'utilisation des organismes.

En outre, les cas d'utilisation ont révélé que les organismes s'appuyaient généralement sur une combinaison d'employés internes et d'experts externes pour concevoir et mettre en œuvre leurs modèles d'IA (tableau 3). Les dépenses engagées pour les projets de cas d'utilisation comprenaient généralement les salaires des employés et le financement de la maintenance et de la surveillance des modèles à long terme. Les organismes ont également parfois payé les frais d'hébergement de la

plateforme et des partenaires externes, mais cela varie selon les cas d'utilisation et les programmes de financement.

<i>Cas d'utilisation</i>	<i>Ressources mobilisées</i>	
	<i>Personnel</i>	<i>Argent</i>
3.1 Détection des changements par ASE	Employés de l'organisme, personnel d'ASE, autorités locales	38 829 \$ CA – 46 595 GBP par zone de 110 × 100 km
3.2 Modèle de la RCRA de l'EPA	Employés de l'organisme, personnel des partenaires universitaires, bureaux régionaux de l'EPA	200 heures ETP d'un développeur principal, 3086 \$ CA en coûts annuels d'hébergement AWS
3.3 Classification supervisée de la SMA	Ingénieur de données de l'organisme, partenaires externes, deux scientifiques de données de l'organisme	Bourse pour le groupe partenaire, salaire des experts internes
3.4 Modèle META d'ECCC	Bureau du dirigeant principal des données d'ECCC, DGAL, d'un à cinq scientifiques des données d'ECCC	Les salaires des scientifiques des données, les salaires des analystes subalternes et 30 000 \$ CA pour l'informatique, ce qui totalise environ 555 000 \$ CA.
3.5 Cipher de SC	Huit employés internes, des entrepreneurs tiers pour l'analyse des données et la plateforme	500 000 \$ CA

Tableau 3 : Ressources mobilisées pour les modèles d'IA à partir des cas d'utilisation.

À mesure que l'AEIC et d'autres organismes environnementaux exécutent des projets d'IA, ils doivent tenir compte de l'incidence de ces facteurs sur la recherche, la conception et la mise en œuvre de l'IA dans leur organisme particulier.

4.7 EnviroVerse – L'avenir de la surveillance réglementaire

À mesure que l'intérêt, le financement et les compétences augmentent, le secteur de l'IA environnementale continuera à progresser. Ces progrès pourraient notamment inclure des algorithmes d'analyse prédictive à grande échelle ainsi qu'une efficacité accrue. L'un de ces progrès pourrait être la possibilité d'effectuer des inspections de sites virtuelles plutôt que sur place. Enviro.AI, une société spécialisée dans la technologie des données et la conformité environnementale, conçoit actuellement EnviroVerse, une couche du métavers qui permettrait aux agents d'application de la loi de procéder à des inspections virtuelles (figure 4). Enviro.AI note également qu'EnviroVerse pourrait être utilisé pour la surveillance, la conformité, l'octroi de permis, la vérification, et plus encore. La plateforme est encore en développement et ne sera pas disponible avant plusieurs années, mais elle offre un exemple des travaux novateurs en matière d'IA environnementale qui pourraient voir le jour dans les années à venir.

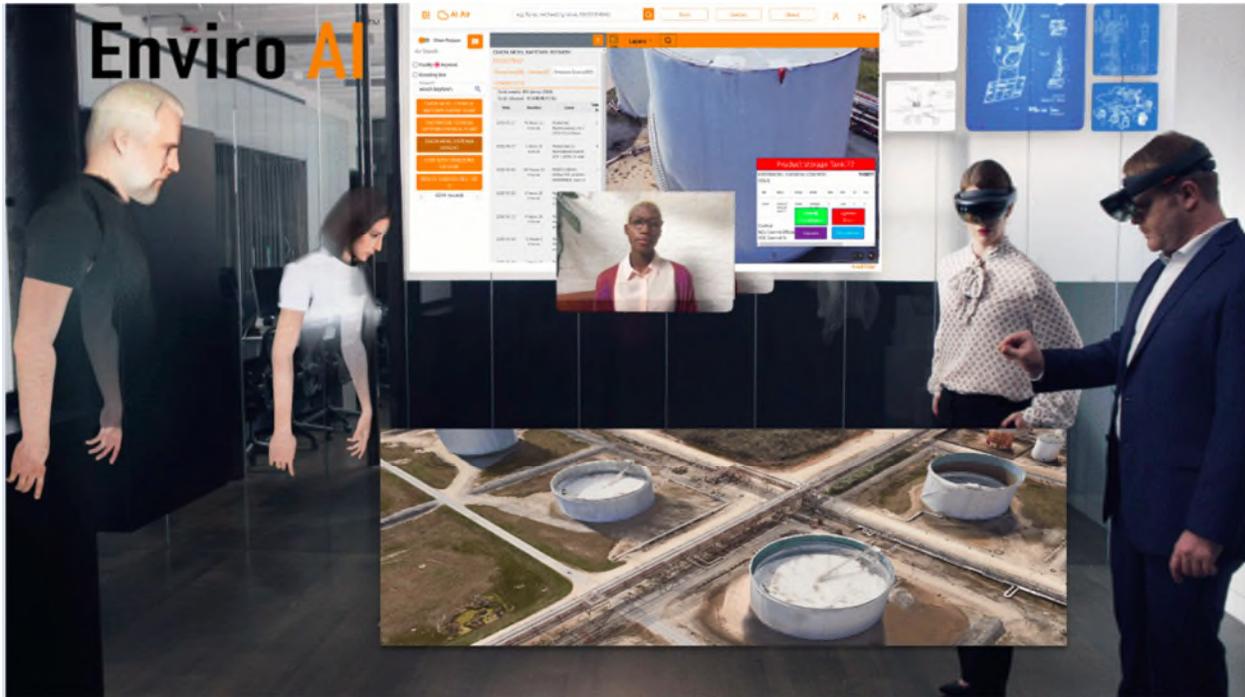


Figure 4 : Matériel promotionnel pour EnviroVerse par Enviro.AI.

5 CRITÈRES D'ADOPTION

Alors que l'AEIC étudie comment les modèles d'IA et d'apprentissage machine peuvent contribuer à accroître l'efficacité de ses activités postérieures aux décisions et à gérer sa charge de travail croissante, elle doit évaluer les applications potentielles dans son propre contexte. Ces considérations comprennent notamment les faits suivants : le programme de conformité, d'application de la loi et de suivi de l'AEIC est relativement nouveau; l'AEIC mène actuellement des activités postérieures aux décisions dans de nombreuses installations, mais plus de déclarations de décision seront diffusées dans les années à venir; l'AEIC compte une petite équipe et dispose de peu de données internes et historiques; l'AEIC travaille actuellement à affecter une partie de son budget à des initiatives d'IA, mais n'a pas, pour le moment, de source de financement continue pour les projets d'IA; et l'AEIC n'emploie pas de scientifiques des données ou d'experts en IA en interne à l'heure actuelle.

Compte tenu de ces circonstances, l'Agence devrait tenir compte de critères particuliers avant les phases de recherche, de conception et de mise en œuvre d'un modèle d'IA, et tout au long de celles-ci. Ces facteurs sont établis à partir des cas d'utilisation (tableau 4) et de l'analyse documentaire.

<i>Cas d'utilisation</i>	<i>Considérations de l'Agence en matière de possibilités</i>					
	<i>Financement</i>	<i>Partenariats</i>	<i>Disponibilité des données</i>	<i>Participation des intervenants</i>	<i>Adhésion de la haute direction</i>	<i>Compétences</i>
3.1 Détection des changements par ASE	X	X			X	
3.2 Modèle de la RCRA de l'EPA	X	X	X	X		
3.3 Classification supervisée de la SMA	X	X	X			X
3.4 Modèle META d'ECCC	X		X		X	X
3.5 CIPHER de SC			X			X

Tableau 4 : *Considérations de l'Agence en matière de possibilités à partir des cas d'utilisation.*

Disponibilité des données – Les données constituent la base d'un projet d'IA et, même si les projets d'IA ne nécessitent pas tous qu'un organisme dispose d'une base de données historiques, de nombreux projets d'IA se fondent sur des données historiques pour concevoir le modèle. En outre, les données disponibles doivent être fiables, de bonne qualité et représentatives [Government Accountability Office des États-Unis, 2021]. Elles doivent être organisées et stockées dans une base de données accessible, où les sources sont indiquées. Idéalement, les données devraient être intégrées et accessibles sur une plateforme centralisée plutôt que dans des dépôts de données discrètes, et des employés de l'organisme devraient se consacrer au maintien de la gouvernance et de la qualité des données (GoDataDriven, s.d.). Il peut être nécessaire de procéder à une analyse des données pour s'assurer que les données historiques répondent à ces facteurs avant de les utiliser pour la conception d'un modèle d'IA.

Ce besoin de données de qualité est un aspect que l'AEIC devra sérieusement évaluer lorsqu'elle envisagera de mettre en place les modèles d'IA, car il influera probablement sur les types de projets qui sont réalisables maintenant et sur ceux qui devront être revus plus tard. Comme il s'agit d'un programme relativement nouveau, l'AEIC devra probablement se concentrer sur les projets d'IA qui ne nécessitent pas de données historiques fiables, comme le modèle de détection des changements par satellite d'ASE. De plus, l'AEIC pourrait devoir évaluer ses données historiques actuelles pour déterminer si elles répondent aux critères de qualité et à d'autres préoccupations. Si ce n'est pas le cas, il pourrait être utile de se concentrer sur l'expansion ou la modification de sa capacité de gestion des données afin que les données qu'elle recueille maintenant puissent être utilisées à l'avenir.

Expertise – Des experts en IA, des scientifiques des données et d'autres experts, selon le type de projet, sont indispensables à la recherche et au développement d'un système d'IA efficace. Ces experts peuvent être des employés d'un organisme, ou des experts externes qui sollicitent les services d'un organisme ou s'associent à ce dernier pour un projet d'IA particulier. Étant donné que l'AEIC ne dispose pas à l'heure actuelle de l'expertise interne pour les phases de recherche, de conception et de mise en œuvre d'un modèle d'IA, elle devra s'associer à des experts externes ou embaucher du personnel ayant l'expertise nécessaire. Au moment de décider qui formera l'équipe du projet d'IA, l'AEIC devra également évaluer qui sera capable d'exploiter et de maintenir le système d'IA une fois qu'il sera mis en œuvre. Il faudrait envisager d'embaucher des professionnels expérimentés de l'IA qui peuvent contribuer à garantir le bon fonctionnement des systèmes d'IA et à former les autres employés (GoDataDriven, s.d.).

En général, lorsqu'un organisme commence à s'intéresser à l'IA, il s'appuie davantage sur des experts externes, car il n'a pas la capacité, le financement ou les ressources nécessaires pour le faire à l'interne. Toutefois, comme l'organisme sera probablement chargé de la maintenance du projet une fois qu'il sera entièrement mis en œuvre, les connaissances sur le système devraient être transférées des parties externes aux employés de l'organisme pour que ces derniers ne soient pas complètement dépendants d'un tiers (GoDataDriven, s.d.). Si l'AEIC engage un tiers pour effectuer les phases de recherche et de conception pour un système d'IA, elle doit s'assurer que le personnel se familiarise avec le système afin de pouvoir en faire la maintenance.

Partenariats – Comme mentionné précédemment, en matière d'expertise, établir des partenariats avec des tiers peut être une voie pour la recherche et la conception d'un modèle d'IA. Un partenariat avec un laboratoire universitaire ou une organisation axée sur l'IA qui dispose déjà de l'expertise nécessaire en IA peut faciliter la mise en place d'un système d'IA pour un organisme aux ressources limitées. Un partenariat pourrait être utile à l'AEIC, car elle ne dispose pas à l'heure actuelle de l'expertise interne nécessaire pour concevoir un modèle d'IA. Cependant, il est impératif pour un organisme de bien communiquer avec ses partenaires. L'établissement de méthodes de communication qui fonctionnent pour les deux parties – comme des réunions hebdomadaires – serait crucial pour garantir que le modèle d'IA réponde aux besoins de l'AEIC et de ses intervenants et s'intègre à la culture de leur organisme.

Gouvernance – Les considérations en matière de gouvernance devraient également jouer un rôle dans l'adoption des systèmes d'IA. L'application de l'IA et les processus de données connexes doivent respecter la loi. Les organisations qui souhaitent adopter l'IA doivent s'engager à faire preuve de transparence, à faire participer les intervenants de manière sérieuse et à éliminer les biais du système. En outre, il convient d'établir un processus de gestion de la mise en œuvre du projet d'IA qui comprend des responsabilités claires et qui intègre divers points de vue et intervenants pour atténuer les risques [Government Accountability Office des États-Unis, 2021].

À cette fin, le gouvernement du Canada a élaboré des principes directeurs et une évaluation de l'incidence algorithmique afin de déterminer dans quelle mesure les solutions d'IA sont acceptables d'un point de vue éthique et humain (citation). L'AEIC devrait étudier des structures de gouvernance et se

référer à l'évaluation de l'incidence algorithmique aux fins d'orientation dans l'élaboration de tout système d'IA.

Participation des intervenants – En tant qu'organisme gouvernemental, l'AEIC a pour mission de soutenir le développement durable, ce qui requiert une coordination avec des intervenants de l'ensemble du Canada. Comme tout système d'IA mis en œuvre par l'Agence aura une incidence sur le personnel de l'Agence et sur le public, il est impératif de faire participer les intervenants concernés tout au long des phases de conception et d'essai d'un projet d'IA.

Offrir aux Canadiens et aux futurs utilisateurs du modèle la possibilité de donner leur avis pourrait contribuer à accroître la crédibilité du modèle et à faciliter son adoption par les utilisateurs. Cela pourrait également contribuer à garantir l'accessibilité et l'utilité des résultats des modèles d'IA. En outre, l'intégration de divers points de vue tout au long de la phase de conception du modèle peut contribuer à atténuer les risques et à répondre à des préoccupations comme l'éthique et les biais [Government Accountability Office des États-Unis, 2021].

Adoption par les utilisateurs – Les considérations relatives à l'adoption par les utilisateurs font référence aux problèmes que l'AEIC peut rencontrer dans le processus d'adoption. Par exemple, quelle serait la facilité d'utilisation de l'application pour le personnel? Quelle formation serait nécessaire pour maîtriser la nouvelle technologie? Le personnel soutient-il l'application? L'AEIC devrait tenir compte de ces questions lorsqu'elle cherche à réaliser un projet d'IA afin de s'assurer que son personnel y adhère et que la mise en œuvre continue du projet est faisable (GoDataDriven, s.d.). En outre, les préoccupations des employés concernant la perte d'emplois liée à l'adoption de l'IA signifient que l'AEIC doit être prête à introduire et à concevoir des solutions d'IA de manière à souligner dans quelle mesure elles amélioreront l'efficacité et l'efficience des employés et ne supprimeront pas leurs emplois [Chenok, 2018].

Adhésion de la haute direction – L'adhésion de la haute direction et des équipes de direction à l'IA est cruciale pour la réussite de la mise en œuvre d'un projet d'IA. Le soutien de la direction peut influencer sur la capacité d'un organisme à embaucher des experts ou à solliciter leurs services, à obtenir un financement, à accéder aux données et à intégrer la plateforme d'IA dans l'organisme. Ainsi, l'AEIC doit s'assurer que les dirigeants de l'Agence adhèrent à l'IA de manière générale, et à tout projet d'IA particulier que le personnel de l'AEIC envisage de mettre en œuvre. La communication avec les cadres dès le début du projet est de la plus haute importance. La direction peut également éliminer les obstacles aux projets d'IA, soutenir l'adoption de l'IA à grande échelle ou les initiatives multiorganisationnelles et aider à obtenir du financement (GoDataDriven, s.d.).

Financement – Comme le financement peut être un obstacle majeur à la mise en œuvre de l'IA, en particulier dans les organismes gouvernementaux, l'AEIC doit prendre en compte le coût de l'IA. L'Agence devra prévoir un budget pour la recherche initiale, la conception et la mise en œuvre de chaque modèle d'IA, ainsi que des fonds pour soutenir les activités récurrentes d'exploitation et de maintenance. Par conséquent, le budget doit tenir compte du coût des critères susmentionnés et inclure

le financement des partenariats avec des experts externes; la collecte, le stockage et l'analyse des données, notamment; la participation des intervenants; le recrutement et la rémunération d'experts internes; l'adoption par les utilisateurs; et les considérations en matière de gouvernance, entre autres facteurs. Étant donné la nécessité d'un financement continu, l'AEIC devrait continuer à chercher à intégrer un financement consacré à l'IA dans son budget, ou à établir d'autres sources de financement continu, comme des bourses, des subventions ou des programmes de financement de la technologie.

L'AEIC devrait également prendre en compte le coût d'un type de modèle d'IA particulier au moment de choisir de le mettre en place ou non. Comme l'illustrent les cas d'utilisation, les coûts varient selon le type de modèle et l'organisme, de sorte que l'AEIC doit évaluer quels types de projets sont réalisables avec le financement qui lui est accordé.

Considérations et limites juridiques – L'utilisation de l'IA dans les processus postérieurs aux décisions pourrait se heurter à des limites juridiques. En 2021, la Commission européenne a publié un projet de règlement sur l'IA au sens duquel les systèmes d'IA utilisés pour l'application de la loi présentent un risque élevé. Le projet de règlement soumet les systèmes d'IA à risque élevé à diverses exigences, comme l'assurance de la qualité des ensembles de données utilisés pour entraîner les systèmes d'IA, l'application d'une supervision humaine, la création de registres permettant d'effectuer des vérifications de la conformité et la fourniture de renseignements pertinents aux utilisateurs. On estime généralement que la mise en œuvre d'un tel règlement prendra des années. Cependant, il est probable que les pays appliqueront des exigences juridiques aux utilisations de l'IA liées à l'application de la loi et mettront l'accent sur la transparence, l'équité et l'explicabilité.

6 PROCHAINES ÉTAPES

Avant de réaliser les phases de recherche, de conception et de mise en œuvre d'un modèle d'IA pour contribuer aux activités postérieures aux décisions, l'AEIC doit franchir certaines étapes. Un modèle de maturité peut être utilisé pour créer une feuille de route pour la mise en œuvre de l'IA à l'Agence [Panetta, 2019]. GoDataDriven, une société de conseil et de formation en matière de données et d'IA établie à Amsterdam et en activité depuis 2009, a préparé un modèle de maturité qui peut être adapté et utilisé par les organismes pour éclairer leur travail sur l'IA. Dans l'ensemble, un organisme doit tenir compte de sa capacité d'analyse et de son adoption opérationnelle au moment de déterminer les prochaines étapes. La capacité d'analyse concerne l'organisation et l'accessibilité des données, les capacités et les compétences des employés, ainsi que la technologie et les outils. L'adoption opérationnelle renvoie au degré d'intégration de l'IA dans l'organisation, qui est mesuré par le soutien de la direction, le financement et la mise en œuvre. Lorsque l'AEIC envisage l'adoption de l'IA, elle doit prendre en compte tous ces aspects.

Pour aider à cette réflexion, l'élaboration d'un plan stratégique pour l'IA dans l'ensemble de l'AEIC constituerait une bonne première étape. De nombreux organismes ont mis à l'essai et expérimenté des projets en cloisonnement et se rendent compte aujourd'hui de la nécessité d'une approche ou d'une

stratégie à l'échelle de l'organisme afin de continuer à mettre à l'essai, expérimenter et mettre en œuvre des modèles d'IA. Souvent, les gouvernements se retrouvent bloqués après avoir mené des projets pilotes en cloisonnement et se rendent compte que l'IA à grande échelle a des exigences différentes [Van Buren, 2021]. Ainsi, l'élaboration d'une stratégie qui intègre des considérations telles que le soutien de la direction, l'expertise interne, les données, l'éthique, les plateformes technologiques et les partenariats peut aider un organisme à adopter plus efficacement l'IA à l'échelle de l'organisme. L'élaboration de ce plan stratégique avant la conception du modèle d'IA pourrait empêcher l'AEIC de se heurter à des obstacles en cours de route et faire en sorte que les projets puissent être plus facilement étendus, qu'ils soient plus efficaces et qu'ils aient une plus forte incidence au sein de l'Agence. Le plan pourrait également fournir une feuille de route permettant de répondre aux critères susmentionnés, y compris le financement, l'expertise, les données, la gouvernance, et plus encore. L'AEIC devrait donc accorder une attention particulière à la réflexion et au traitement de ces critères dans le plan stratégique.

Comme l'AEIC commence tout juste à s'intéresser à l'IA, elle se trouve dans la phase d'initialisation de la conception de l'IA selon le modèle de maturité de GoDataDriven. Cette phase consiste à déterminer les possibilités en matière d'IA, à préparer les données, le personnel et la technologie pour la mise au point du premier modèle d'IA et à lancer le premier modèle d'IA. À l'heure actuelle, l'AEIC reconnaît que l'IA pourrait être utile et a fait appel à l'ELI pour déterminer les cas d'utilisation, mais elle ne dispose pas d'experts internes en IA, n'a peut-être pas les données historiques ni le financement nécessaire à l'heure actuelle pour développer l'IA au sein de l'Agence et n'a pas encore lancé de modèle d'IA. Après avoir élaboré un plan stratégique, l'AEIC devrait se concentrer sur les étapes de la phase d'initialisation.

À terme, l'AEIC devra prendre des mesures pour passer à la phase suivante du modèle de maturité, soit l'expérimentation continue. À ce moment, l'AEIC devra se concentrer sur le recrutement d'experts internes, la création d'une infrastructure robuste pour l'IA et l'adoption continue d'autres modèles d'IA que ceux conçus pendant la phase d'initialisation.

En résumé, le présent rapport détermine les utilisations potentielles de l'IA qui pourraient être employées par l'AEIC et d'autres organismes environnementaux lors de la réalisation d'activités de conformité, d'application de la loi et de suivi. Il recense six cas d'utilisation de l'IA dans les organismes canadiens, aux États-Unis et dans d'autres pays qui pourraient être pertinents pour les activités postérieures aux décisions de l'AEIC. Il contient également des recommandations quant aux prochaines étapes que l'AEIC pourrait suivre pour intégrer l'IA à son travail afin d'accroître son efficacité et son efficacité.

7 RÉFÉRENCES

Chenok, D. (2018). How can AI help government improve? <https://gcn.com/2018/11/how-can-ai-help-government-improve/300028/>

Deloitte AI Institute. (2021). The Government & Public Services AI Dossier. <https://www2.deloitte.com/us/en/pages/consulting/articles/ai-dossier-government-public-services.html>

GoDataDriven. (2019). The AI Maturity Journey: Learn what it takes to become an AI-driven organization. <https://godatadriven.com/topic/whitepaper-ai-maturity-journeys/>

Government Accountability Office des États-Unis. (2021). Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities. <https://drive.google.com/file/d/13dJ7bYt6Dd5RAcoqeHAt4n9iHXhBXd9D/view>

Greenstone, M. (2018). Using Predictive Analytics to Improve Inspection Targeting. <https://www.ecos.org/wp-content/uploads/2019/10/UChicago-Slide-Deck.pdf>

Martinho-Truswell, E. (2018). How AI Could Help the Public Sector. <https://hbr.org/2018/01/how-ai-could-help-the-public-sector>

Nalbandian, L. (2021). Canada should be transparent in how it uses AI to screen immigrants. <https://theconversation.com/canada-should-be-transparent-in-how-it-uses-ai-to-screen-immigrants-157841>

Panetta, Kasey (2019). CIOs can separate AI hype from reality by considering these areas of risk and opportunity. <https://www.gartner.com/smarterwithgartner/the-cios-guide-to-artificial-intelligence>

Shubhendu, S. et Vijay, J.F. (2013). Applicability of Artificial Intelligence in Different Fields of Life. <https://www.semanticscholar.org/paper/Applicability-of-Artificial-Intelligence-in-Fields-Shubhendu-Vijay/2480a71ef5e5a2b1f4a9217a0432c0c974c6c28c?p2df>

Van Buren, E., Eggers, W. D., Austin, T., Mariani, J. et Kishnani, P. K. (2021). Scaling AI in Government. Deloitte. <https://www2.deloitte.com/us/en/insights/industry/public-sector/government-ai-survey.html>