

The logo for 'obvia' is displayed in a white, lowercase, sans-serif font. The letter 'o' is stylized with a small gap at the top. The background of the entire page is a dark purple gradient with a faint, semi-transparent image of a person's hands typing on a laptop keyboard. Overlaid on the laptop screen are various digital interface elements, including a folder icon with a lock, a bar chart, a list of documents, and a central graphic with the letters 'AI' inside a square frame with radiating lines.

L'IA dans le secteur public Cas d'utilisation et enjeux éthiques

Steve Jacob et Sébastien Brousseau



Auteurs

Steve Jacob est professeur titulaire de science politique à l'Université Laval et directeur scientifique de la collaboration avec les acteurs publics à l'Observatoire international sur les impacts sociétaux de l'intelligence artificielle et du numérique (Obvia).

Sébastien Brousseau est professionnel de recherche à l'Université Laval et coordonnateur du Pôle d'expertise en cybersécurité et impacts sociétaux (Obvia).

Remerciements

Nous remercions les membres de l'axe de recherche Éthique, gouvernance et démocratie de l'Observatoire international sur les impacts sociétaux de l'IA et du numérique (Obvia) qui ont contribué aux ateliers de réflexion éthique que nous avons organisés pour approfondir la compréhension des cas d'application de l'IA dans le secteur public décrits dans ce rapport. Les participantes et participants aux ateliers de réflexion éthique qui ont consenti à ce que leur nom figure dans ce rapport sont : Luc Bégin, Frédéric Bruneault, Charles-Étienne Daniel, Hazar Haidar, Roxanne Lépine, Emmanuelle Marceau, Allison Marchildon, Aude Marie Marcoux, Dominic Martin, Frédérique Romy-Godin, Andréane Sabourin Laflamme, Jean Rouat et Joé Trempe-Martineau.

Le contenu de ce rapport n'engage que ses auteurs.



Pavillon Charles-De Koninck, local 2489
1030, avenue des Sciences-Humaines
Université Laval
Québec (Québec) G1V 0A6

Produit avec le soutien financier du ministère de la
Cybersécurité et du Numérique du Québec

**Cybersécurité
et Numérique**

Québec 

ISBN : 978-2-925138-45-7

Jacob, S., & Brousseau, S. (2024). *L'IA dans le secteur public : Cas d'utilisation et enjeux éthiques*. Observatoire international sur les impacts sociétaux de l'IA et du numérique. <https://doi.org/10.61737/FCXM4981>

Table des matières

1. Introduction	5
<hr/>	
2. Méthodologie	8
<hr/>	
3. Catégorisation des SIA dans le secteur public	12
<hr/>	
4. Cas d'application de l'utilisation de l'IA dans le secteur public	21
La Bonne Boîte	22
GoSafe	26
Chorus	31
ABLE	37
ToxCast	44
SKALA	51
Surtrac	57
Corti	63
Tengai	70
AuroraAI	76
SSOSA	86
<hr/>	
5. Conclusion	97
<hr/>	
6. Bibliographie	101

Liste des tableaux

Tableau 1.	Catégorisation des SIA dans le secteur public	14
Tableau 2.	Description synthétique de La Bonne Boîte	22
Tableau 3.	Description synthétique de GoSafe	26
Tableau 4.	Description synthétique de Chorus	31
Tableau 5.	Description synthétique d'ABLE	37
Tableau 6.	Description synthétique de ToxCast	44
Tableau 7.	Description synthétique de SKALA	51
Tableau 8.	Description synthétique de Surtrac	57
Tableau 9.	Description synthétique de Corti	63
Tableau 10.	Description synthétique de Tengai	70
Tableau 11.	Description synthétique d'AuroraAI	76
Tableau 12.	Description synthétique de SSOSA	86

1. Introduction

1. Introduction

L'introduction de l'intelligence artificielle (IA)¹ dans l'administration publique est en pleine expansion et représente la phase plus récente de la transformation numérique. L'utilisation de systèmes d'IA est souvent associée à l'amélioration de l'efficacité et de la productivité, car elle permet d'exécuter rapidement des tâches routinières, de réaliser des activités à une plus grande échelle et de rendre les services accessibles à un plus grand nombre de personnes (Etscheid, 2019 ; Toll et al., 2019). Cette technologie a donc le potentiel de transformer plusieurs aspects du travail administratif comme la prise de décision automatisée, l'analyse prédictive, le traitement de documents et l'interaction avec les citoyens².

Cependant, les études indiquent également que le déploiement³ et l'utilisation des solutions d'IA ne sont pas sans poser des défis aux décideurs, gestionnaires et employés du secteur public (Wirtz et al., 2019 ; Champion et al., 2020). L'automatisation des tâches et des processus suscite un vaste débat aussi bien dans le milieu professionnel que scientifique. Les interrogations portent notamment sur les effets induits par l'IA sur le fonctionnement des organisations ainsi que sur l'avenir et la qualité de la relation entre les organismes publics et les citoyens (Zerilli, 2021). D'autres défis d'ordre sécuritaire, législatif ou encore éthique sont également mentionnés dans les travaux portant sur l'introduction de l'IA dans le secteur public.

Au Québec comme ailleurs, des organisations publiques ont commencé à se doter de cette technologie depuis plusieurs années. À l'heure actuelle, elles sont nombreuses à vouloir tirer avantage de l'évolution rapide de l'IA pour moderniser les services offerts aux citoyens, améliorer la qualité des services et accroître leur performance ou leur agilité (Jacob & Souissi, 2022). Parallèlement, différentes instances locales, nationales et internationales élaborent des documents d'orientation et d'encadrement concernant le développement et l'utilisation de ces technologies. En juin 2021, le Gouvernement du Québec a dévoilé sa *Stratégie d'intégration de l'intelligence artificielle dans l'administration publique*. Avec cette stratégie, le Gouvernement du Québec entend profiter pleinement du potentiel de l'IA comme outil d'optimisation et de modernisation des processus administratifs. Le déploiement de cette stratégie vise à cibler les possibilités les plus prometteuses en matière d'amélioration des services publics et balise les travaux pour que l'évolution de ces outils technologiques se fasse de façon responsable et sécuritaire⁴.

À l'heure où les technologies d'intelligence artificielle occupent une place grandissante dans la vie des citoyens et que les administrations publiques introduisent progressivement ces technologies au sein de leurs services, il semble que les fonctions et possibilités associées aux systèmes d'IA (SIA)⁵ ne sont pas toujours clairement connues ou comprises par les décideurs, les usagers ou les employés du secteur public. La présente recherche a pour but de décrire des cas concrets d'utilisation de l'IA dans les organisations publiques de pays de l'OCDE. L'intention est d'enrichir la connaissance et la compréhension des potentialités de l'IA en présentant un éventail de systèmes aux capacités techniques variées.

1 Il existe de nombreuses définitions de l'IA. Dans ce rapport, l'IA désigne « l'ensemble des techniques qui permettent à une machine de simuler l'intelligence humaine, notamment pour apprendre, prédire, prendre des décisions et percevoir le monde environnant ». *Déclaration de Montréal pour un développement responsable de l'IA*, 2018, p.19.

2 *White House launches artificial intelligence information portal*, 2021.

3 Dans ce rapport, nous privilégions l'utilisation du terme « déploiement » (Ensemble des opérations consistant à introduire un nouveau produit informatique dans son environnement d'utilisation prévu, à grande échelle et de façon coordonnée. OQLF, 2021) et faisons un usage limité du terme « implémentation » (Opération qui consiste à réaliser la phase finale d'élaboration d'un système, afin de le rendre fonctionnel. OQLF, 2003). La notion de « déploiement » nous apparaît plus représentative de l'ensemble du processus d'introduction d'un système d'intelligence artificielle dans une organisation. L'OQLF précise en ce sens qu'en plus de l'installation du produit, du réglage de ses paramètres et de son intégration aux systèmes existants, le « déploiement » comprend la planification et le soutien qui lui sont afférents, par exemple la mise sur pied du calendrier et la formation des utilisateurs.

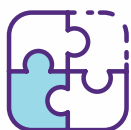
4 Secrétariat du Conseil du trésor, *Stratégie d'intégration de l'intelligence artificielle dans l'administration publique 2021-2026*, 2021.

5 Nous utilisons l'appellation systèmes d'intelligence artificielle (SIA) dans le sens que lui donne Andrew Critch du MILA dans le texte *AI Research Considerations for Human Existential Safety (ARCHES)*: « By "AI system", we refer to any collection of one or more automated decision-making units. The units are not assumed to be cooperating or competing, and are not assumed to have been created by cooperating or competing stakeholders. Hence, the term "system" is intentionally general and agent-agnostic, and is meant to encompass simple and complex artifacts of engineering that could variously be called "decision-making tools", "agents", "multi-agent systems", "societies of machines", or none of the above » (Critch, 2020, p. 12).

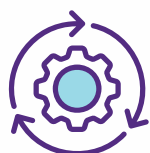
Cet exercice nous amène à recenser et documenter des applications en vue de répondre aux questions suivantes :



Quel est l'objectif poursuivi par le SIA ?



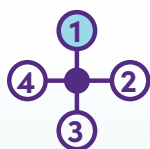
Quelles sont les étapes de développement du SIA ?



Comment fonctionne le SIA ?



Qui est le bénéficiaire direct du travail du SIA ?



Quel est le degré d'autonomie du SIA ?



Quelles sont les préoccupations éthiques soulevées par le déploiement et l'utilisation du SIA ?

Pour répondre à la dernière question relative aux enjeux éthiques des SIA, nous avons organisé des groupes de discussion avec des membres de l'axe de recherche Éthique, Gouvernance et Démocratie de l'Observatoire international sur les impacts sociétaux de l'IA et du numérique (Obvia).

La présentation des cas et des questions éthiques qu'ils soulèvent permettront une meilleure compréhension des implications du déploiement des solutions d'IA et contribueront à déterminer les pratiques prometteuses et les pistes d'action pour favoriser le succès d'initiatives québécoises d'intégration de l'IA dans les organisations publiques.

2. Méthodologie

2. Méthodologie

Premier volet

Le **premier volet** de cette recherche repose sur l'étude d'une douzaine de cas de déploiement de systèmes d'intelligence artificielle par des organisations publiques de pays de l'OCDE. L'étude de cas est une stratégie de recherche utilisée en sciences sociales pour étudier les phénomènes complexes dans leurs contextes réels. Elle est mobilisée dans les recherches de type exploratoire en vue de comprendre le fonctionnement d'un phénomène à travers une plongée dans ses éléments constitutifs (Mucchielli, 2007). La méthode de l'étude de cas porte sur quelques cas, voire un seul, sur lesquels on recueille une grande quantité d'informations suivant différentes méthodes de recherche. Le cas peut être un individu, un groupe, un événement, une institution, etc. (Latzko-Toth, 2009). Dans la présente recherche, les cas étudiés sont des instances de déploiement de systèmes d'intelligence artificielle (SIA) par des organisations publiques.

Bien que le secteur public accuse un certain retard sur le privé en ce qui concerne l'utilisation de l'IA, son usage y est de plus en plus répandu et il existe une grande quantité d'exemples d'utilisation de cette technologie dans des organisations publiques de pays de l'OCDE. Notre objectif étant de présenter un éventail varié d'applications de l'IA, nous avons d'abord développé un outil de catégorisation des SIA de façon à faciliter la sélection de cas à documenter. Cette catégorisation composée de deux dimensions (*Bénéficiaires directs du travail du SIA* et *Degré d'autonomie du SIA*) est présentée en détail dans la prochaine section du rapport.

La documentation des cas s'est faite à partir de recherches bibliographiques et sur Internet. Notre travail repose uniquement sur des informations relevant du domaine public. La nature des documents consultés est variée et comprend : la littérature scientifique concernant des cas de déploiement de SIA ou concernant des technologies d'IA spécifiques, les documents publiés par les gouvernements ou par des organismes faisant la promotion ou s'intéressant à l'IA, les sites internet d'entreprises privées fournisseurs de solutions d'IA, les sites Internet de centres de recherche ou de laboratoires développant des SIA, etc.

Lorsque cela était possible, nous avons fait usage d'une variété de sources pour la description de chaque cas, accordant davantage d'attention aux articles scientifiques lorsqu'ils existaient. Dans les cas pour lesquels l'information disponible provenait principalement de fournisseur de solution d'IA et/ou d'organisations publiques intéressées, nous avons porté une attention particulière à la distinction entre l'information factuelle et l'activité promotionnelle, de façon à faire une présentation la plus neutre possible de chaque cas.

Pour qu'un cas soit sélectionné, nous devons être en mesure de documenter avec une certaine précision les éléments suivants :

Objectif poursuivi par l'organisation en déployant le SIA

Développement du SIA

- Lieu (pays, région, municipalité, etc.), niveau de gouvernement, administration concernée, besoin du client et étapes de développement;
- Développeur et processus d'acquisition (interne ou appel au privé) ;
- Année du déploiement et statut actuel du SIA (phase de test, prêt à être utilisé, système stable ou ajustements nécessaires, etc.).

Fonctionnement du SIA

- Que fait le SIA ?
- Comment fonctionne le SIA ?
- Quel est l'impact souhaité ou observé de l'utilisation du SIA pour le secteur public qui l'utilise ?
- Qui bénéficie directement du travail du SIA ?
- Quel est le degré d'autonomie du SIA ?
- Quel rôle l'humain remplit-il dans le fonctionnement du SIA ?

Nous avons volontairement exclu les cas trop « problématiques », ayant fait l'objet de contestations ou de critiques publiques comme ce fut le cas aux Pays-Bas où un scandale a entraîné la chute du Gouvernement en raison d'un algorithme ayant conduit à accuser injustement des milliers de familles de fraude fiscale liée à leurs allocations sociales. De plus, la mise en évidence des enjeux éthiques ou des effets positifs et négatifs du fonctionnement et de l'utilisation du SIA n'était pas au centre de nos préoccupations pour ce premier volet du projet de recherche. La réflexion concernant les enjeux éthiques et les effets ou conséquences du déploiement et de l'utilisation des applications d'IA présentées est l'objet du second volet de notre projet de recherche.

Second volet

Pour ce **second volet**, le rapport présente le résultat d'ateliers de discussion avec des membres de l'axe de recherche Éthique, Gouvernance et Démocratie de l'Observatoire international sur les impacts sociétaux de l'IA et du numérique (Obvia). Se déroulant au cours de l'été 2023, trois ateliers semi-dirigés ont permis de réunir des participants aux profils variés et aux intérêts de recherche diversifiés : médecine sociale et préventive, droit, éthique appliquée, génie informatique et génie logiciel, neurosciences, bioéthique, éthique organisationnelle, etc.

Chacune des trois rencontres a permis de discuter de quatre cas d'application :

- les ateliers se sont déroulés entre le 15 juin et le 5 juillet 2023 ;
- les cas d'application analysés lors des ateliers étaient répartis par type de bénéficiaires. Le premier atelier portait sur les applications d'IA bénéficiant directement aux usagers. Le deuxième atelier concernait les applications bénéficiant aux employés. Le troisième atelier réunissait les applications qui bénéficient principalement aux organisations ;
- un atelier réunissait 5 à 7 participant.e.s ;
- la durée des ateliers est comprise entre 90 à 120 minutes.

Afin de maximiser le temps alloué aux échanges, les participants ont reçu à l'avance les documents de présentation des quatre cas visés par chacun des ateliers, ainsi qu'un court texte permettant d'orienter la discussion autour des thèmes et questions suivants :

Thème 1 : L'application d'IA en général

1. Quels sont vos **constats généraux** relatifs à cette application d'IA (ex. inquiétudes, craintes) ?
2. Quels sont les **avantages et les inconvénients** de cette application d'IA pour les différentes parties prenantes (ex. usager, employé, organisation, etc.) ?

Thème 2 : Les considérations éthiques

1. L'application d'IA soulève-t-elle des **enjeux ou préoccupations éthiques** ?
2. Des valeurs sont-elles en **tension** dans le(s) cas présenté(s) ? Si oui,
 - quelle est la source de la tension ?
 - quelles sont les conséquences de cette tension et pour quelles parties prenantes ?
 - quelles solutions peut-on envisager pour surmonter la tension ?
3. Quel(s) **principe(s) éthique(s) faudrait-il réaffirmer** pour tendre vers un développement et une utilisation responsables de cette application ?
4. Quels outils et/ou mécanismes seraient nécessaires pour appuyer les personnes et l'organisation en vue d'une évaluation et utilisation adéquate de cette application ?

Les ateliers ont été enregistrés (avec le consentement des participants), retranscrits, anonymisés et synthétisés pour en analyser les principaux thèmes. Afin de faciliter la lecture, les synthèses des propos recueillis au cours des ateliers ont été incorporées à la présentation de chacun des cas d'application (Section 6 – Pistes de réflexions éthiques).

Il est à noter que notre rapport met en lumière les principales préoccupations éthiques qu'il serait pertinent de considérer dans l'éventualité du déploiement, au Québec, de SIA similaires à ceux présentés, et qu'un travail d'analyse plus approfondie reste à faire pour effectuer un constat plus complet et concret en lien avec la réalité politico-administrative québécoise.

De plus, pour cette section du rapport, nous repons sur les témoignages d'experts qui utilisent parfois des termes distincts pour parler de concepts similaires. Une certaine uniformisation terminologique sera nécessaire, mais celle-ci n'est pas l'objet du présent rapport.

3. Catégorisation des SIA dans le secteur public

3. Catégorisation des SIA dans le secteur public

Les systèmes d'intelligence artificielle (SIA), qu'ils soient déployés dans la sphère privée ou publique, s'insèrent dans une structure déjà existante. Ils s'immiscent dans un ensemble de relations organisationnelles et sont donc en relation avec différents éléments, composantes, acteurs et systèmes déjà présents lors de leur introduction. Étant donné la diversité des usages et impacts de l'IA dans le secteur public, nous proposons de répartir les cas d'application selon une catégorisation en deux dimensions nous permettant de classer les différents cas de SIA selon :

- Les bénéficiaires directs du travail du SIA
- Le degré d'autonomie du SIA

En classant les systèmes selon le *bénéficiaire direct* du travail du SIA d'un côté, et du *degré d'autonomie* du SIA de l'autre, nous serons en mesure d'observer et de qualifier la composante relationnelle particulière qui émerge à l'intérieur des organisations dans lesquelles les systèmes s'insèrent.

La première dimension, *Bénéficiaires directs*, est composée de trois éléments :

- 1 Usager
- 2 Employé
- 3 Organisation

La seconde dimension, *Degré d'autonomie*, est divisée en quatre niveaux :

- 1 Automatique
- 2 Autonomie limitée
- 3 Semi-autonome
- 4 Autonome

Il est à noter que la catégorisation que nous proposons est d'ordre idéal-typique, c'est-à-dire qu'il s'agit d'une représentation conceptuelle nous permettant de distinguer les systèmes étudiés en étant conscients qu'il peut y avoir, dans certains cas, des chevauchements entre des caractéristiques (Coenen-Huther, 2003, p. 531). Ceci implique qu'il y aura toujours des cas limites, difficiles à catégoriser précisément, pour lesquels des réflexions et discussions supplémentaires seront nécessaires. Cette catégorisation vise à susciter ces réflexions et discussions. Elle est flexible, mais assez circonscrite et précise pour être à même de démarrer ce processus de réflexion sur des bases solides.

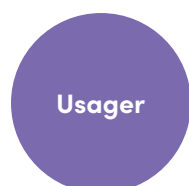
Tableau 1. Catégorisation des SIA dans le secteur public

DEGRÉ D'AUTONOMIE	BÉNÉFICIAIRES DIRECTS		
	Usager	Employé	Organisation
Degré 1 - Automatique	La Bonne Boîte	GoSafe	Chorus
Degré 2 - Autonomie limitée	ABLE	ToxCast	SKALA
Degré 3 - Semi-autonome	Surtrac	Corti	Tengai
Degré 4 - Autonome	AuroraAI	SSOSA	SSOSA et AuroraAI

Bénéficiaires directs du travail du SIA

Les SIA remplissent des fonctions qui profitent à différents bénéficiaires. Dans le secteur public, les principaux bénéficiaires sont : les usagers, les employés et les organisations. Pour chacun des cas potentiels, nous nous sommes demandé qui bénéficie directement du fonctionnement du SIA. Est-ce l'utilisateur, bénéficiant de la fonction du SIA dans le cadre de son interaction avec un organisme public ? Est-ce l'employé, le SIA ayant pour tâche de l'assister ou de le suppléer dans son travail ? Ou est-ce l'organisation elle-même, le SIA jouant un rôle dans l'orientation et/ou la direction de l'organisme, du ministère, etc.⁶

En indiquant les acteurs qui bénéficient principalement du travail effectué par les SIA, cette catégorisation donne une idée de la nature et de la direction de la relation induite par un système donné à l'intérieur d'un organisme public. Afin de permettre l'exploration de cette dimension relationnelle, nous considérons et présentons les SIA comme des acteurs ou des agents dont certains manifestent une forme embryonnaire, mais de plus en plus concrète d'intelligence. Nous pourrions ainsi observer l'espace qu'ils occupent et le travail concret qu'ils effectuent au sein de leur domaine d'intervention respectif.



« L'utilisateur du service public est la personne physique ou morale qui a recours individuellement, personnellement, effectivement et directement à une prestation » (Morio, 2014b, p. 515). L'utilisateur peut donc être un individu, une entreprise ou un organisme ayant recours à toute prestation de service public allant de la simple demande d'information, à l'octroi d'une subvention ou d'un service, en passant par l'assistance pour remplir un formulaire.

Puisque l'utilisateur, souvent un citoyen, entre en relation avec un service public dans le but d'avoir recours à une prestation, la fonction des SIA dont bénéficie l'utilisateur est avant tout d'interagir avec lui. Les fonctions générales des systèmes concernent l'amélioration de l'accès à l'information pour l'utilisateur et la délivrance des services publics aux citoyens et entreprises. Les SIA peuvent également contribuer à l'amélioration et l'innovation dans la production et la fourniture des services (van Noordt & Misuraca, 2022, p. 2-4).

6 Le bénéficiaire direct du travail d'un SIA peut parfois être difficile à établir ou être matière à débat. Il ne s'agit pas, dans le cadre de notre catégorisation, de déterminer le travail de qui ou de quoi un SIA donné a pour fonction d'assister ou de remplacer, mais bien de déterminer qui bénéficie directement du travail du SIA.

Par exemple: dans le cas d'un *chatbot* ayant pour fonction de traiter les requêtes simples des citoyens d'une grande ville, bien que le SIA remplace un employé municipal, le bénéficiaire direct est l'utilisateur (le citoyen) qui reçoit réponse à sa question ou se voit rediriger vers le service pertinent. Si un système similaire est destiné à assister l'employé municipal dans son interaction avec le citoyen (un *chatbot* assistant l'employé dans sa réponse à une requête citoyenne, par exemple), le bénéficiaire direct sera alors l'employé lui-même, le citoyen bénéficiant indirectement de l'assistance apportée à l'employé par le système.

Employé

En ce qui concerne la fonction publique, les employés sont pour la plupart des fonctionnaires, des agents publics titulaires, définis comme des agents nommés, occupant un emploi permanent et titularisé dans un grade de la hiérarchie administrative (Morio, 2014a, p. 234). La notion d'agents publics renvoie également aux agents qui, même s'ils sont employés par une personne publique, ne sont pas titulaires et/ou permanent au sein de leur organisation. La catégorie « Employé » regroupe les personnes qui occupent des postes dans l'ensemble des échelons hiérarchiques y compris la direction.

Les SIA bénéficiant à l'employé occupent généralement des fonctions liées à la gestion des affaires internes d'une organisation, ainsi qu'à l'assistance et l'efficacité du travail des employés. Les fonctions suivantes en font partie : amélioration du traitement de l'information, accélération du traitement des dossiers, amélioration de l'attribution des dossiers, substitution de la main-d'œuvre, réduction de la bureaucratie (Medaglia et al., 2023, p. 129).

Organisation

Dans le secteur public, l'organisation désigne « la structure des services administratifs et sa modification éventuelle en fonction des objectifs et des politiques publiques » (Guglielmi, 2014, p. 358). Ce bénéficiaire peut être une organisation dans son ensemble ou une partie de celle-ci (ex. un service, une direction, un département). Les SIA dont le travail bénéficie directement à l'organisation occupent généralement des fonctions liées à l'analyse des situations et problèmes qui sont de sa responsabilité, l'établissement et le suivi des orientations et des règles. Les fonctions suivantes en font partie : détecter, documenter et/ou contribuer à la résolution des problèmes relevant de la responsabilité d'un organisme public ; améliorer les décisions de politique publique (et estimer les effets potentiels de la politique) ; contrôler la mise en œuvre des politiques (et évaluer les politiques existantes) ; renforcer la participation des citoyens à l'élaboration des politiques (van Noordt & Misuraca, 2022, p. 2-4).

Degré d'autonomie du SIA

Pour la seconde dimension, nous classons les systèmes selon leur degré d'autonomie dans l'accomplissement de leurs tâches⁷. Dans le contexte de notre classification, le degré d'autonomie réfère à la capacité plus ou moins grande d'un SIA à fonctionner de manière indépendante et à prendre des décisions sans intervention humaine. L'autonomie implique la capacité de percevoir, de raisonner et d'agir dans un environnement physique et/ou virtuel, dans le but d'atteindre un objectif ou d'accomplir une tâche. La notion d'autonomie est éminemment relationnelle, étant lié aux interactions sociales et à celles, plus générale, avec l'environnement (Selten & Meijer, 2021). Le degré d'autonomie fait référence à l'ampleur du pouvoir discrétionnaire⁸ que possède un SIA en vue d'atteindre son objectif ou d'accomplir sa tâche. Ce pouvoir dépend de ce que l'organisation qui déploie le SIA est prête à lui conférer et des capacités technologiques disponibles au moment de son développement. L'ampleur du pouvoir discrétionnaire du SIA pourra ensuite évoluer au gré des usages et des transformations organisationnelles entraînées par son déploiement⁹.

7 Il ne s'agit pas ici du modèle par *niveau d'autonomie* (0-5) pour les voitures autonomes (Kaber, 2018, pp. 8-9) ou de la taxonomie par *level of automation* de Sheridan et Verplank (Parasuraman et al., 2000). Ces modèles s'intéressent surtout à l'impact de différents niveaux d'automatisation sur l'opérateur humain en termes de responsabilité, d'implication et d'attention de ce dernier, et au degré d'automatisation souhaitable selon le contexte.

8 « Le pouvoir discrétionnaire permet à un agent public de prendre une décision en fonction d'un éventail de possibilités. L'agent public mobilise alors son jugement professionnel afin de tenir compte des circonstances et du contexte dans lequel il intervient. La littérature en administration publique emploie également le terme « discrétion administrative » pour parler de cette autonomie dont jouissent les agents dans leur travail. [I]l arrive que des ordinateurs ou des machines remplacent les humains dans la prise de décision en disposant eux-mêmes d'un pouvoir discrétionnaire. Dans ce cas, nous parlerons de discrétion numérique ». Dans Steve Jacob et Seïma Souissi (2020) *La fourniture de services publics à l'ère numérique : Évolution du rôle et des compétences des employés de première ligne*. Cahiers de recherche sur l'administration publique à l'ère numérique [# 02]. Dans le présent rapport, nous utilisons le terme *pouvoir discrétionnaire* et non *discrétion numérique* puisque nous désirons insister sur le partage de pouvoir discrétionnaire entre les SIA d'un côté, et les usagers, employés et organisations de l'autre.

9 Nous nous intéressons davantage à ce que le système fait concrètement et moins à ses capacités techniques particulières, et ceci pour deux raisons principales : premièrement, une organisation publique peut décider de faire usage de certaines fonctionnalités d'un système et d'en garder certaines autres latentes. Le degré d'autonomie étant déterminé à partir de la relation observée entre le système et son environnement, notre intérêt porte sur l'autonomie manifeste du système et non sur son autonomie potentielle (sa capacité technique). Deuxièmement, il y a souvent une différence entre ce pour quoi un système est conçu et l'usage qui en est fait. Le degré d'autonomie étant déterminé à partir de l'observation du système en action, les différents usages qui en sont faits à travers le déploiement, l'appropriation par l'organisation, et l'utilisation par toutes les parties concernées sont prises en compte.

Dans le contexte de notre catégorisation, le degré d'autonomie varie d'*automatique* à *autonome* : les systèmes automatiques ne manifestent aucune autonomie, car leurs *outputs* ou actions reposent strictement sur leur programmation (sur les personnes responsables de celle-ci), tandis que les systèmes autonomes peuvent s'adapter à leur environnement, fonctionner de manière indépendante avec une implication humaine minimale et manifester une capacité de prise de décision leur conférant un pouvoir discrétionnaire important (White & Searle, 2015, pp.60-64).

L'enthousiasme récent pour l'IA, en tant qu'apprentissage automatique, signale un changement dans la perception et l'utilisation des technologies numériques. Il faut toutefois garder à l'esprit qu'en termes d'utilisation par différentes administrations publiques, l'apprentissage automatique s'inscrit dans une trajectoire de technologies, de systèmes et d'algorithmes numériques, plutôt que dans une rupture fondamentale (Henman, 2020, p. 210). Par exemple, bon nombre des préoccupations juridiques, éthiques et sociales liées à l'utilisation de l'apprentissage automatique sont similaires à celles qui émergent lorsque des algorithmes moins complexes sont utilisés. C'est pourquoi il nous a semblé important d'utiliser une catégorisation large, sous forme de continuum, laissant place aux systèmes *automatiques* (automatisation des tâches, systèmes experts, etc.), souvent ignorés dans les discussions contemporaines concernant l'IA (Ng & Alarcon, 2020, pp. 79-81), ainsi qu'aux systèmes complètement autonomes, encore inexistant, mais nous permettant d'entrevoir l'horizon des développements à venir.

Bien que notre catégorisation soit large, elle ne concerne pas la distinction entre *artificial narrow intelligence* (ANI) et *artificial general intelligence* (AGI) : comme il sera explicité dans la description des degrés d'autonomie, chacun de ceux-ci correspondent à l'ANI dans le sens où ils réfèrent tous à des capacités techniques limitées à des tâches particulières et non généralisables ou supérieures à l'ensemble des capacités de l'intellect humain (Harrison & Luna-Reyes, 2022).

La catégorisation par *degré d'autonomie* est un exercice de distinction des SIA selon plusieurs éléments qui ont chacun leur importance : adaptabilité, capacités de prise de décision (Leyer et al., 2020), pouvoir discrétionnaire (Peeters, 2020 ; Young et al., 2019), agentivité (Mykhailychenko, 2019, pp. 99-120 ; Russell et al., 2010, pp. 39-40), type de techniques utilisées (European Commission. Joint Research Centre, 2020, p. 16), mémoire (Desouza et al., 2020, pp. 2-3 ; Kim et al., 2022)¹⁰, fonctions (Pereira & Lopes, 2020, pp. 30-31 ; van Noordt & Misuraca, 2022, pp. 2-4), etc. Des classifications plus circonscrites et moins multifactorielles existent¹¹, comme la catégorisation des SIA par type de technologie employée, par leurs fonctions ou par secteur d'intervention publique dans lesquels ils sont employés. Notre effort de catégorisation vise plutôt à inclure une multiplicité de facteurs afin de représenter de façon intuitive et englobante la composante relationnelle associée au déploiement et à l'utilisation de systèmes d'IA dans des organisations publiques. C'est donc au prix d'une certaine imprécision en ce qui concerne la fine ligne séparant chaque degré d'autonomie que nous proposons d'apporter un éclairage sur les dynamiques d'agentivité, de prise de décisions et de contrôle, liées à l'introduction des différents systèmes. Considérant le caractère éminemment social du secteur public, l'exploration de ces dimensions relationnelles, voire politiques, de l'introduction des différents systèmes nous semble prometteuse pour stimuler la réflexion éthique (Selten & Meijer, 2021 ; Veale & Brass, 2020 ; Bovens & Zouridis, 2002 ; Lehner & Knoll, 2022, pp. 39-41). La notion de « pouvoir discrétionnaire » nous semble importante puisque les nouveaux partages de pouvoir et de contrôle que les SIA introduisent sont au cœur des potentielles préoccupations soulevées par leur déploiement. En effet, étant déployés par des organisations publiques, ces systèmes effectuent des tâches qui sont souvent directement liées à l'offre de services aux citoyens et, dans une certaine mesure, à l'exercice d'un pouvoir et d'un contrôle sur ces derniers.

10 « Memory is essential for AI systems because it enables them to learn from and adapt to their environment. By incorporating memory, AI can make more informed decisions, refine its understanding of the world, and evolve over time. This adaptability is particularly important when addressing complex, dynamic problems or when an AI system must interact with unpredictable environments ». Source: *Memory in Artificial Intelligence: The Key to Advanced Learning* - Cow-Shed Startup, 2023, en ligne : <https://www.cow-shed.com/blog/memory-in-artificial-intelligence-the-key-to-advanced-learning>.

11 Voir par exemple: *Artificial Intelligence in Government: Taking Stock and Moving Forward* (Medaglia et al., 2023, pp. 126-132) ; *Responsible AI and Ethical Issues for Businesses and Government* (Vassileva & Zwilling, 2021, pp. 6-11) ; *AI Watch: European landscape on the use of artificial intelligence by the public sector* (Joint Research Centre., 2022, pp. 30-34) ; *What Do You Mean by "AI"?* (Wang, 2008).

3. Catégorisation des SIA dans le secteur public

Il nous semble qu'une classification strictement technique des systèmes aurait réduit l'analyse à ce que *sont* les systèmes, et à l'impact potentiel sur la sphère publique selon le type d'algorithme utilisé. Or, nous nous sommes davantage intéressés à ce que les systèmes *font* (les tâches qu'ils effectuent concrètement) et moins à ce qu'ils *sont* (leurs caractéristiques strictement techniques). Nous sommes conscients que la dimension technique est importante puisqu'elle est à la source des capacités d'un système et de sa possibilité de manifester un degré d'autonomie ou un autre. C'est pourquoi la description technique des systèmes occupera une partie importante de la présentation des applications. Le niveau de précision et de détail peut cependant varier en fonction de la disponibilité de l'information pour chacun des cas présentés. Il nous semble tout de même que la relation n'est pas directe entre les caractéristiques techniques d'un système et son degré d'autonomie. Plusieurs facteurs viennent préciser et expliciter le degré d'autonomie, dont la composante relationnelle sur laquelle nous insistons (Nickles et al., 2004, pp.41-53). Chaque SIA est constitué et fait partie d'autres systèmes, qu'ils soient techniques, politiques, économiques, légaux, culturels, etc. C'est l'insertion des SIA dans cet éventail de structures préexistantes qui nous intéresse particulièrement et qui informe le degré d'autonomie effectif de chaque SIA. Que ce soit au niveau de l'interaction, du contrôle ou de la collaboration, l'introduction de technologie dite « intelligente » possédant différents degrés d'autonomie et bénéficiant à différents acteurs de la sphère publique, favorise l'émergence de nouvelles réalités que nous tenterons d'éclairer par les cas étudiés.

Les cas présentés sont intentionnellement variés et ne représentent en rien une gradation, en termes de degré d'autonomie, de système aux objectifs similaires. Comme nous l'avons déjà précisé, chaque SIA doit être considéré avant tout à l'égard de ce qu'il fait, du travail concret qu'il accomplit à l'intérieur de l'organisation publique qui le déploie. L'envergure du système ou l'étendue de son impact n'est pas ici au centre de nos préoccupations, et n'est donc pas un critère de sélection des cas présentés. C'est ainsi qu'un SIA qui centralise les informations relatives aux finances publiques d'un État peut manifester un degré d'autonomie minimale, même si son impact et son envergure sont grands, tandis qu'un SIA qui coordonne les feux de circulation des artères principales d'une ville peut fonctionner avec un degré d'autonomie élevé, même si son envergure est moindre : si l'importance du second système est moindre en termes d'impact, son degré d'autonomie à l'égard de sa tâche reste plus grand.



Les SIA automatiques sont basés sur des séquences d'activités prescrites, fixées à l'avance. Les actions et 'décisions' de ces systèmes reposent entièrement sur leur programme et il s'agit de systèmes qui font preuve de peu de flexibilité (Chatila et al., 2021, p. 26-27). Nous y retrouvons surtout des logiciels automatisés dépourvus de processus décisionnels *cognitifs*. Ils exécutent des tâches répétitives telles que l'ouverture et le classement de courriels et de pièces jointes, l'extraction de données sur le web, et quantité d'autres opérations plus ou moins complexes permettant d'économiser temps et main-d'œuvre (Harrison & Luna-Reyes, 2022, p. 496). Bien que certains des systèmes appartenant à cette catégorie fassent usage d'algorithmes d'apprentissage automatique leur permettant de classer des informations et d'automatiser des tâches (Veale & Brass, 2020, pp. 4-5 ; Peeters, 2020, pp. 512-513), ils reposent généralement sur les algorithmes 'classiques' ou 'symboliques' qui ont dominé la recherche sur l'IA du milieu des années 1950 jusqu'à la fin des années 1980. Ils constituent des SIA rudimentaires, mais sont encore largement utilisés de nos jours, bénéficiant de l'accroissement de la puissance de calcul et des données disponibles. Il ne s'agit pas, pour ces SIA, d'apprendre, d'innover ou de faire preuve d'initiative dans la façon d'accomplir des tâches, mais d'*automatiser* des tâches existantes (Rao & Verweij, 2017). Ainsi, même si le produit (*output*) de ces SIA peut être perçu et interprété comme une décision prise par ces systèmes, ce produit relève complètement et uniquement de la programmation, et donc des choix humains concernant la marche à suivre et le résultat recherché (Young et al., 2019, pp. 302-305). C'est pourquoi, dans le contexte de notre catégorisation, les SIA automatiques ne possèdent aucune autonomie décisionnelle (aucun pouvoir discrétionnaire), même si leur *output* automatique n'est pas sans conséquence pour l'utilisateur.

Ces systèmes d'IA s'appuient sur les données qui leur sont fournies et des règles préprogrammées pour effectuer des tâches (Nickles et al., 2004, p. 45)¹². Ils ne peuvent pas modifier leur fonctionnement de manière autonome ni s'adapter aux circonstances changeantes de leur environnement. Les exemples incluent les *rule-based expert systems*, l'*automatisation robotisée des processus (robotic process automation ou RPA)*, les *chatbots* simples et, de façon générale, les algorithmes de classification de type 'if-then' (Osman, 2019)¹³. Les systèmes dont la tâche est de comparer les caractéristiques d'un cas et de le situer par rapport à un ensemble de cas dont les caractéristiques sont déjà connues sont aussi inclus dans cette catégorie.

Lorsque ce genre de SIA est introduit dans une organisation, l'humain joue un rôle prépondérant en définissant ses règles de fonctionnement et les tâches qui lui sont confiées. Le système rend les décisions automatiquement en suivant les règles prédéfinies sans apprentissage ni adaptation. La relation se caractérise par le fait que l'opérateur humain s'appuie sur le système d'IA pour automatiser certaines tâches de façon prévisible.



Des algorithmes d'apprentissage automatique sont utilisés pour entraîner ces SIA à partir d'un ensemble de données et d'exemples. Les décisions ou *outputs* du système reposent avant tout sur l'ensemble d'apprentissages fournis (Chatila et al., 2021, pp. 26-27). Ces systèmes visent l'optimisation des actions et des tâches et ont une certaine capacité d'adaptation à l'intérieur d'un cadre restreint (Rao & Verweij, 2017). Ils n'ont pas la capacité d'apprendre directement au contact de l'environnement plus large dans lequel ils opèrent, mais ils peuvent modifier leur fonctionnement en s'entraînant sur les nouveaux ensembles de données fournis ou à la suite des rétroactions des concepteurs/utilisateurs.

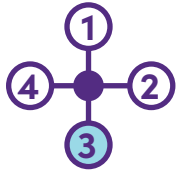
Les SIA dotés d'une autonomie limitée peuvent effectuer des tâches de manière indépendante dans un cadre étroit et bien défini. Ces systèmes peuvent analyser des données et fournir des recommandations, mais la décision quant aux choix à faire ou aux actions à poser dépendent des opérateurs/utilisateurs. Les moteurs de recommandation, les assistants virtuels aux fonctionnalités limitées, les systèmes de détection de fraude et les systèmes de reconnaissance d'images de base en sont des exemples.

À ce degré d'autonomie, les systèmes n'apprennent pas au contact de l'environnement, virtuel ou physique, concerné par leur tâche. Ils ne peuvent donc pas s'y adapter ou l'influencer directement. S'ils ont accès à l'environnement, c'est pour en extraire de l'information à laquelle ils appliquent un ensemble de règles qui sont fournies par les opérateurs/utilisateurs (corpus de lois, hypothèses scientifiques, règles du jeu, plan de marketing, etc.). L'autonomie dans l'accomplissement de leur tâche est limitée, entre autres, par la quantité et la qualité des données et exemples rendus disponibles par les opérateurs/utilisateurs (ou présents dans l'environnement limité auquel les opérateurs/utilisateurs leur donnent accès). Ces systèmes ayant la capacité d'apprendre et d'analyser, ils sont à même de générer, à partir du lot d'exemples et de données auquel ils ont accès, des corrélations difficilement identifiables par l'humain ou des hypothèses non explorées précédemment. Ils manifestent en cela un pouvoir discrétionnaire limité (Young et al., 2019, p. 304).

Lorsque ce genre de système est introduit dans une organisation, il est à même de fournir des recommandations, d'offrir un soutien aux décisions ou d'effectuer des analyses prédictives, tout cela dans un cadre étroit, tandis que l'utilisateur prend les décisions finales sur la base des résultats du système. La relation se caractérise par le fait que le SIA assiste l'utilisateur dans l'accomplissement de tâches spécifiques, tandis que ce dernier conserve le contrôle et la responsabilité du résultat final.

¹² **Executive autonomy:** the Agent is not allowed to decide anything but the execution of the entire delegated plan.

¹³ Voir aussi *AI-augmented government: Using cognitive technology redesign public sector work*, Eggers et al., 2017, Deloitte Center for Government Insights. En ligne : <https://www2.deloitte.com/us/en/insights/focus/cognitive-technologies/artificial-intelligence-government.html>



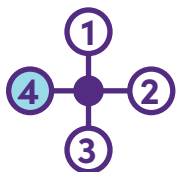
Degré d'autonomie 3 Semi-autonome

Les SIA semi-autonomes peuvent effectuer des tâches avec une intervention humaine réduite. Ils ont des capacités d'analyse et de prise de décision, et peuvent s'adapter à de nouvelles situations en tirant des leçons de leur expérience. Ces systèmes peuvent gérer des tâches complexes de manière autonome, mais ils dépendent toujours de la supervision et de l'apport humains pour les décisions critiques, la définition des objectifs et la correction des erreurs. Parmi les exemples, on peut citer les systèmes de diagnostic médical pilotés par l'IA, les systèmes avancés de traitement du langage naturel et les systèmes de réponse aux urgences qui fournissent des recommandations aux humains sur le meilleur plan d'action à mettre en place.

Pour ces SIA, le comportement du système évolue grâce aux interactions/rétroactions avec l'environnement. Dans les systèmes de ce type, les décisions sont effectivement prises suite au contact avec l'environnement physique ou virtuel concerné par la tâche (Chatila et al., 2021, p. 26-27). Le SIA peut renforcer/ajuster par lui-même le *poids* de différentes données ou informations afin de favoriser l'accomplissement du but ou de la tâche qui lui est assigné. Il peut donc se modifier et innover dans la façon d'accomplir sa tâche et manifeste en cela une autonomie de planification (Nickles et al., 2004, p. 45)¹⁴. Certains parlent d'*intelligence augmentée* : de systèmes avancés qui augmentent ou suppléent la prise de décision humaine et apprennent de leurs interactions avec les humains et l'environnement (Rao & Verweij, 2017).

Ces systèmes peuvent faire preuve d'initiative quant à l'accomplissement de leurs tâches en les accomplissant de façon originale, en ciblant des problèmes, des lacunes ou des redondances, et en suggérant des solutions et des améliorations. Ils manifestent en cela un pouvoir discrétionnaire d'intensité modérée à élevée, selon la nature de leur tâche et de l'environnement auquel ils ont accès (Young et al., 2019, p. 304).

Lorsque ce genre de système est introduit dans une organisation, le SIA et l'utilisateur collaborent plus étroitement. Cette collaboration se caractérise par une relation plus équilibrée, où l'opérateur assure la supervision, la définition des objectifs et la correction des erreurs, tandis que le système joue un rôle plus important dans l'accomplissement des tâches, la prise de décision et la suggestion d'amélioration.



Degré d'autonomie 4 Autonome

Les SIA autonomes peuvent planifier, établir des priorités et accomplir des tâches avec une supervision humaine minimale. Ils possèdent des capacités de décision avancées, peuvent fixer leurs propres objectifs et s'adapter à de nouvelles situations en tirant les leçons de leurs propres expériences. Ces systèmes peuvent analyser et optimiser leurs performances de manière indépendante, même dans des environnements complexes et incertains. Parmi les exemples d'applications potentielles, citons les systèmes autonomes de gestion du trafic ou les plateformes décisionnelles pilotées par l'IA pour les systèmes à grande échelle tels que les villes intelligentes (Cugurullo, 2020, p.11) ou les marchés financiers.

Le SIA peut prioriser ses actions (son *output*) à partir de sa propre perception de l'environnement dans lequel il évolue. À ce degré d'autonomie, le SIA possède une 'conception du monde' plus ou moins précise nourrie par ses apprentissages et ses interactions (ex : *map of distributed knowledge* dans un système multi-agents). Le système peut manifester un comportement 'personnalisé' et agir de façon autonome en relation avec les humains et les structures de l'organisation dans laquelle il s'insère.

¹⁴ **Planning autonomy:** the Agent is allowed to plan by itself, the delegated action is not completely specified in the delegation itself.

Dans cette dernière catégorie de SIA, les décisions font intervenir des algorithmes basés sur des principes/ motivations internes et des *croyances* sur la situation actuelle. Les décisions sont prises par le système, elles ne sont pas fixées à l'avance et ne sont pas directement déterminées par l'environnement ou l'entraînement du système (Chatila et al., 2021, p. 26-27). On parle ici d'*intelligence autonome* puisque le processus de prise de décision requiert un minimum de supervision humaine, et que ces systèmes peuvent s'adapter à différentes situations et agir sans assistance humaine (Rao & Verweij, 2017). Il s'agit de SIA apprenant à identifier et à classer les instances d'un phénomène sur la base des caractéristiques des données elles-mêmes (Harrison & Luna-Reyes, 2022, p. 497).

À ce degré d'autonomie, les systèmes priorisent leurs actions par eux-mêmes et peuvent 'recalibrer' leur but à partir de leur modèle du monde et de leur conception de la tâche à accomplir. Ils manifestent en cela une autonomie concernant le ou les objectif(s) à prioriser pour assurer l'accomplissement de leur tâche plus large (Nickles et al., 2004, p. 45)¹⁵. Le système peut adapter sa compréhension de son environnement au contact de ce dernier et adapter son comportement en conséquence.

Aucun système d'IA ne possède actuellement une autonomie complète dans un sens pratique. Le déploiement d'un tel système impliquerait qu'il planifie, hiérarchise et exécute des tâches de manière indépendante avec une intervention humaine minimale ou inexistante. La relation entre humain et système se caractériserait par le fait que le SIA assumerait le rôle principal dans l'accomplissement des tâches, tandis que l'opérateur humain serait chargé de surveiller ses performances, de veiller à l'alignement sur les valeurs et les objectifs humains et d'intervenir en cas de situations exceptionnelles.

Bien qu'aucun système complètement autonome n'existe présentement, certains sont envisagés dans le secteur public. Nous présenterons donc des projets semi-spéculatifs en guise d'exemples.

15 *Goal autonomy: the Agent is allowed to have/find its own goals* ».

4. Cas d'application de l'utilisation de l'IA dans le secteur public

4. Cas d'application de l'utilisation de l'IA dans le secteur public

La Bonne Boîte

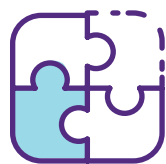
Tableau 2. Description synthétique de *La Bonne Boîte*

Chronologie	2015 : projet pilote 2016 : lancement national
Pays	France
Organisation	Pôle emploi (agence publique française pour l'emploi) - Maintenant France travail (2024)
Développeur	Public (Pôle emploi en collaboration avec l'European Employment Services Network/PES)
Bénéficiaire direct	Usagers (personnes à la recherche d'emploi)
Degré d'autonomie	1. Automatique



Objectif poursuivi

La Bonne Boîte est un outil en ligne¹⁶ qui identifie des entreprises ayant une forte probabilité d'embaucher dans les six mois à venir dans le secteur et la zone géographique sélectionnés par un demandeur d'emploi¹⁷. L'objectif principal de ce système est de faciliter le processus des personnes en recherche d'emploi en leur permettant d'avoir accès à des offres potentielles qui ne sont pas publiées par les employeurs (les « offres d'emploi cachées »). Le moteur de recherche améliore ainsi les chances d'un demandeur d'emploi d'obtenir un entretien d'embauche sur la base de candidatures spontanées en lui permettant de concentrer ses efforts sur les entreprises qui ont le plus de chances d'embaucher dans un avenir rapproché. Ceci réduit le temps passé à chercher un emploi et améliore l'efficacité globale du marché de l'emploi en France.



Étapes de développement

La mise en œuvre de *La Bonne Boîte* remonte au milieu des années 2010 alors que l'on constatait que près de sept entreprises sur dix examinent d'abord les candidatures spontanées qu'elles ont reçues lorsqu'elles cherchent à recruter de nouveaux employés. Afin de faciliter l'accès à ce marché caché de l'emploi, un outil est développé par un conseiller de Pôle emploi, l'agence publique française pour l'emploi. C'est ainsi que *La Bonne Boîte* voit le jour. Elle sera lancée, en 2015, sous forme d'un projet pilote visant à tester l'efficacité de la plateforme et recueillir les commentaires des utilisateurs¹⁸.

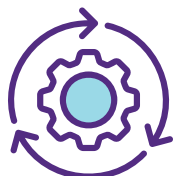
16 Interface web *La Bonne Boîte* : <https://labonneboite.pole-emploi.fr/>

17 European Commission-*La Bonne Boîte* : <https://ec.europa.eu/social/BlobServlet?docId=20605&langId=en>; *La Bonne Boîte, Mieux cibler les entreprises pour vos candidatures spontanées* : <https://www.youtube.com/watch?v=AAUNWmVm2Y&t=66s>

18 *La Bonne Boîte, un outil pour cibler ses candidatures spontanées* : https://www.bfmtv.com/economie/emploi/la-bonne-boite-un-outil-pour-cibler-ses-candidatures-spontanees_AN-201808180050.html

4. Cas d'application de l'utilisation de l'IA dans le secteur public

Cette initiative a été appuyée par La Grande Fabrique de Pôle emploi et l'outil a été cofinancé par le *Fonds Social Européen* dans le cadre du *Programme opérationnel national « emploi et inclusion »*¹⁹. Après le succès du projet pilote, *La Bonne Boîte* a été officiellement lancée à l'échelle nationale en 2016. Depuis, la plateforme est continuellement mise à jour afin d'améliorer son efficacité globale et on y dénombre plus de 5 millions de consultations par mois.



Fonctionnement du système

Le service proposé par l'outil à partir de son site internet permet à l'utilisateur de rechercher des entreprises ou organismes susceptibles de recruter des candidats de profils similaires au sien, en fonction des données qu'il a saisies (métier, secteur géographique)²⁰.

Les entreprises ou organismes qui apparaissent dans les résultats de recherche peuvent être classés :

- en fonction de leur proximité géographique par rapport au périmètre défini par l'utilisateur ;
- en fonction du potentiel d'embauche des entreprises, calculé mensuellement par Pôle emploi et mis à jour quotidiennement en fonction des demandes formulées explicitement par les entreprises auprès des services *La Bonne Boîte*.

Selon Pôle emploi, le système attribue un potentiel d'embauche à chaque entreprise. Ce potentiel est déterminé par l'étude approfondie de millions de données. Il devient alors possible de prédire les entreprises qui devraient embaucher, dans les semaines à venir, des candidats au profil correspondant à l'utilisateur. Seules les entreprises ayant un fort potentiel d'embauche apparaissent dans les listes fournies par *La Bonne Boîte*²¹.

Pour parvenir à cette prédiction, *La Bonne Boîte* s'appuie sur une base de données mise à jour chaque mois et constituée à partir des déclarations préalables à l'embauche (DPAE), une formalité que les entreprises doivent accomplir auprès de l'URSSAF (Unions de Recouvrement des Cotisations de Sécurité Sociale et d'Allocations Familiales) lorsqu'elles recrutent²². Les données sont ensuite traitées à l'aide d'un algorithme *open source* conçu à l'interne²³ qui permet de trier les entreprises en fonction des postes et de leur localisation, et de noter leur potentiel d'embauche dans les 6 mois à venir.

Les demandeurs d'emploi peuvent accéder en ligne à la plateforme *La Bonne Boîte* et saisir la profession qu'ils souhaitent exercer ainsi que la zone géographique ciblée. La plateforme génère ensuite une liste personnalisée d'entreprises susceptibles de recruter dans la région et la profession sélectionnées, classées par ordre de probabilité d'embauche.

Le système est continuellement mis à jour avec de nouvelles données, et ses algorithmes sont affinés au fil du temps pour améliorer la précision de ses recommandations. Les commentaires des utilisateurs sont également recueillis pour apporter des ajustements et des améliorations à la plateforme.

19 Interlocuteur des instances internationales : <https://www.reseau-emploi-med.com/en/cooperation-2/>

20 Conditions Générales d'Utilisation : <https://labonneboite.pole-emploi.fr/conditions-generales>

21 Emploi Store/La Bonne Boîte : <https://www.emploi-store.fr/portail/services/laBonneBoite>

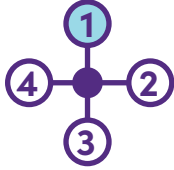
22 *La Bonne Boîte, un outil pour cibler ses candidatures spontanées* : https://www.bfmtv.com/economie/emploi/la-bonne-boite-un-outil-pour-cibler-ses-candidatures-spontanees_AN-201808180050.html

23 Startups d'État Pôle emploi/La Bonne Boîte : <https://github.com/StartupsPoleEmploi/labonneboite>



Bénéficiaire direct du système

Le travail effectué par cet outil bénéficie avant tout à la personne en recherche d'emploi qui se voit recommander les entreprises ou organisations qui ont de bonnes probabilités d'embaucher dans les mois à venir. Les employés responsables d'assister les citoyens dans leur recherche d'emploi bénéficient indirectement du système, ainsi que l'organisation Pôle emploi qui voit une partie de sa mission favorisée par sa mise en place.



Degré d'autonomie du système

La Bonne Boîte est essentiellement un outil de recommandation qui met en contact les personnes en recherche d'emploi avec des entreprises ou organisations avant même que ces dernières n'affichent les offres publiquement. Le système croise les caractéristiques saisies par le demandeur d'emploi avec les besoins potentiels des entreprises et organisations établies à partir d'un traitement statistique des données colligées. Il s'agit d'un outil dépourvu de capacités d'apprentissage ou de décision, qui permet d'automatiser l'adéquation entre l'offre et les besoins de main-d'œuvre. Ce système correspond au degré d'autonomie 1 (automatique).



Pistes de réflexions éthiques

Bénéfices et avantages du SIA

- Permet d'optimiser les pratiques de recherche d'emploi, en offrant un accès à des informations sur des employeurs potentiels et des emplois éventuellement disponibles qui ne seraient pas accessibles autrement.
- Utilisation optionnelle. Les demandeurs d'emploi ne sont pas contraints de modifier leurs préférences ou de remplacer leurs méthodes habituelles de recherche d'emplois.
- Permet aux utilisateurs de postuler directement en ligne aux emplois qui les intéressent.
- Fait preuve d'une certaine transparence en requérant l'acceptation des conditions d'utilisation et en se référant au *Règlement général sur la protection des données* (RGPD) de l'Union européenne.

Inconvénients et risques du SIA

- Manque de clarté concernant l'utilité et l'utilisation des cookies que l'utilisateur doit accepter lorsqu'il accède au site internet de l'application.
- Inquiétude sur le stockage et le partage des données, en particulier concernant la protection de la vie privée des utilisateurs. Les données sont-elles préservées, partagées ou revendues à des tiers ? Pendant combien de temps sont-elles conservées ? Est-ce qu'une corrélation des données permettrait d'identifier les habitudes de certains individus, de façon à leur envoyer de la publicité ciblée ? Qu'en est-il de la protection de la vie privée pour les chômeurs ou les bénéficiaires de l'aide sociale ? Est-ce que leur comportement sur l'application est « tracé » par des agents publics et pourrait avoir des répercussions sur leurs prestations sociales ? Est-ce que les compagnies d'assurance pourraient avoir accès à ces informations ?
- Manque de précision concernant le traitement et la protection des données des entreprises entraînant un risque potentiel au niveau de la compétitivité (une entreprise pourrait-elle être désavantagée lorsque des informations à son sujet sont rendues disponibles par l'application, par rapport à une autre entreprise pour laquelle ce ne serait pas le cas ?). Il n'est pas certain que les entreprises aient l'occasion de consentir ou de s'opposer à l'utilisation de leurs données qu'elles ont transmises à l'État à une autre fin.

4. Cas d'application de l'utilisation de l'IA dans le secteur public

- Bien que la fonctionnalité « Je postule » soit optionnelle, elle nécessite la création d'un profil et la transmission d'informations personnelles pour postuler à un emploi en ligne, directement à partir de l'application. Ceci soulève de nouvelles questions concernant la protection des renseignements personnels.
- La fonction de recommandation des emplois à des utilisateurs ayant un profil similaire implique un profilage et pourrait potentiellement mener à une discrimination puisque des offres d'emplois sont transmises à certain(e)s personnes/groupes, mais pas à d'autres. Cette fonctionnalité soulève la préoccupation plus large du regroupement par catégories d'individus (selon le sexe, le lieu de résidence, l'origine ethnique, etc.) qui est associé au profilage dans différents SIA.
- Des inquiétudes existent quant à la possibilité de discrimination basée sur des critères simples tels que l'adresse et le code postal de l'utilisateur.

Valeurs à prendre en considération pour le déploiement et l'utilisation de ce SIA (ou d'un système similaire)

- Autonomie. La recherche d'emploi est un domaine sensible (l'impact peut se faire sentir au niveau monétaire, familial, etc.). Un filtrage excessif des offres d'emploi basé sur les informations personnelles risque d'entraver l'autonomie des utilisateurs dans leur recherche d'emploi en les empêchant d'avoir une vue exhaustive sur les offres disponibles.
- Équité et non-discrimination. Le système offre la possibilité de recommandation des emplois à des utilisateurs ayant un profil similaire. Ceci implique qu'il y a comparaison ou corrélation entre les données fournies par différents utilisateurs. Ce profilage peut mener à une forme de discrimination : le système doit présenter les résultats de façon équitable sans discriminer selon le sexe, le lieu de résidence, l'origine ethnique, etc.
- Optionnalité et accessibilité. La Bonne Boîte est un outil optionnel, et non un prérequis à la recherche d'emploi. Ceci étant dit, l'outil doit d'être accessible à tous et conçu pour être utilisable par un large éventail de personnes.
- Protection de la vie privée. Les utilisateurs doivent être assurés que leurs données personnelles seront protégées et non partagées de manière inappropriée.
- Protection des données d'entreprises. La compétitivité de certaines entreprises pourrait être mise à mal par l'information transmise par l'application.
- Transparence. L'utilité des cookies et le traitement des données des utilisateurs devraient être plus transparents.

Comment protéger ou restaurer les valeurs affectées par le SIA?

- Autonomie. Les utilisateurs devraient être en mesure de voir un grand nombre d'options disponibles et de choisir parmi elles, sans que l'application ne filtre excessivement les options à la place des utilisateurs.
- Équité et non-discrimination. Il faut veiller à ce que le système ne favorise pas certains utilisateurs au détriment d'autres, notamment en évitant les biais liés à l'adresse, au code postal ou à d'autres caractéristiques.
- Optionnalité et accessibilité. Il est important de maintenir l'accessibilité de l'application pour tous les utilisateurs et de garantir que l'utilisation de l'application reste facultative.
- Protection de la vie privée. Des mesures techniques et organisationnelles peuvent être mises en place pour préserver la sécurité des données personnelles.
- Protection des données d'entreprise. Obtenir le consentement des entreprises dont les données sont utilisées.
- Transparence. Une politique de confidentialité claire et facilement accessible aux utilisateurs est nécessaire afin d'expliquer la manière dont les cookies et les données sont utilisés par l'application, notamment en ce qui concerne le partage et le stockage des données.

Tableau 3. Description synthétique de *GoSafe*

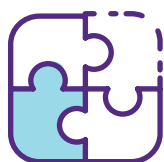
Chronologie	2010
Pays	Irlande
Organisation	<i>An Garda Síochána</i> (AGS) communément appelé <i>Gardaí</i> (National Police and Security Service of Ireland)
Développeur	Partenariat public-privé (National Police and Security Service of Ireland et <i>GoSafe</i>)
Bénéficiaire direct	Employés
Degré d'autonomie	1. Automatique



Objectif poursuivi

Le système *GoSafe* est une initiative de sécurité routière qui vise principalement à réduire le nombre d'accidents, de blessures et de décès sur les routes irlandaises. Le système utilise un réseau de radars, mobiles et fixes, pour faire respecter les limites de vitesse et décourager les excès de vitesse.

Plus spécifiquement, *GoSafe* contribue au respect des limitations de vitesse et à la conduite responsable en infligeant des amendes à ceux qui dépassent les limites. Le système fait aussi office de sensibilisation des conducteurs à l'importance de la sécurité routière ; la présence du système *GoSafe* rappelle aux conducteurs l'importance de respecter les limitations de vitesse et d'adopter des habitudes de conduite sécuritaires afin de se protéger et de protéger les autres usagers de la route. Finalement, le système aide les autorités à identifier les zones où les excès de vitesse et les accidents sont les plus fréquents, ce qui permet de cibler les interventions et d'améliorer les mesures de sécurité routière.



Étapes de développement

La mise en œuvre du système *GoSafe* en Irlande remonte à la fin des années 2000. Elle s'inscrit dans le cadre d'un effort plus large du gouvernement irlandais et des autorités chargées de la sécurité routière visant à réduire le nombre d'accidents de la route, de blessures et de décès²⁴ :

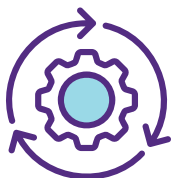
- en 2007, le gouvernement irlandais a publié la *Stratégie de sécurité routière 2007-2012*, qui présentait diverses initiatives visant à améliorer la sécurité routière dans le pays²⁵. L'une de ces initiatives était l'introduction d'un nouveau système de contrôle par caméra pour dissuader les excès de vitesse et faire respecter les limitations de vitesse ;

²⁴ A Citizen's Guide to Artificial Intelligence. MIT Press. <https://doi.org/10.7551/mitpress/12518.001.0001> (Zerilli, 2021, pp. 133-134).

²⁵ Ireland's new Road Safety Authority – and strategy 2007-2012 | CIECA, s. d.: <https://www.cieca.eu/news/125>

4. Cas d'application de l'utilisation de l'IA dans le secteur public

- en 2009, le gouvernement irlandais a attribué le contrat de gestion et d'exploitation du nouveau système de contrôle par radars à *GoSafe*, un consortium composé de *Redflex* (Australie), *Spectra* (Irlande) et *Egis Projects* (France)²⁶ ;
- en novembre 2010, le système *GoSafe* est devenu officiellement opérationnel, avec le déploiement de camionnettes radars mobiles dans tout le pays. Le déploiement initial comprenait environ 600 zones de radars, sélectionnées en fonction des données d'accidents et du risque d'incidents liés à la vitesse ;
- au fil des ans, le nombre de zones de radars a augmenté et le système a été continuellement mis à jour afin d'améliorer son efficacité. Le réseau de caméras de sécurité s'est développé pour couvrir à la fois des emplacements mobiles et fixes, et le système est devenu un élément important de la stratégie de sécurité routière en Irlande ;
- en 2013, le gouvernement a publié une nouvelle *Stratégie de sécurité routière 2013-2020*, qui continue de souligner l'importance du système *GoSafe* et du contrôle de la vitesse pour atteindre les objectifs de sécurité routière du pays²⁷ ;
- suite à l'évaluation positive des impacts du système sur la sécurité routière²⁸, le nombre de caméras est fortement augmenté en février 2020²⁹ ;
- en 2022, une nouvelle expansion du système était annoncée, ajoutant 61 nouvelles zones couvertes par des caméras aux 1312 déjà utilisées sur le territoire irlandais³⁰.



Fonctionnement du système

Le système *GoSafe* utilise une combinaison de radars mobiles et fixes pour contrôler la vitesse des véhicules et faire respecter les limitations de vitesse. Il s'appuie sur une technologie avancée pour mesurer avec précision la vitesse des véhicules et traiter les données collectées. Selon *Redflex*, l'un des partenaires qui opèrent *GoSafe* en sous-traitance pour les forces de police irlandaises, le système comprend des caméras mobiles, un système de gestion des caméras numériques (pour gérer à distance les caméras embarquées dans les véhicules) et un système de traitement des images et des infractions (IIPS) pour statuer sur les infractions³¹.

Le système *GoSafe* utilise à la fois des radars mobiles et des radars fixes. Les radars mobiles sont généralement montés dans des camionnettes garées dans des zones désignées, tandis que les radars fixes sont installés de manière permanente à des endroits spécifiques. Les radars utilisent la technologie radar ou LIDAR pour mesurer la vitesse des véhicules en mouvement³².

Lorsqu'un véhicule circule au-dessus de la limite de vitesse permise, le radar enregistre la vitesse du véhicule, l'heure et la date de l'infraction. Il photographie le véhicule et la plaque d'immatriculation. Ces données sont ensuite envoyées à un centre de traitement pour analyse et vérification.

26 *Redflex Holdings Limited Annual Report 2012*, 2012, p. 11: https://redflex.com/annual-reports/?attachment_id=3069&download=1; *Redflex Traffic Systems - Verdict Traffic*: https://www.roadtraffic-technology.com/contractors/traffic_man/redflex-traffic-systems/#company-details; *Egis in Ireland*, s. d.: <https://www.egis-group.com/locations/europe-central-asia/ireland>

27 *gov.ie - Road Safety Strategy 2013-2020*, s. d.: <https://www.gov.ie/en/publication/8ed7fd-road-safety-strategy-2013-2020/>

28 *Study shows Irish speed cameras provide five-fold benefit* | ITS International, s. d.: <https://www.itsinternational.com/its2/feature/study-shows-irish-speed-cameras-provide-five-fold-benefit>

29 *Safety Camera Expansion 10th February 2020*, s. d.: <https://www.rsa.ie/news-events/news/details/2020/02/10/safety-camera-expansion-10th-february-2020>

30 *Safety Cameras - Garda*, s. d.: <https://www.garda.ie/en/roads-policing/safety-cameras/>

31 *Redflex Holdings Limited Annual Report 2012*, 2012, p. 11: https://redflex.com/annual-reports/?attachment_id=3069&download=1

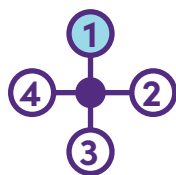
32 *Speed cameras in Ireland: How they work and types* | Carzone Advice, 2023: <https://www.carzone.ie/motoring-advice/speed-cameras-in-ireland-how-they-work-and-types/2925>; *Our Cameras* | *GoSafe Welsh Road Casualty Reduction Partnership*, s. d.: <https://www.gosafe.org/learn/how-we-enforce/our-cameras/>

Au centre de traitement, les données enregistrées sont examinées et vérifiées pour s'assurer de l'exactitude de la mesure de la vitesse et confirmer que le véhicule était bien en infraction. Une fois l'infraction confirmée, le propriétaire enregistré du véhicule reçoit un avis d'infraction pour excès de vitesse, qui comprend une amende et l'indication sur le nombre de points de pénalité ajoutés à son dossier de conduite. L'avis est envoyé par la poste à l'adresse associée à l'immatriculation du véhicule.



Bénéficiaire direct du système

Ce système effectue avant tout un travail de mise en application des règles dans le but de favoriser la sécurité sur les voies publiques. Bien que les usagers de la route bénéficient de la sécurité accrue favorisée par le système, il n'est pas ici question d'améliorer la durée de leur déplacement, leur confort ou la fluidité de la circulation. Les bénéficiaires directs sont les employés qui effectuent les tâches liées à l'application des règles du Code de la route. L'organisation (la police nationale) bénéficie indirectement du travail du système qui contribue à l'atteinte d'objectifs concernant la sécurité routière sur le territoire irlandais.



Degré d'autonomie du système

Ce système automatisé utilise la réflexion radar pour calculer la vitesse des véhicules venant en sens inverse, déterminer s'ils dépassent la limite de vitesse, puis prendre une photo. Il ne repose pas sur des techniques de programmation particulièrement complexes et ne constitue pas une technique d'IA avancée³³. Il met en œuvre une série d'actions programmées à l'avance et ne manifeste pas de capacités d'apprentissage ou de prise de décision. Ce type de système correspond au degré d'autonomie 1 (automatique).



Pistes de réflexions éthiques

Bénéfices et avantages du SIA

- Assiste les policiers dans l'application de la réglementation routière. L'emploi de ce système permet de détecter automatiquement les véhicules en infraction, de prendre des photos, et de transmettre l'information pour qu'elle soit traitée par un agent. Cette automatisation peut augmenter l'efficacité des services policiers en permettant aux agents de se concentrer sur d'autres tâches.
- Contribue au respect des règles de sécurité routière. La simple présence de ces systèmes de surveillance peut encourager les individus à réguler leur comportement et à respecter les règles, ce qui contribue à améliorer la sécurité routière.
- Favorise la sécurité de tous les usagers de la route. Ce système contribue à la sécurité routière en dissuadant les conducteurs de dépasser les limites de vitesse, en particulier dans les zones à risque comme les zones scolaires ou les zones où de nombreux accidents se sont produits dans le passé.
- Harmonisation de traitement. Contrairement à la surveillance humaine qui peut être sporadique ou discriminatoire, *GoSafe* peut potentiellement capter toutes les infractions, ce qui permet une application plus uniforme des lois et réglementations.
- Proportionnalité. Contrairement à d'autres systèmes de surveillance, *GoSafe* ne collecte que des informations limitées et spécifiques. L'équilibre entre la collecte d'informations personnelles et la finalité du système apparaît comme raisonnable et proportionné.

33 A Citizen's Guide to Artificial Intelligence. MIT Press. <https://doi.org/10.7551/mitpress/12518.001.0001> (Zerilli, 2021, pp. 133-134)

Inconvénients et risques du SIA

- Information et contestation. Puisqu'il n'y a pas d'interaction au moment où l'infraction est constatée, il peut être difficile d'obtenir de l'information sur l'infraction, ou de contester les infractions détectées par ce système.
- Risques de discrimination. Il y a un risque de discrimination basée sur l'endroit où les systèmes sont installés (si certaines populations sont ciblées) ou sur l'éducation, la langue ou la culture (si certains individus sont moins informés ou conscients des nouvelles méthodes de surveillance).
- Risques pour la démocratie. Dans les sociétés autoritaires ou celles où la démocratie est fragile, ce type de système pourrait devenir un outil de contrôle supplémentaire pour le pouvoir en place.
- Sécurité des données. Les données collectées pourraient être utilisées de manière malveillante ou non désirée.
- Surveillance accrue des citoyens. L'utilisation de ce type de système peut accentuer l'omniprésence de la surveillance, suscitant des questions sur le type de société dans lequel nous vivons.
- Utilisation secondaire des données. Une fois que les données sont utilisées pour la fonction principale du système (la régulation de la vitesse par radar), elles pourraient être utilisées pour d'autres fonctions, comme la localisation et l'arrestation de suspects, la création de profils d'individus ou la modulation des coûts d'assurance.

Valeurs à prendre en considération pour le déploiement et l'utilisation de ce SIA (ou d'un système similaire)

- Encadrement réglementaire. Une réglementation claire doit encadrer l'utilisation de ces systèmes, notamment en termes de sécurisation, d'utilisation et d'accès aux données.
- Équité. Le système ne doit pas cibler ou discriminer certaines populations. Les emplacements des systèmes de surveillance doivent être choisis de manière équitable et non discriminatoire.
- Justice. Les citoyens doivent avoir la possibilité de contester les infractions détectées par le système. Un processus de recours doit être mis en place pour remédier aux erreurs potentielles du système.
- Pertinence et proportionnalité. L'usage de ces systèmes de surveillance doit être justifié et proportionné. La surveillance omniprésente peut avoir un impact négatif sur la société et doit donc être utilisée avec parcimonie.
- Respect de la démocratie. L'utilisation de ces systèmes doit se faire dans le respect des valeurs démocratiques. Des garanties doivent être en place pour éviter que ces systèmes deviennent un outil d'oppression par le pouvoir en place.
- Respect de la vie privée. La collecte de données doit être limitée à ce qui est strictement nécessaire pour accomplir la finalité du système. Le partage de ces données avec des tiers (par exemple, des compagnies d'assurance) doit être réglementé.
- Sécurité des données. Les données collectées doivent être sécurisées pour éviter leur utilisation malveillante ou indésirée.
- Supervision humaine. Bien que le système puisse fonctionner de manière automatique, une supervision humaine est nécessaire pour valider et attester les infractions détectées par le système.
- Transparence. Les citoyens doivent être informés de la présence et du fonctionnement de ces systèmes. Ceci est particulièrement important pour éviter de désavantager les individus moins informés ou moins conscients de ces nouvelles méthodes de surveillance.

Comment protéger ou restaurer les valeurs affectées par le SIA?

- Bénéfices collectifs versus droits individuels. Il convient de trouver un équilibre entre les risques associés à ces systèmes en termes de droits individuels (comme la surveillance omniprésente et l'atteinte à la vie privée) et les bénéfices collectifs en termes de respect des lois et règlements. Cela implique un débat ouvert et une évaluation approfondie des bénéfices et des inconvénients.
- Cadre réglementaire. Pour protéger la vie privée, assurer la sécurité des données et éviter une utilisation abusive, un cadre réglementaire clair sera nécessaire. Ce cadre devrait inclure des règles sur le type de données collectées, la manière dont elles sont stockées, les personnes qui y ont accès et les raisons légitimes de consultation. Une attention particulière devra être portée aux usages secondaires indésirables des données collectées, comme leur utilisation par les compagnies d'assurance pour moduler les coûts de couverture, par exemple.
- Contrôle et supervision. Il est proposé qu'une entité supervise l'installation et le fonctionnement du système. Cela peut aider à garantir que les procédures appropriées soient suivies et que les valeurs clés soient respectées.
- Équité. Des mesures doivent être mises en place pour éviter toute forme de discrimination, que ce soit en fonction de l'endroit où les systèmes sont installés ou en lien avec l'éducation, la langue ou la culture des individus surveillés.
- Évaluation régulière. Une évaluation régulière de l'efficacité et de l'impact de ces systèmes peut aider à identifier et à corriger les problèmes potentiels, tout en assurant que ces systèmes produisent les bénéfices escomptés.
- Possibilité de contestation. Une procédure simple et rapide devrait permettre aux conducteurs d'obtenir des informations sur la nature de l'infraction et de contester les infractions détectées par le système.
- Sécurité des données. Une attention particulière doit être portée à la sécurisation des données afin d'éviter leur utilisation malveillante ou indésirée. Cela comprend la protection contre les cyberattaques et la sécurisation des communications entre les caméras et les serveurs où les données sont stockées.
- Transparence. Il est suggéré d'indiquer clairement au public la présence de ces systèmes de surveillance. Cela pourrait aider à renforcer la confiance du public et à garantir que les individus soient conscients de la surveillance à laquelle ils sont soumis.

Chorus

Tableau 4. Description synthétique de *Chorus*

Chronologie	2007 : lancement initial 2011 : intégration et utilisation par toutes les administrations des ministères français, au niveau national et infranational
Pays	France
Organisation	Agence pour l'informatique financière de l'État (AIFE)
Développeur	Partenariat public-privé : l'Agence pour l'informatique financière de l'État (AIFE) et SAP SE (German software company) pour le logiciel et Sopra Steria et Capgemini France pour la maintenance.
Bénéficiaire direct	Organisation (l'appareil gouvernemental français)
Degré d'autonomie	1. Automatique

**Objectif poursuivi**

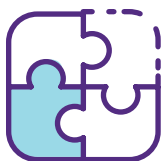
Chorus est le système d'information financière (SIF) des administrations centrales de l'État en France. Il a été développé pour faciliter la gestion et le suivi des dépenses publiques. Son objectif principal est de rationaliser et de moderniser les processus financiers du gouvernement français, en améliorant l'efficacité, la transparence et le contrôle.

Chorus couvre différents domaines fonctionnels : l'allocation des ressources de l'État, le traitement des dépenses de l'État, la gestion des actifs de l'État, la collecte des recettes non fiscales et l'intégration de trois types de comptabilité (comptabilité générale, comptabilité budgétaire et comptabilité analytique). Elle permet aux décideurs politiques de connaître exactement et en temps réel la situation financière de l'État central, d'un ministère ou d'un programme politique donné³⁴.

Chorus vise ainsi à consolider les multiples systèmes financiers utilisés par les différents ministères en une seule plateforme. Cela permet de garantir la cohérence des processus, des structures de données et des rapports entre toutes les entités, et aide le gouvernement à mieux contrôler et gérer son budget³⁵. Ce faisant, ce système de systèmes contribue à réduire les charges administratives et les coûts liés à la gestion des finances publiques, ce qui se traduit par une utilisation plus efficiente des ressources de l'État.

34 *Chorus: An Integrated Financial Information System for Policy Design in France*: <https://tropico-project.eu/cases/administration-costs-for-bureaucracy/chorus-an-integrated-financial-information-system-for-policy-design-in-france/>; Tropico/Comparative case Studies, , p. 15 : <https://tropico-project.eu/download/d4-3-research-report-on-comparative-case-studies/?wpdmdl=1052&refresh=603771627fc6e1614246242>

35 Ibid, pp. 25-26



Étapes de développement

La mise en place du système *Chorus* fait suite à l'adoption de la *Loi organique relative aux lois de finances* (LOLF) de 2001 qui visait à réformer le cadre budgétaire français et à jeter les bases d'une gestion des finances publiques davantage axée sur la performance et la transparence. Cette réforme a nécessité la mise en place d'un système d'information financière adapté pour répondre aux nouvelles exigences, ainsi que la création, en 2005, d'une organisation dédiée à sa conception, son développement, sa mise en œuvre et sa maintenance (l'Agence pour l'informatique financière de l'État/AIFE)³⁶.

Le déploiement national du système par l'AIFE se fera en plusieurs étapes. Le projet pilote, visant à tester la plateforme et ses fonctionnalités, a débuté en 2007 avec la participation d'un petit nombre d'agences gouvernementales (Défense, Enseignement supérieur et Budget).

À la suite du succès du projet pilote, le gouvernement français a déployé *Chorus* à l'échelle nationale³⁷ :

Between 2008 and 2011, the implementation of Chorus was organized in six waves which targeted different ministries and their central and subnational services. Through this process, AIFE intended to consider the different concerns of each ministry while developing the new financial information systems (FIS). From a technical point of view, the stepwise process also aimed at warranting that the budget and accounting information of the state were properly handled by the new FIS. By 1 January 2011, all ministries used Chorus and had implemented the shared FIS³⁸.

Depuis son déploiement complet, le système *Chorus* a été continuellement mis à jour et amélioré pour tenir compte des commentaires des utilisateurs, améliorer les fonctionnalités et intégrer de nouvelles avancées technologiques :

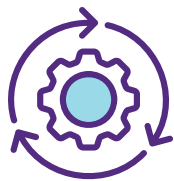
En 2021, l'AIFE a mis à la disposition des ministères des produits innovants (un captcha souverain, un traitement automatisé de certaines dépenses à caractère récurrent). Elle a développé des RPA (robots [pour l'automatisation des processus]), expérimentés par des ministères, pour traiter des tâches à moindre valeur ajoutée. Elle a créé un groupe thématique pour porter les projets innovants³⁹.

36 A transformative change through a coordination process and a steering agency. The case of the financial information system of the French central state (Defacqz & Dupuy, 2021, p. 7): <https://journals-sagepub-com.acces.bibl.ulaval.ca/doi/epub/10.1177/00208523211058859>

37 Tropico/*Comparative case Studies*, p. 22 : <https://tropico-project.eu/download/d4-3-research-report-on-comparative-case-studies/?wpdmml=1052&refresh=603771627fc6e1614246242>

38 A transformative change through a coordination process and a steering agency. The case of the financial information system of the French central state (Defacqz & Dupuy, 2021, pp. 9-11): <https://journals-sagepub-com.acces.bibl.ulaval.ca/doi/epub/10.1177/00208523211058859>

39 2021, une année de continuité et de transition pour l'AIFE : <https://aife.economie.gouv.fr/>



Fonctionnement du système

Le système *Chorus* est essentiellement un système d'information financière (SIF) qui centralise et normalise la gestion des finances publiques en France⁴⁰.

Centralisation des données. *Chorus* consolide les données financières des différents ministères et agences gouvernementales en une seule plateforme. Cela permet d'avoir des structures de données cohérentes et facilite le partage d'informations entre les différentes entités.

Automatisation des processus. Le système automatise de nombreux processus financiers, tels que la planification, l'exécution et le suivi du budget, ainsi que les achats et la comptabilité. Cela permet de réduire les tâches manuelles, de minimiser les erreurs et d'améliorer l'efficacité globale de la gestion financière.

Rapports et analyses. *Chorus* fournit une large gamme d'outils de reddition de comptes et d'analyse pour aider les agents publics à contrôler les dépenses publiques, à identifier les tendances et à prendre des décisions plus éclairées.

Le système fonctionne sur la plateforme SAP ERP⁴¹ de SAP SE, entreprise basée en Allemagne, et sa maintenance est assurée par Steria et Capgemini France⁴². Un passage vers la solution infonuagique SAP S/4HANA⁴³ était annoncé en 2021.



Bénéficiaire direct du système

Les principaux bénéficiaires du système *Chorus* sont les ministères et les organismes publics. Le système rationalise et centralise les processus financiers pour ces entités, en facilitant la gestion de leurs budgets, de leurs achats, de leur comptabilité et de leurs rapports. Le système contribue à réduire les charges administratives et à accroître l'efficacité, ce qui permet aux ministères et aux agences gouvernementales de mieux allouer les ressources. En fournissant une vue d'ensemble des dépenses publiques, *Chorus* permet aux gestionnaires publics de prendre des décisions plus éclairées en matière d'allocation budgétaire, de gestion financière et de formulation de politiques. Le bénéficiaire direct du système est donc l'organisation publique de l'État français dans son ensemble.

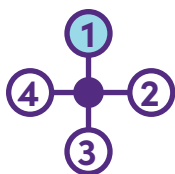
Chorus bénéficie indirectement aux gestionnaires, contrôleurs financiers et autres employés en simplifiant et automatisant de nombreuses tâches de gestion financière, leur permettant de se concentrer sur d'autres activités. Favorisant la transparence et la responsabilité dans les dépenses publiques, le déploiement de *Chorus* bénéficie également indirectement aux contribuables en favorisant une gestion transparente et responsable des fonds publics.

40 Rapport d'activité 2021 de l'AIFE : <https://aife.economie.gouv.fr/wp-content/uploads/2022/05/Rapport-annuel-2021-de-lAIFE.pdf> , pp. 6-9 Tropico/ Comparative case Studies, pp. 25-26 : <https://tropico-project.eu/download/d4-3-research-report-on-comparative-case-studies/?wpdmdl=1052&refresh=603771627fc6e1614246242>

41 SAP ERP is a software developed by the German company SAP SE (Systems, Applications & Products in Data Processing). This German-based corporation is a European multinational enterprise (Societas Europaea), which provides software to manage business operations and customer relations. Voir: SAP/What is SAP ERP?: <https://www.sap.com/products/erp/what-is-sap-erp.html>

42 Steria et Capgemini décrochent la maintenance de Chorus : <https://www.lemondeinformatique.fr/actualites/lire-steria-et-capgemini-decrochent-la-maintenance-de-chorus-32619.html>; Sopra Steria/AIFE – (Agency for State Financial Information Technology) Chorus: <https://www.soprasteria.com/industries/government/aife-chorus>

43 SAP S/4HANA Cloud : <https://www.sap.com/products/erp/s4hana.html>



Degré d'autonomie du système

Chorus est conçu pour centraliser, rationaliser et automatiser divers processus financiers. Le système aide les agents publics à gérer les finances publiques en automatisant des tâches routinières, en fournissant des données financières consolidées et en proposant des outils de reddition de comptes et d'analyse. Le système ne possède aucune autonomie en ce qui a trait à la planification, l'exécution, le suivi et le contrôle du budget, ainsi que dans la définition des politiques et la prise de décisions stratégiques. Il s'agit d'un système de degré d'autonomie 1 (automatique)⁴⁴.

Des documents promotionnels destinés aux entreprises concernant le logiciel *Chorus* indiquent que le système propose des fonctionnalités faisant usage de techniques d'apprentissage automatique :

Our autonomous digital workers and technologies mimic and learn business processes like people and can be augmented with AI (artificial intelligence) and ML (machine learning) capabilities e.g., knowledge, visual perception, learning and problem-solving⁴⁵.

D'autres documents de même nature indiquent que *Chorus* offre des fonctionnalités lui permettant l'automatisation de la prise de décision :

Work orchestration. Improve collaboration and customer interactions through automated decision making. Integrate with other systems and applications to orchestrate work across your entire organization.

Rules management. Define, deploy, execute, monitor and maintain a defined set of rules – including policies, requirements and conditional statements – to control or influence decisions across a specific process or your entire business⁴⁶.

Les fonctionnalités faisant usage de techniques d'apprentissage automatique pour automatiser les décisions justifieraient la catégorisation du SIA *Chorus* dans un degré d'autonomie supérieure (autonomie assistée, voire semi-autonome selon le type de fonctionnalité). Or, nous n'avons aucune confirmation ou indication que l'État français fait usage de ces fonctionnalités liées à l'apprentissage automatique. Tous les documents auxquels nous avons eu accès concernant l'implémentation et l'usage du système⁴⁷ concernent des fonctionnalités qui correspondent à notre description du degré d'autonomie 1 (automatique). *Chorus* est en cela un bon exemple de la différence qui existe parfois entre les capacités techniques potentielles d'un système et le travail effectivement accompli par le système à la suite de son déploiement au sein d'une organisation. C'est également un bon exemple du caractère évolutif de ces systèmes : les techniques d'apprentissage automatique permettant l'assistance à la décision dont semble doté *Chorus*, si elles ne sont pas déjà utilisées par l'État français, le seront peut-être à l'avenir.

44 Depuis quelques années, l'AIFE a doté différentes applications du système *Chorus* d'assistants virtuels pour dialoguer avec les utilisateurs (citoyens ou employés). En 2021, trois applications du SIA *Chorus* utilisent déjà ces solutions : *Chorus Pro* avec le *chatbot* « ClaudIA », *Chorus Cœur* et *Chorus Formulaire* avec « IZIA » et *Chorus DT* avec « lanis ». Ces outils conversationnels constituent des cas d'application en eux-mêmes dans le sens où ils servent d'interface utilisateur avec le système et ne sont pas nécessaires à son fonctionnement. Dans la présentation qui précède (et en particulier pour la détermination du degré d'autonomie), nous nous concentrons sur *Chorus* en tant que système d'information financière et ignorons les interfaces de type *chatbot*.

45 SS&C | Blue Prism® Enterprise/*Chorus* : <https://files.blueprism.com/uploads/resources/pdf/enterprise-datasheet.pdf>

46 SS&C | Blue Prism®/*Chorus* : <https://files.blueprism.com/uploads/resources/pdf/chorus-data-sheet.pdf>

47 En plus des références déjà citées : *Le système d'information Chorus : développements et perspectives* (Diyani, 2017) : <https://www.cairn.info/revue-gestion-et-finances-publiques-2017-1-page-102.htm>



Pistes de réflexions éthiques

Bénéfices et avantages du SIA

- Augmente l'efficacité des organisations publiques. *Chorus* vise à renforcer l'efficacité et la performance de l'administration publique française en centralisant les données de divers ministères et agences, permettant ainsi une meilleure gestion des finances publiques. Cette centralisation des informations permet également une plus grande cohérence dans les processus administratifs.
- Améliore l'accès à l'information pour les gestionnaires. *Chorus* offre une meilleure vue d'ensemble sur les informations financières pour les gestionnaires des différents ministères et agences. Cet accès à l'information permet de résorber le problème fréquent d'asymétrie d'information entre les entités administratives. De plus, le système promeut la transparence et l'accessibilité de l'information, ce qui peut être bénéfique pour la reddition de comptes.
- Automatise des tâches routinières. L'automatisation de certaines tâches de gestion réduit la saisie manuelle des données et autres tâches chronophages, ce qui peut libérer du temps pour que les employés puissent se consacrer à d'autres tâches. *Chorus* permet également de générer automatiquement divers rapports, y compris des rapports financiers, ce qui facilite la reddition de comptes.

Inconvénients et risques du SIA

- Accessibilité et transparence de l'information. Il y a des inquiétudes concernant l'accès aux informations pour les utilisateurs (fonctionnaires et autres employés de l'État). Est-ce que le système sera suffisamment transparent et l'interface suffisamment conviviale pour permettre un accès uniforme à l'information ? De plus, bien que la transparence interne de l'appareil public semble améliorée avec *Chorus*, il n'est pas assuré que les citoyens bénéficient également de cette transparence.
- Acceptabilité et résistance au changement. Certains ministères ou agences pourraient être réticents face à la centralisation et au partage des informations sensibles qu'ils géraient auparavant en interne. Cette centralisation de l'information pourrait être perçue comme un risque de perte de contrôle sur certaines informations (problème de chasse gardée).
- Automatisation des tâches. L'automatisation pourrait entraîner des abolitions de poste et une surcharge de travail pour les employés restants. De plus, l'automatisation peut réduire le pouvoir discrétionnaire des employés, entraîner une diminution de leur autonomie et une transformation de leurs tâches. L'introduction de systèmes automatisés peut mener à négliger certaines valeurs telles que l'empathie, l'autonomie et l'indépendance des employés.
- Coût du système. Les coûts liés au déploiement, à la maintenance et à l'actualisation de ce système pourraient ne pas être justifiés par les gains en efficacité ou en efficience.
- Dépendance vis-à-vis du secteur privé. Le fait que le logiciel *Chorus* et sa maintenance soient gérés par le secteur privé (SAP) suscite des inquiétudes concernant la dépendance du secteur public envers les entreprises privées. Cette dépendance pourrait être coûteuse et pas toujours avantageuse pour le secteur public. De plus, cette dépendance vis-à-vis du secteur privé risque d'entraîner une perte d'expertise interne au secteur public.
- Risques d'abus de pouvoir et atteintes à la vie privée. L'accès facilité à l'information pour les gestionnaires peut conduire à des abus de pouvoir, des perturbations de l'organisation du travail et des atteintes à la vie privée des citoyens.
- Sécurité du système. La centralisation des données et des informations augmente la vulnérabilité du système face aux cyberattaques. Une seule intrusion pourrait donner accès à l'ensemble des données. De plus, la centralisation peut rendre le système fragile, ce qui pose des questions sur la robustesse de l'administration publique en cas de dysfonctionnement d'une partie du système qui pourrait potentiellement le perturber dans sa totalité.

Valeurs à prendre en considération pour le déploiement et l'utilisation de ce SIA (ou d'un système similaire)

- Autonomie des employés. L'automatisation des tâches peut entraîner une diminution de l'autonomie des employés. Il est important de veiller à ce que les employés aient un contrôle suffisant sur leur travail et qu'ils comprennent bien les systèmes automatisés qui sont utilisés.
- Continuité du service public. Le déploiement du SIA doit garantir la continuité des services publics, et ne pas compromettre la fourniture de ces services. La centralisation des informations dans un seul système soulève des questions quant à la robustesse de *Chorus*. Il est nécessaire de garantir que le système reste opérationnel même en cas de panne d'une de ses composantes. De plus, la centralisation des données peut augmenter le risque de cyberattaques. Il est donc essentiel d'assurer la sécurité du système.
- Humains dans la boucle. Il est important de maintenir une présence humaine pour assurer une supervision des processus automatisés et évaluer l'impact du système.
- Indépendance du secteur public. Le recours à des sociétés privées pour la maintenance du logiciel de *Chorus* suscite des préoccupations quant à la dépendance du secteur public envers le secteur privé. Le secteur public doit maintenir un contrôle suffisant sur ses propres systèmes.
- Reddition de comptes. Il est nécessaire d'avoir une reddition de comptes claire et transparente concernant le fonctionnement du système lui-même.
- Rentabilité. Le déploiement du système représente un investissement significatif, et il est donc nécessaire de s'assurer que les gains en matière d'efficacité justifient les coûts.
- Transparence pour le citoyen. Une réflexion s'impose quant à la manière dont le système peut améliorer la transparence de la gestion des finances publiques pour les citoyens, et non seulement pour les employés de l'État.

Comment protéger ou restaurer les valeurs affectées par le SIA?

- Automatisation. Il est important d'évaluer les impacts de l'automatisation sur le travail, notamment en termes de réduction de l'autonomie des employés et d'éventuelles abolitions de postes. Les tâches à automatiser devraient être clairement définies, et le processus d'automatisation devrait être transparent.
- Continuité du service public. Pour assurer la continuité des services, il est recommandé d'avoir des redondances et des systèmes de sauvegarde des données pour protéger les informations en cas de piratage ou de dysfonctionnement.
- Évaluation continue de l'efficacité du système. Pour assurer une supervision et évaluer l'impact du système, il est suggéré de maintenir une présence humaine dans le processus (humain dans la boucle). Cette évaluation d'impact devrait être effectuée en continu afin de mesurer l'efficacité du système et d'identifier rapidement les problèmes potentiels.
- Robustesse du système. Il est suggéré d'avoir des systèmes basés sur des modules indépendants pour éviter une centralisation excessive et renforcer la robustesse du système. Il est également suggéré de maintenir un certain niveau d'autonomie locale.
- Transparence et accessibilité de l'information. Pour s'assurer que l'information soit transparente et accessible, il est nécessaire de rendre l'information concernant le fonctionnement du système disponible pour un processus de vérification. Il devrait être possible de savoir d'où provient l'information, à quel endroit elle est hébergée et de quelle façon elle est traitée. Il est par ailleurs suggéré d'avoir un second niveau de reddition de comptes concernant le fonctionnement du système lui-même.

ABLE

Tableau 5. Description synthétique d'ABLE

Chronologie	2019 : projet pilote avec Siemens 2020 : mise en place par Hayden AI
Pays	États-Unis
Organisation	NYC Transit en collaboration avec le New York City Department of Transportation (NYCDOT)
Développeur	Privé (Hayden AI)
Bénéficiaire direct	Usagers
Degré d'autonomie	2. Autonomie limitée



Objectif poursuivi

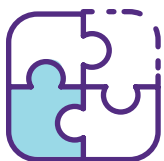
L'objectif du système ABLE (Automated Bus Lane Enforcement) de New York est de réduire les embouteillages, d'améliorer le respect des couloirs réservés aux autobus et de renforcer l'efficacité et la fiabilité des services de transport par autobus de la ville. Cet objectif s'inscrit dans le cadre du plan *Fast Forward* de NYC Transit⁴⁸, qui vise à améliorer les services d'autobus publics afin d'attirer de nouveaux usagers⁴⁹. Outre le NYC Transit, le programme implique le Metropolitan Transportation Authority (MTA), le New York City's Department of Transportation (NYCDOT) et le New York Police Department (NYPD) et vise initialement à faire respecter les règles relatives aux couloirs d'autobus dans les zones à forte congestion routière⁵⁰. **Pour ce faire, les différentes instances ont collaboré avec des entreprises privées pour mettre en place un système permettant d'enregistrer les plaques d'immatriculation et autres informations pertinentes concernant les véhicules en infraction, afin de pouvoir émettre des contraventions⁵¹.** Permettant de détecter et pénaliser les véhicules qui occupent illégalement les voies réservées aux autobus, le système vise l'amélioration du temps de déplacement pour les usagers et contribue à un système de transport urbain plus fluide et plus respectueux de l'environnement.

48 *Fast Forward: The Plan to Modernize New York City Transit*, s. d.) : https://static1.squarespace.com/static/5afef986c3c16a2dc6705929/t/5b072571f950b7a5e621a4ff/1527194994914/Fast+Forward+Plan_05-24-2018_3.15PM.pdf

49 *Automated bus lane traffic enforcement begins October 7 with new camera technology on MTA buses* | Mass Transit, 2019: <https://www.masstransitmag.com/bus/vehicles/press-release/21107006/mta-new-york-city-transit-automated-bus-lane-traffic-enforcement-begins-october-7-with-new-camera-technology-on-mta-buses>

50 *New York City to install Siemens camera system for bus lane enforcement* | Traffic Technology Today, 2019: <https://www.trafficechnologytoday.com/news/enforcement/new-york-city-to-install-siemens-camera-system-for-bus-lane-enforcement.html>

51 *New York City Uses Mobile Camera System to Catch Vehicles in Bus Lanes* | Vision Systems Design, 2022: <https://www.vision-systems.com/embedded/article/14282607/new-york-city-uses-mobile-camera-system-to-catch-vehicles-in-bus-lanes>



Étapes de développement

En 2019, NYC Transit et l'entreprise allemande Siemens ont collaboré à un projet pilote nommé *ABLE*, visant l'installation de caméras LaneWatch de Siemens sur trois parcours d'autobus à Manhattan et Brooklyn⁵². Le projet visait à évaluer l'efficacité du contrôle automatisé des voies réservées aux autobus et son impact sur la durée et la vitesse de déplacement. Le projet pilote fait suite à une validation de concept par NYC Transit, qui a constaté que ce type de système pouvait recueillir suffisamment d'informations pour procéder à l'émission de constats d'infraction au Code de la route dans les couloirs d'autobus. La MTA a donc attribué un contrat de 6.2 millions de dollars à Siemens Mobility Inc. pour l'achat et l'installation des caméras et des systèmes informatiques correspondants. Les caméras installées sur les bus de la MTA s'ajoutent aux caméras fixes installées dans les rues par le NYCDOT pour faire respecter les couloirs d'autobus, et les initiatives de la police de New York pour dissuader le stationnement dans ces mêmes couloirs. Dans le cadre du projet pilote, le système *ABLE* est déployé sur 123 autobus, et il est prévu de l'étendre avec un budget de 85 millions de dollars dans le plan d'investissement pour 2020-2024⁵³.

Dans le cadre de ce programme, les conducteurs qui restent dans une voie réservée aux autobus sans en sortir au premier virage à droite possible ou ceux bloquant la voie réservée à deux autobus successifs sont identifiés par le système comme étant en situation potentielle d'infraction. Les éléments de preuves sont alors rendus disponibles aux agents de contrôle qui les révisent et émettent un constat d'infraction au besoin⁵⁴. Selon Siemens, le système a détecté 9 700 infractions dans les deux mois suivant son déploiement. L'entreprise indique aussi qu'à la fin de cette période de deux mois, la vitesse moyenne des autobus avait augmenté de 55 % sur les trajets visés par le projet pilote⁵⁵. Les responsables des transports publics ont de leur côté déclaré que l'assistance technologique au respect des règles par le système *ABLE* permettait d'économiser les ressources de la police et de libérer les chauffeurs d'autobus de l'obligation d'intervenir auprès des contrevenants⁵⁶.

Suite aux résultats positifs du projet pilote, deux entreprises privées américaines saisissent l'opportunité et collaborent à l'expansion du programme. Il s'agit de Conduent Transportation et Hayden AI. La première est spécialisée dans les services et solutions en matière de processus d'entreprise, et la seconde se présente comme innovatrice en matière de mobilité urbaine et de solutions pour la ville numérique. La technologie de Hayden AI est conçue pour atteindre les mêmes objectifs que le système de Siemens utilisé dans le projet pilote *ABLE*⁵⁷. L'idée derrière cette technologie est simple : si les couloirs d'autobus sont dégagés, les autobus circuleront plus vite et les usagers emprunteront davantage les transports en commun⁵⁸. Du côté d'Hayden AI, il est prévu d'optimiser cette technologie et d'en étendre le rôle à d'autres aspects du transport et de la sécurité dans les villes.

52 *First-ever mobile bus lane enforcement solution in New York | Short News | Siemens Mobility USA*, s. d.: <https://press.siemens.com/global/en/pressrelease/siemens-mobility-launches-first-ever-mobile-bus-lane-enforcement-solution-new-york>; *New York City to install Siemens camera system for bus lane enforcement | Traffic Technology Today*, 2019: <https://www.traffictechnologytoday.com/news/enforcement/new-york-city-to-install-siemens-camera-system-for-bus-lane-enforcement.html>

53 *New York City Transit expands bus lane enforcement program - Bus - Metro Magazine*, 2019: <https://www.metro-magazine.com/10030873/new-york-city-transit-expands-bus-lane-enforcement-program>

54 Pendant les 60 premiers jours de la période de grâce initiale, les automobilistes bloquant les voies de bus reçoivent un avertissement sans amende. À la fin de cette période, les automobilistes qui continuent de bloquer les voies de bus seront passibles d'une amende allant de 50 dollars pour la première infraction, à 250 dollars pour la cinquième infraction et chaque infraction supplémentaires au cours d'une période de 12 mois. Dans *New York City Transit expands bus lane enforcement program - Bus - Metro Magazine*, 2019: <https://www.metro-magazine.com/10030873/new-york-city-transit-expands-bus-lane-enforcement-program>

55 *New York City Uses Mobile Camera System to Catch Vehicles in Bus Lanes | Vision Systems Design*, 2022: <https://www.vision-systems.com/embedded/article/14282607/new-york-city-uses-mobile-camera-system-to-catch-vehicles-in-bus-lanes>

56 *New York City Deploys Video Bus Lane Enforcement*, 2022: <https://www.govtech.com/fs/new-york-city-deploys-video-bus-lane-enforcement>

57 *Conduent Transportation and Hayden AI Announce Technology*, 2021 : <https://www.globenewswire.com/news-release/2021/04/06/2205051/0/en/Conduent-Transportation-and-Hayden-AI-Announce-Technology-Partnership-to-Improve-Bus-Lane-Performance-and-Traffic-Safety.html>; *MTA to Install ABLE Cameras on 300 Additional Buses*, 2022 : <https://new.mta.info/press-release/mta-install-able-cameras-300-additional-buses-across-nine-routes-end-of-2022>

58 *Conduent, Hayden AI Team Up to Automate Bus Lane Enforcement*, 2021: <https://www.govtech.com/biz/conduent-hayden-ai-team-up-to-automate-bus-lane-enforcement.html>

4. Cas d'application de l'utilisation de l'IA dans le secteur public

Dans la foulée de l'installation réussie des 300 premiers systèmes de contrôle automatisé des voies de circulation réservées aux autobus (ABLE) par Hayden AI, une nouvelle expansion du programme a été convenue par le MTA et le NYCDOT, à l'automne 2022, pour couvrir jusqu'à 85% des voies d'autobus existantes d'ici la fin de 2023, ajoutant 300 nouveaux systèmes de vision embarquée d'Hayden AI.

En plus de collecter des données permettant de faire respecter le Code de la route, le système de Hayden AI recueille des données variées que les villes peuvent utiliser pour la gestion des stationnements, le suivi des actifs de la ville, l'analyse des schémas de circulation, les mises à jour cartographiques en temps réel, l'identification des dangers sur les routes et les trottoirs, et bien d'autres choses encore.

En effet, utilisant les différents capteurs et équipements d'analyse introduits dans les autobus dans le contexte du programme ABLE, l'entreprise et son système sont à même de générer des alternatives ou des améliorations pour optimiser l'organisation urbaine, s'insérant dans les processus de prise de décision et dans les processus d'application de règles et règlements. L'entreprise dit proposer une vision innovante de l'urbanisme avec une technologie qui facilite la résolution de problèmes de sécurité et d'efficacité urbaine :

Hayden AI is quickly becoming a world leader in smart enforcement technologies. The company was founded on the belief that by combining computer vision with artificial intelligence, we can help governments automate manual processes, thus making public agencies more effective. To that end, we have leveraged our patented deep learning platform to partner with transit agencies across the globe as they look to better understand the obstacles they face in providing riders with the quickest, safest and most positive ridership experience possible. In addition to the enforcement benefits of our systems, Hayden AI technology captures valuable traffic data that can also be used to increase performance. By providing agencies with real-time insights into how effective their transit systems are operating, we are empowering them to make the best decisions possible with their dedicated riders in mind⁵⁹.

Ainsi, l'entreprise propose une plateforme de gestion autonome du trafic conçue pour améliorer la mobilité urbaine et la qualité de vie des communautés. Elle utilise un nombre croissant d'appareils mobiles connectés au réseau permettant à ses clients, souvent des organisations municipales, d'accéder à des solutions de mobilité alimentées par l'intelligence artificielle, capables de détecter, de comprendre et de déterminer la cause des incidents de la circulation⁶⁰.

L'introduction du système ABLE semble avoir eu un impact significatif sur le secteur des transports publics de la ville de New York. Il a permis d'améliorer la fluidité du trafic et d'accroître l'efficacité globale des services. En outre, il a permis d'augmenter les recettes générées par les amendes, qui peuvent être réinvesties dans le système de transport collectif.

Pour Hayden AI, la mise en place des équipements nécessaires au fonctionnement du programme ABLE offre de nombreuses possibilités menant au développement de nouvelles applications brevetées comme le brevet 11,322,017⁶¹, émis en 2022 (*Systems and methods for managing traffic rules using multiple mapping layers with traffic management semantics*).

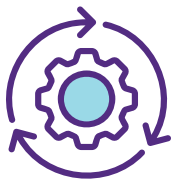
59 Transit Buses – Hayden AI, s. d : <https://www.hayden.ai/transit-buses>

60 Hayden AI Company Profile: Valuation & Investors | PitchBook, s. d.: <https://pitchbook.com/profiles/company/268302-52>

61 US Patent for Systems and methods for managing traffic rules using multiple mapping layers with traffic management semantics Patent (Patent # 11,322,017 issued May 3, 2022) - Justia Patents Search, s. d.: <https://patents.justia.com/patent/11322017>

The new patented technology facilitates the management of complex traffic rules using a traffic enforcement layer situated on top of a semantic map layer. The map layer, which consists of street elements such as intersections and stop signs, is continuously updated with data captured by Hayden AI edge devices already installed on municipal vehicles for automated enforcement. The enforcement layer consists of traffic rules that can be adjusted directly by the user or automatically via a traffic insight layer, which is generated based on traffic violations and traffic conditions. By analyzing the impact of traffic rules, the insight layer also provides suggestions for adjustments⁶².

En fournissant plusieurs méthodes pour ajouter et modifier les règles de circulation à même l'application, le système peut aider les organisations publiques à modifier les restrictions inefficaces et à s'assurer que l'application des règles entraîne une réduction des embouteillages, tout en améliorant la sécurité routière et l'efficacité des transports.



Fonctionnement du système

Le système *ABLE* est un système de contrôle automatisé des voies de circulation réservée des autobus équipé de deux caméras fixées sur le pare-brise intérieur des autobus. La première enregistre les informations relatives aux plaques d'immatriculation, tandis que la seconde filme un plan à 80 degrés montrant le contexte de l'infraction, tel que la marque et le modèle du véhicule, l'emplacement et le sens de circulation. Un logiciel intégré examine ces données et transfère les métadonnées et vidéos associées à une potentielle infraction vers un logiciel basé dans le *cloud*. Les informations non pertinentes sont effacées, puis les agents de contrôle examinent les données et décident s'il convient d'émettre une contravention.

Le système est basé sur des strates d'algorithmes logiciels qui travaillent ensemble pour traiter et analyser efficacement les données captées. Les algorithmes sont développés à l'interne, spécialement pour le système *ABLE*, par une équipe de scientifiques spécialisés dans la vision informatique et l'apprentissage automatique⁶³. Ces algorithmes comprennent :

- une strate qui analyse les images vidéo pour identifier les véhicules et leurs positions relatives ;
- une strate de cartographie et de réglementation dans laquelle des cartes sémantiques sont intégrées à un référentiel de règles et de lois applicables en matière de stationnement ;
- un module d'inférence, qui combine les données de plusieurs capteurs pour reconstruire numériquement l'événement en 3D ;
- un module de raisonnement, capable de comprendre le contexte et la gravité des situations afin de filtrer les erreurs.

62 Hayden AI Granted Patent for Traffic Rule Management Technology, 2022: <https://www.prnewswire.com/news-releases/hayden-ai-granted-patent-for-traffic-rule-management-technology-301576449.html>

63 New York City Uses Mobile Camera System to Catch Vehicles in Bus Lanes | Vision Systems Design, 2022: <https://www.vision-systems.com/embedded/article/14282607/new-york-city-uses-mobile-camera-system-to-catch-vehicles-in-bus-lanes>

4. Cas d'application de l'utilisation de l'IA dans le secteur public

Les différentes strates d'algorithmes travaillent ensemble pour traiter et analyser efficacement les données captées afin de déterminer si une infraction a été commise. Voici quelques éléments clés du système⁶⁴ :

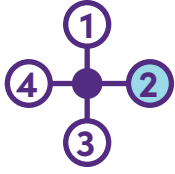
- **Prétraitement des images.** Les images brutes et les vidéos captées par les caméras installées sur les bus contiennent souvent du bruit, des conditions d'éclairage variables et d'autres artefacts visuels. Le système d'IA prétraite ces images afin d'en améliorer la qualité et de les rendre compatibles avec une analyse plus poussée.
- **Détection et reconnaissance des plaques d'immatriculation.** Après avoir détecté les véhicules, le système d'IA se concentre sur la localisation et la reconnaissance des plaques d'immatriculation. Il utilise des algorithmes spécialisés conçus pour identifier les caractéristiques uniques des plaques d'immatriculation, telles que leur forme, leur taille et les motifs des caractères. Des techniques de reconnaissance optique de caractères (OCR) sont ensuite utilisées pour extraire les numéros de plaques d'immatriculation.
- **Détermination des infractions.** Le système d'IA utilise des informations contextuelles telles que les panneaux de signalisation, le marquage au sol et les restrictions relatives aux voies réservées pour déterminer si un véhicule enfreint les règles relatives aux voies d'autobus. Ce processus implique l'analyse de la position, du mouvement et de la durée d'occupation du véhicule dans le couloir réservé. Le système prend également en compte des facteurs tels que l'heure, les conditions météorologiques et les exemptions temporaires (par exemple, les véhicules effectuant des virages à droite autorisés).
- **Apprentissage automatique.** Le système *ABLE* s'appuie sur des techniques d'apprentissage automatique supervisé et non supervisé pour améliorer sa précision et ses performances au fil du temps. Au fur et à mesure que le système traite davantage de données et apprend à partir d'exemples réels, il devient plus performant pour détecter les violations et minimiser les faux positifs ou négatifs.



Bénéficiaire direct du système

Le système *ABLE* est mis en place, opéré et consulté par différentes instances, dont le NYC Transit, le NYCDOT, le MTA et le département de la police de New York (NYPD). Ces agences collaborent pour faire respecter les réglementations relatives aux couloirs d'autobus et s'assurer qu'elles sont bien respectées. L'objectif du système *ABLE* (le travail qu'il accomplit) est de rendre plus efficient le service de transport collectif par autobus de la ville de New York en documentant les infractions à ces réglementations. En ce sens, les passagers des autobus sont les bénéficiaires directs du système qui contribue à l'amélioration de leur vitesse de déplacement et à leur appréciation générale du service offert. L'organisation bénéficie indirectement du système qui assiste le NYC Transit dans sa mission d'optimisation de la qualité du service offert à la clientèle. D'autres acteurs tels que les conducteurs des autobus, les policiers ou autres employés responsables d'intervenir lors des infractions en sont également des bénéficiaires indirects.

64 Transit Buses – Hayden AI, s. d.: <https://www.hayden.ai/applications/transit-buses>; *Making Cities More Efficient with Computer Vision and Edge AI*, a Presentation from Hayden AI - Edge AI and Vision Alliance, 2022: <https://www.edge-ai-vision.com/2022/10/making-cities-more-efficient-with-computer-vision-and-edge-ai-a-presentation-from-hayden-ai/>



Degré d'autonomie du système

Les capacités de détection et d'application basées sur l'IA du système *ABLE* minimisent le besoin d'intervention humaine, garantissant que les infractions sont détectées et sanctionnées de manière opportune et efficace. Cependant, il repose sur des règles et règlements prédéfinis (Code de la route et règles concernant les couloirs réservés aux autobus) et nécessite la vérification humaine pour garantir l'exactitude de l'information transmise par le système, prendre les décisions concernant l'émission des constats d'infraction et traiter les litiges potentiels. Le système *ABLE* fonctionne avec un degré d'autonomie 2 (autonomie limitée)⁶⁵.



Pistes de réflexions éthiques

Bénéfices et avantages du SIA

- Optimise le service de transport en commun. Le système contribue à améliorer la circulation en identifiant et sanctionnant les comportements non conformes à la réglementation applicable aux voies réservées au transport collectif.
- Favorise l'utilisation du transport en commun. En dissuadant les comportements non conformes, le système contribue à rendre le transport collectif plus attrayant et encourage les citoyens à opter pour ce mode de transport.
- Accentue le respect des règles de circulation en général et favorise le développement durable.
- La confirmation des infractions et l'émission des constats d'infraction sont effectuées par un employé qui valide les informations transmises par le système.

Inconvénients et risques du SIA

- Surveillance permanente. Ce système implique une surveillance constante qui peut être considérée intrusive par certains usagers de la voie publique. D'autant plus qu'il capte des images/informations concernant des individus et des situations qui ne sont pas en lien avec l'application de la réglementation des voies réservées aux autobus. Par exemple, si un piéton se trouve à côté de la voiture en infraction, il sera également filmé. Se pose également la question de savoir si les données seront conservées pour en extraire plus tard des informations qui ne sont pas liées aux infractions ciblées par le système.
- Augmentation potentielle de la criminalisation même s'il n'y a pas davantage d'infractions qu'avant le déploiement du système. Des comportements ou actions qui étaient acceptés ou tolérés sont maintenant mis en évidence et potentiellement punis par un constat d'infraction, suite à l'application stricte et systématique des règles par le système.
- Utilisation et partage des données. Les données collectées par le système (sauvegardées dans le *cloud*) pourraient être utilisées de manière malveillante si elles ne sont pas correctement sécurisées ou si elles sont partagées avec d'autres acteurs (même si les gestionnaires du système en font une utilisation responsable et appropriée). Il semble que les données/informations non pertinentes à la tâche du système sont effacées suite à l'analyse et la validation par un employé. Qui est chargé de déterminer quelles sont les données/informations non pertinentes et selon quels critères?
- Risque de profilage. Les algorithmes d'apprentissage automatique du SIA pourraient mener à un ciblage disproportionné de certains groupes ou comportements spécifiques, ce qui pourrait conduire à des discriminations. Par exemple, si le système constate que les femmes conduisant de vieilles voitures noires à 8h00 le matin commettent davantage d'infractions, est-ce qu'il en viendra, en s'entraînant sur les données, à accentuer sa surveillance de ces individus/véhicules plus que les autres?

⁶⁵ Il est question, avec des technologies nouvellement brevetées par Hayden AI, d'utiliser l'apprentissage automatique de manière à ce que le système puisse suggérer de façon autonome des approches novatrices d'optimisation et de sécurisation de la circulation sur les voies publiques. Ceci pourrait faire passer ce SIA vers une catégorisation de degrés d'autonomie 3 (semi-autonome).

4. Cas d'application de l'utilisation de l'IA dans le secteur public

- Perception des circonstances atténuantes. Le système pourrait ne pas être capable de comprendre ou d'interpréter correctement les situations complexes ou nuancées.
- Risque d'excès de confiance dans le système. Avec un taux de succès élevé d'identification des infractions par le système, les *humains dans la boucle* pourraient devenir moins vigilants et accepter mécaniquement les propositions/décisions du SIA. D'autant plus que les employés responsables de confirmer les infractions n'ont accès qu'aux données et informations contextuelles fournies par le système (il n'y a aucun humain dans la boucle sur le terrain pour analyser le contexte de la situation d'infraction constatée par le système).

Valeurs à prendre en considération pour le déploiement et l'utilisation de ce SIA (ou d'un système similaire)

- Équité. Le système doit être conçu de manière à éviter le ciblage discriminatoire ou injuste.
- Pouvoir discrétionnaire. Bien que le SIA prenne certaines décisions, il est important d'avoir un *humain dans la boucle* ayant accès à suffisamment d'information pour valider les décisions et prendre en compte les nuances que le système pourrait manquer ou ne pas inclure dans son analyse.
- Proportionnalité. Les moyens employés pour atteindre les objectifs du système doivent être proportionnels et non excessifs ou intrusifs.
- Respect de la vie privée. L'omniprésence de la surveillance et l'accumulation de données concernant les citoyens (souvent à leur insu) doivent être considérées de manière appropriée. Avec ses deux caméras embarquées sur chacun des autobus, ce système capte une grande quantité d'informations superflues, qui ne sont pas pertinentes pour accomplir sa tâche. Ces données/informations doivent être gérées dans le respect des usagers de la voie publique.
- Sécurité des données. Des mesures appropriées doivent être prises pour assurer la sécurité des données collectées par le système.
- Transparence. Les usagers doivent être informés de l'existence du système, de la manière dont il fonctionne et des implications possibles.

Comment protéger ou restaurer les valeurs affectées par le SIA?

- Équité. Veiller à éviter la discrimination ou le ciblage d'individus ou de comportements en surveillant l'évolution de l'apprentissage automatique du système et en intervenant si nécessaire.
- Pouvoir discrétionnaire. Maintenir un élément humain dans le processus de décision pour vérifier les informations fournies par le système et déterminer leur pertinence et leur exactitude. La vigilance humaine est importante surtout pour les cas complexes ou nuancés. Réfléchir au risque d'excès de confiance dans le système, en particulier lorsque la performance de celui-ci augmente.
- Proportionnalité. Examiner si les moyens employés pour atteindre les objectifs du système sont proportionnels et ne sont pas excessifs ou intrusifs. Se demander s'il existe des moyens moins invasifs permettant d'aboutir au même résultat.
- Respect de la vie privée. S'interroger sur les considérations éthiques générales soulevées par ce SIA, notamment la surveillance constante et ses impacts sur les usagers de la voie publique.
- Sécurité des données. Mettre en place des mécanismes pour sécuriser les données collectées et éviter leur utilisation malveillante, en particulier par des tiers.
- Transparence. Informer les conducteurs (et les autres usagers de la voie publique) de l'existence du système de détection et du fait qu'ils peuvent être filmés, ou recevoir des amendes, suite à leur identification par le système.

Tableau 6. Description synthétique de ToxCast

Chronologie	2007 : projet pilote en 2007 2010 : phase 1 2013 : phase 2 2018 : phase 3
Pays	États-Unis
Organisation	Agence américaine de protection de l'environnement (EPA)
Développeur	Partenariat public-privé (EPA et des milieux de recherche publics et privés)
Bénéficiaire direct	Employés
Degré d'autonomie	2. Autonomie limitée



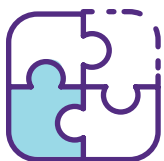
Objectif poursuivi

Les analystes de l'Agence américaine de protection de l'environnement (EPA) ont développé un programme appelé *ToxCast* pour aider l'agence à prédire la toxicité des composés chimiques. La toxicité des produits chimiques est traditionnellement établie à l'aide de tests sur les animaux, mais ces techniques de laboratoire sont coûteuses et prennent beaucoup de temps, sans compter qu'elles sont souvent nocives pour les animaux. Les technologies automatisées de criblage chimique, appelées tests de criblage à haut débit (high-throughput screening, HTS), sont utilisées pour exposer des cellules vivantes, des protéines isolées ou d'autres molécules biologiques à des produits chimiques. Les cellules ou les protéines sont ensuite analysées pour détecter les changements dans l'activité biologique qui peuvent suggérer des effets toxiques potentiels. Ces méthodes innovantes peuvent limiter le nombre de tests de toxicité sur les animaux de laboratoire, tout en permettant de cribler rapidement et efficacement des milliers de produits chimiques pour en déterminer les effets potentiels sur la santé⁶⁶.

Confrontée à des dizaines de milliers de produits chimiques susceptibles d'être soumis à la réglementation de l'EPA, *ToxCast* permet de déterminer, parmi la multitude de produits chimiques en production, ceux qui devraient faire l'objet de tests plus approfondis. *ToxCast* applique des algorithmes d'apprentissage automatique – en particulier l'analyse discriminante linéaire – aux données sur les interactions des produits chimiques obtenues lors de tests in vitro afin de prédire leur toxicité⁶⁷. La raison d'être du programme *ToxCast* est ainsi de générer des données qui permettent de classer les produits chimiques par ordre de priorité en vue d'essais supplémentaires, de soutenir l'évaluation des risques chimiques et d'alimenter la prise de décision réglementaire.

66 *NAMs Training: Getting Started* | US EPA: <https://www.epa.gov/chemical-research/nams-training-getting-started>

67 *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era* (Coglianese & Lehr, 2017, p. 1162): <https://www.law.upenn.edu/live/files/6329-coglianesi-and-lehr-regulating-by-robot-penn-ile>



Étapes de développement

ToxCast a débuté en tant que programme pilote exploratoire au sein du National Center for Computational Toxicology (NCCT) nouvellement créé par l'EPA en 2007, l'année même où le National Research Council (NRC) indiquait un changement de paradigme en publiant son rapport intitulé *Toxicity Testing in the 21st Century*⁶⁸. Le programme vise alors à évaluer l'utilisation d'analyses *in vitro* pour comprendre les types de perturbations causées par l'exposition aux produits chimiques, et les modèles de hiérarchisation informatisés pour la toxicité *in vivo*. Dans la phase I, les chercheurs ont choisi 309 produits chimiques pour lesquels des données de toxicité étaient déjà disponibles. À l'aide de 467 analyses *in vitro* relevant de neuf technologies, ils ont étudié un large éventail d'activités chimiques au niveau moléculaire et au niveau des voies de transmission. La comparaison de ces résultats *in vitro* avec les données existantes a aidé les chercheurs à élaborer des modèles de hiérarchisation pour prédire la toxicité de produits chimiques similaires, mais non testés :

Chemical toxicity testing is being transformed by advances in biology and computer modeling, concerns over animal use, and the thousands of environmental chemicals lacking toxicity data. The U.S. Environmental Protection Agency's ToxCast program aims to address these concerns by screening and prioritizing chemicals for potential human toxicity using *in vitro* assays and *in silico* approaches⁶⁹.

Sur la base des exemples de la phase I, les chercheurs ont été convaincus que les données issues des tests de criblage à haut débit peuvent aider à prédire les mécanismes d'action de nombreux autres produits chimiques bien étudiés et indiquer quelles autres voies biologiques peuvent également être activées. Ces données jettent les bases d'un dépistage des produits chimiques non testés et permettent de fournir des orientations essentielles pour les futures analyses⁷⁰. La phase I des tests s'est officiellement achevée avec la publication des données *ToxCast* en janvier 2010. La phase II du programme a élargi la taille et la diversité de la chimiothèque soumise au dépistage *ToxCast* à 1878 produits chimiques et s'est officiellement achevée avec la publication des données en décembre 2013.

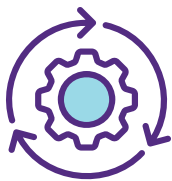
À l'issue de la phase III en 2018, *ToxCast* a analysé plus de 4 500 produits chimiques dans plus de 700 criblages à haut débit. Parallèlement aux trois phases d'expansion de la bibliothèque, le pipeline d'analyse des données utilisé pour traiter les résultats des analyses a fait l'objet de mises à jour régulières jusqu'à aujourd'hui. Ces changements résultent de l'évolution des technologies de dosage, des technologies informatiques d'analyse et de prédiction, et des priorités changeantes des programmes de l'EPA⁷¹.

68 *Toxicity Testing in the 21st Century: A Vision and a Strategy* | The National Academies Press, 2007: <https://nap.nationalacademies.org/catalog/11970/toxicity-testing-in-the-21st-century-a-vision-and-a>

69 *In vitro screening of environmental chemicals for targeted testing prioritization: the ToxCast project* - PubMed, 2010: <https://pubmed.ncbi.nlm.nih.gov/20368123/>

70 *ToxCast on Target : In Vitro Assays and Computer Modeling Show Promise for Screening Chemicals* (Haynes, 2010): <https://doi.org/10.1289/ehp.118-a172a>

71 *Purposes and Uses of EPA's Toxicity ForeCaster (ToxCast™)*, 2022: <https://www.americanchemistry.com/better-policy-regulation/research/long-range-research-initiative-lri/resources/purposes-and-uses-of-epa-s-toxicity-forecaster-toxcast>



Fonctionnement du système

Puisque le modèle informatique constitue un élément parmi un ensemble de technologies utilisées dans le cadre plus large du programme *ToxCast*, nous commencerons par définir les grandes lignes de ce programme, pour ensuite nous attarder aux particularités des technologies d'intelligence artificielle qui y sont utilisées.

Programme de recherche ToxCast

1. Dans le cadre du programme *ToxCast*, on commence par sélectionner les produits chimiques à évaluer. Le programme se concentre sur les produits chimiques qui sont largement utilisés dans le commerce ou qui sont particulièrement préoccupants en raison de leur toxicité potentielle ou des niveaux d'exposition.
2. On procède ensuite à des analyses *in vitro* pour tester les effets des produits chimiques sur des cibles cellulaires et moléculaires spécifiques. On utilise pour ce faire le criblage à haut débit (*high-throughput screening*, HTS). Ces tests sont conçus pour imiter divers aspects de la biologie humaine, tels que la signalisation hormonale, les réponses au stress cellulaire et les processus de développement.
3. Chaque analyse génère une grande quantité de données sur les effets de la substance chimique sur la cible. Par exemple, une analyse peut mesurer les changements dans l'expression des gènes, les niveaux de protéines ou la viabilité des cellules en réponse à un produit chimique.

Successful implementation of any high-throughput screening (HTS) initiative requires substantial data science efforts to store and process the large quantities of data generated. The diversity of the ToxCast data necessitated the development of a flexible HTS data pipeline capable of efficiently processing and storing large volumes of heterogenous data (Filer et al., 2017, p. 618).

4. On utilise ensuite des modèles informatiques pour analyser les données et prédire la toxicité potentielle du produit chimique. Les modèles intègrent les données de plusieurs analyses pour fournir une image plus complète des effets du produit chimique sur la santé humaine. Les prédictions faites par *ToxCast* sont basées sur des modèles statistiques qui utilisent les données des analyses de criblage à haut débit (HTS) pour estimer la probabilité d'effets néfastes. Ces modèles sont entraînés sur un ensemble de produits chimiques dont la toxicité a été testée précédemment, ce qui permet de prédire la toxicité de nouveaux produits chimiques qui n'ont pas encore été testés. Les modèles informatiques permettent d'obtenir une estimation de la puissance de l'interaction et de la réaction indésirable probable⁷².
5. Les données générées peuvent être utilisées pour éclairer l'évaluation des risques chimiques et la prise de décision réglementaire. Par exemple, les données peuvent être utilisées pour hiérarchiser les produits chimiques devant faire l'objet de tests supplémentaires ou pour soutenir le développement d'alternatives plus sûres.

72 *The Next Generation Blueprint of Computational Toxicology at the U.S. Environmental Protection Agency* (Thomas et al., 2019) : <https://doi.org/10.1093/toxsci/kfz058>.

Approche technique

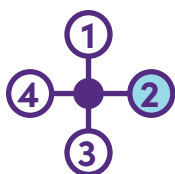
Différentes approches de traitement de données ont été appliquées au programme *ToxCast* depuis ses débuts en 2007. Il s'agit d'un bon exemple d'évolution de techniques d'IA utilisées pour remplir une tâche précise, les méthodes d'analyse des données évoluant et s'améliorant avec le temps⁷³. Le programme *ToxCast* vise précisément la recherche et le développement de modèles prédictifs permettant d'accélérer et de rendre plus efficaces les analyses de composantes et de produits toxiques présents dans l'environnement. À travers le temps, des techniques d'apprentissage supervisé et non supervisé ont été utilisées, en particulier l'approche de la carte auto-organisatrice (self-organizing map/SOM)⁷⁴. Des modèles d'apprentissage profond sont également testés et utilisés, tels les réseaux neuronaux convolutifs (CNN) et les réseaux neuronaux récurrents (RNN) qui sont considérés comme étant particulièrement appropriés pour analyser des données chimiques et biologiques⁷⁵.



Bénéficiaire direct du système

Les employés dont le travail est d'évaluer l'impact de différents composés chimiques sur l'humain sont ceux qui bénéficient directement du travail de ce SIA. Les algorithmes d'intelligence artificielle de ce système permettent d'accélérer et de rendre plus efficace l'analyse de milliers de composants chimiques et de leur impact sur l'humain. Cette assistance à l'évaluation des composants chimiques permet aux employés de se concentrer sur la prise de décision quant aux analyses supplémentaires à prioriser, ainsi que sur les mesures requises pour protéger le public des composants dangereux présents dans certains produits.

L'Agence américaine de protection de l'environnement (EPA) bénéficie indirectement du SIA qui l'aide à remplir un volet de sa mission de protection de la santé publique et de l'environnement, soit l'examen de sécurité des produits chimiques présents sur le marché.



Degré d'autonomie du système

L'analyse prédictive effectuée par ce système repose totalement sur la collecte, l'évaluation et la curation de données issues d'analyse in vivo et in vitro. La sélection des nouveaux composants chimiques faisant l'objet d'analyses prédictives est effectuée par les employés responsables de l'évaluation des risques. L'analyse prédictive permet de prioriser, parmi les nombreux exemples qui lui sont soumis, les composés chimiques qui devraient faire l'objet d'une attention particulière. *ToxCast* peut donc effectuer des tâches de manière indépendante dans un cadre spécifique, à savoir prédire la toxicité des produits chimiques à l'aide d'algorithmes d'apprentissage automatique. Il peut analyser des données et générer des prédictions basées sur les interactions entre les produits chimiques et les cibles biologiques. Le SIA sert ici d'outil d'aide à la décision, fournissant des informations pour éclairer le processus décisionnel des responsables de la santé publique et des autorités de réglementation. Il correspond au degré d'autonomie 2 (autonomie limitée).

73 *Purposes and Uses of EPA's Toxicity ForeCaster (ToxCast™)*, 2022: <https://www.americanchemistry.com/better-policy-regulation/research/long-range-research-initiative-lri/resources/purposes-and-uses-of-epa-s-toxicity-forecaster-toxcast>

74 SOM is a powerful statistical tool used for visualizing high-dimensional data, and exploring and clustering samples while preserving their distances. Sources: https://en.wikipedia.org/wiki/Self-organizing_map; *Self Organizing Maps-Medium*, 2018: https://medium.com/@abhinavr8/self-organizing-maps-ff5853a118d4_What_Are_Self_Organizing_Maps | *Beginner's Guide To Kohonen Map- Simplilearn*, 2023: <https://www.simplilearn.com/self-organizing-kohonen-maps-article>; *Towards Whole Health Toxicology : In-Silico Prediction of Diseases Sensitive to Multi-Chemical Exposures* (Arowolo et al., 2022, p. 9): <https://www.mdpi.com/2305-6304/10/12/764>

75 *From machine learning to deep learning : Progress in machine intelligence for rational drug discovery* (Zhang et al., 2017): <https://doi.org/10.1016/j.drudis.2017.08.010>, *DeepTox : Toxicity Prediction using Deep Learning* (Mayr et al., 2016): <https://doi.org/10.3389/fenvs.2015.00080>, *The Next Generation Blueprint of Computational Toxicology at the U.S. Environmental Protection Agency* (Thomas et al., 2019) : <https://doi.org/10.1093/toxsci/kfz058>



Pistes de réflexions éthiques

Bénéfices et avantages du SIA

- Accroît l'efficacité des tests de toxicité en identifiant parmi les nouveaux produits ceux qui nécessitent une attention particulière et prioritaire pour des tests plus avancés.
- Améliore l'efficacité (temps et argent) en ciblant les produits qui présentent le plus grand risque.
- Offre une assistance dans le processus de prise de décisions réglementaires concernant des produits potentiellement nocifs pour les humains, les êtres vivants ou l'environnement.
- Réduit l'utilisation d'animaux de laboratoire. En effectuant *in vitro* et *in silico* des tests de toxicité qui sont habituellement réalisés sur des animaux vivants, *ToxCast* permet de minimiser ou d'éviter l'expérimentation sur les animaux, ce qui réduit les impacts négatifs sur ceux-ci.

Inconvénients et risques du SIA

- L'évaluation toxicologique est un processus complexe et les recommandations concernant les concentrations acceptables varient souvent et sont différentes selon les juridictions, posant des questions sur la capacité du système à s'adapter à des réglementations et des standards scientifiques en évolution.
- Inquiétude quant à l'efficacité du système à imiter la biologie humaine et à produire des résultats aussi précis que les tests sur les animaux. En effet, le SIA teste les produits dans un environnement virtuel, ce qui soulève des questions sur la transposabilité des résultats dans des environnements réels complexes. De plus, si on s'appuie trop sur le système, on risque de manquer des cas très rares de toxicité. Cela pourrait conduire à la commercialisation de produits toxiques qui n'auraient pas été identifiés comme tels par le SIA.
- Les modèles statistiques du SIA sont basés sur les données disponibles. Si ces données sont limitées ou biaisées, cela pourrait affecter la précision des résultats du SIA.
- Incertitude quant à la responsabilité en cas de préjudice causé par des erreurs du système.
- L'introduction du SIA pourrait mener à une diminution du nombre d'employés, une perte d'expertise humaine pour le secteur public et un risque de surcharge de travail pour les employés restants.
- Le partenariat public-privé au cœur du SIA soulève des questions sur la gestion et la sécurité des données, en particulier si les pratiques du secteur privé ne sont pas en adéquation avec celles du secteur public.
- Préoccupations quant à l'accès aux algorithmes et à l'utilisation des données. Puisque les résultats fournis par le système sont utilisés pour mettre en place des décisions réglementaires concernant différents produits, certains acteurs pourraient être tentés de détourner ou influencer la réglementation s'ils ont accès aux algorithmes utilisés ou à des informations concernant leur fonctionnement.
- Inquiétude quant à une utilisation élargie abusive ou inappropriée du SIA. Constatant l'efficacité du système pour tester la toxicité de certaines molécules, des acteurs pourraient être tentés de le rentabiliser en généralisant son utilisation afin, par exemple, de tester des molécules ou phénomènes qui ne sont pas pris en compte dans la base de données initiale, ou pour lesquels les algorithmes d'apprentissage automatique du système sont inappropriés.

Valeurs à prendre en considération pour le déploiement et l'utilisation de ce SIA (ou d'un système similaire)

- **Fiabilité.** Les résultats fournis par *ToxCast* doivent être fiables, car les décisions prises à partir de ceux-ci peuvent avoir un impact important sur la santé des humains, des êtres vivants et de l'environnement.
- **Intégrité.** Le système doit être protégé contre l'utilisation malveillante ou l'exploitation à des fins différentes que celles initialement prévues. Cela est particulièrement pertinent lorsque le système est développé ou utilisé dans le cadre d'un partenariat public-privé.
- **Pertinence et proportionnalité.** Les décideurs doivent évaluer attentivement le rapport bénéfices-risques associé à l'utilisation du SIA, en tenant compte de toutes les incertitudes et des possibles effets indésirables.
- **Précaution / Prudence.** Compte tenu des risques liés à la toxicité des produits analysés, l'approche de précaution doit prévaloir en ce qui concerne les protocoles d'utilisation, l'implication de l'humain dans les processus de décision, et la mise en place de mesures de sécurité. Il sera important de vérifier jusqu'à quel point le système peut manquer des cas rares de toxicité et d'évaluer constamment ce risque.
- **Respect des réglementations.** Le SIA doit respecter les différentes réglementations concernant la toxicité et les concentrations acceptables, qui peuvent varier d'une juridiction à l'autre et changer régulièrement.
- **Responsabilité.** Il faut déterminer qui est responsable en cas de préjudice individuel ou collectif. Les protocoles d'indemnisation et les mécanismes de responsabilité doivent être clairement établis.
- **Supervision humaine.** Il est important de maintenir une présence humaine dans la boucle de décision pour superviser et valider les résultats du SIA. De plus, le système ne doit pas remplacer complètement l'expertise humaine, mais plutôt la compléter. Le maintien d'une expertise de pointe dans le secteur public est un enjeu important et qui s'accroît lorsqu'elle permet de s'assurer de la sécurité dans des domaines sensibles ou névralgiques.
- **Transparence.** Les processus décisionnels, les algorithmes utilisés, et les méthodes d'évaluation des risques doivent être transparents et compréhensibles par tous les acteurs concernés. Cela comprend la communication claire des limites du système et des paramètres qu'il utilise.

Comment protéger ou restaurer les valeurs affectées par le SIA?

- Créer des protocoles d'utilisation précis. Il est recommandé d'avoir des règles claires sur l'utilisation, ainsi que des mécanismes de surveillance, pour éviter un usage inapproprié ou excessif du système. Il faudra veiller à éviter, par exemple, un usage dérivé des données ou du système pour des tâches pour lesquelles il n'a pas été conçu.
- Impliquer un humain dans le processus décisionnel. Il est important de toujours avoir un humain impliqué dans la revue et la validation des résultats fournis par le SIA.
- Mettre en place un comité d'experts indépendants. Un comité externe devrait être établi pour effectuer des inspections aléatoires, évaluer l'adéquation de nouvelles expérimentations et maintenir une surveillance indépendante du système.
- Prévoir des plans d'intervention en cas de préjudice. Les décideurs devraient déterminer qui sera responsable en cas de préjudice individuel ou collectif, et établir la nature des dédommagements en cas de préjudice subi.
- Principe de prudence. Avant de mettre en place le système, les décideurs doivent s'assurer de son efficacité et de sa sécurité. Cela pourrait nécessiter de rassembler une diversité d'expertises et de procéder à une évaluation approfondie avant le déploiement du système.
- Hiérarchiser les priorités. Il est suggéré d'appliquer le cadre éthique *Accountability for Reasonableness* (Responsabilité de prendre des décisions raisonnables – RR⁷⁶) qui vise à garantir que le processus décisionnel soit transparent et équitable. Cette approche permet de discuter de valeurs concurrentes lorsque des priorités doivent être établies.
- Renforcer la protection des données. Étant donné l'implication du secteur privé dans le partenariat public-privé, il est crucial de renforcer les normes et de veiller à ce que les données soient correctement gérées, sécurisées et partagées.

76 « Élaboré à l'origine par Norman Daniels, le cadre de la « Responsabilité du caractère raisonnables » (RR) a été appliqué à un vaste éventail de scénarios relatifs à l'établissement de priorités, à l'allocation des ressources et au triage. La RR a notamment été utilisée dans le cadre de la planification des activités de lutte contre les pandémies, comme le SRAS et la grippe H1N1, mais elle peut également s'appliquer à des problèmes plus généraux du système de santé privé et public. Le cadre de la RR est généralement caractérisé par quatre valeurs touchant la procédure : la pertinence, la communication, les révisions et appels et la mise en application ». Source : Annexe I de <https://www.canada.ca/fr/service-correctionnel/organisation/bibliotheque/delinquants/sante/orientation-programme-troubles-lies-consommation-dopioides-16-aout-2021.html#annexe1>.

SKALA

Tableau 7. Description synthétique de SKALA

Chronologie	2015 : début du projet pilote dans deux districts. 2017 : extension de la zone couverte par le projet pilote à 4 districts supplémentaires. 2018 : déploiement officiel de SKALA auprès des autorités policières de 16 des districts de Rhénanie-du-Nord-Westphalie.
Pays	Allemagne
Organisation	Office criminel du Land de Rhénanie-du-Nord-Westphalie/The State Office for Criminal Investigation of North Rhine-Westphalia (LKA NRW)
Développeur	Public (Interne à LKA NRW). Collaboration avec l'Université de Constance
Bénéficiaire direct	Organisation
Degré d'autonomie	2. Autonomie limitée



Objectif poursuivi

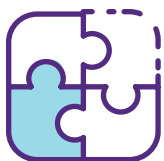
SKALA (System for Crime Analysis and Anticipation in Germany) est une technologie de prévention de la criminalité utilisée en Allemagne. L'objectif principal de ce système est de prédire et de prévenir les entrées par effractions, les cambriolages et autres délits liés à la propriété en utilisant l'analyse de données de crimes passés et de données socio-démographiques :

Professional burglars usually follow certain patterns, like the way they pursue and select their targets. The police use data-driven predictions – predictive policing – primarily to predict in which areas of a city houses and flats are likely to receive uninvited visitors and when. Domestic burglaries were one of the most common offenses in Germany with 117,000 registered offenses in 2017, 82 % of these remaining unsolved⁷⁷.

En identifiant des tendances et des corrélations dans les données, le système génère des prévisions de risques pour des zones spécifiques, ce qui permet aux forces de l'ordre d'allouer leurs ressources plus efficacement, de déployer des agents dans des zones à haut risque de manière proactive et de faire de la sensibilisation ciblée auprès des citoyens et commerçants des zones identifiées comme étant à risque de subir des délits.

En prédisant les points chauds potentiels de la criminalité et en fournissant ces informations aux forces de l'ordre, SKALA vise à décourager les activités criminelles, réduire le nombre d'effractions et de cambriolages, et améliorer la sécurité publique en général.

77 The new digital detective—Thanks to predictive policing, modern investigators can now selectively look into the future (2018): <https://www.t-systems.com/de/en/newsroom/best-practice/02-2018-brilliant-prospects/predictive-policing-prevent-crimes-instead-of-investigating>



Étapes de développement

En 2015, le ministère de l'Intérieur et des Affaires communales du Land de Rhénanie-du-Nord-Westphalie a chargé l'Office de la police criminelle (LKA NRW) de réaliser un projet de recherche sur la police prédictive. Les objectifs du projet étaient d'examiner les possibilités et les limites de la prévision des points chauds de la criminalité et de tester l'efficacité et l'efficience des interventions policières basées sur cette prévision⁷⁸. Pour des raisons de réduction de la complexité, le projet s'est d'abord intéressé à un seul type de délits, le vol avec effraction. Ce délit peut être influencé par des interventions policières et est variable dans l'espace et le temps. Il convenait également à l'étude en raison du nombre élevé de cas recensés et de la demande d'action policière qui en résulte. Il s'agissait en outre de vérifier si la procédure était adaptée à d'autres domaines de criminalité⁷⁹.

Lancé en février 2015, le projet pilote (alors simplement nommé *Predictive policing*) impliquait les autorités policières des districts de Duisbourg et de Cologne. Le choix s'est porté sur ces grands districts parce qu'ils étaient structurellement similaires, qu'ils avaient une charge de criminalité élevée, une structure de quartier hétérogène ainsi que des points chauds de criminalité spécifiques. En janvier 2017, le projet se voit attribuer le nom SKALA (System for Crime Analysis and Anticipation in Germany) et le projet pilote est étendu aux autorités policières des districts de Düsseldorf, Essen, Gelsenkirchen et Bonn⁸⁰.

Le système a été développé à l'interne⁸¹, en collaboration avec l'Université de Constance. Selon le rapport final sur cette initiative préliminaire, le projet pilote a donné des résultats prometteurs, le système identifiant avec succès les zones à haut risque et permettant aux forces de l'ordre de déployer des ressources de manière proactive pour décourager les activités criminelles⁸². Le projet a également fait l'objet d'une évaluation incluant un avis scientifique externe par la *Société pour la recherche sociale innovante et la planification sociale*⁸³.

Suite au projet pilote, SKALA a été intégré, en 2018, au fonctionnement quotidien des autorités policières de 16 des districts de Rhénanie-du-Nord-Westphalie. Le type de délits dont le système fait l'analyse et la prédiction a également été étendu pour inclure, outre les cambriolages résidentiels, les vols par effraction dans des locaux commerciaux et les délits liés aux véhicules.

78 *Predictive Policing in Germany* (Seidensticker, 2018, pp. 4-5): <https://kops.uni-konstanz.de/bitstreams/10477c12-a4b9-46b2-b9d0-5b58cbd127bf/download>

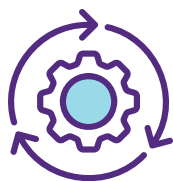
79 *Projekt SKALA: Predictive Policing in Nordrhein-Westfalen* (2018): https://www.bka.de/SharedDocs/Downloads/DE/Publikationen/ForumKI/ForumKI2018/kiforum2018BodeAbstract.pdf?__blob=publicationFile&v=6

80 *Projekt SKALA Abschlussbericht – Kurzfassung* : https://polizei.nrw/sites/default/files/2019-01/180821_Abschlussbericht_SKALA_Kurzfassung_0.pdf , *Beyond Effectiveness : Legitimising Predictive Policing in Germany* (Vepřek et al., 2020, p. 430): <https://www.kriminologie.de/index.php/krimoj/article/view/67/65>

81 *Projekt SKALA: Predictive Policing in Nordrhein-Westfalen* (2018): https://www.bka.de/SharedDocs/Downloads/DE/Publikationen/ForumKI/ForumKI2018/kiforum2018BodeAbstract.pdf?__blob=publicationFile&v=6

82 *Projekt SKALA Abschlussbericht* (Projekt SKALA Rapport final) : https://lka.polizei.nrw/sites/default/files/2018-06/180208_Abschlussbericht_SKALA.pdf

83 *Kooperative Evaluation des Projekts „SKALA“* (Évaluation conjointe du projet SKALA par The Society for Innovative Social Research and Social Planning (GISS), le Zentralstelle Evaluation (ZEVA) et Direktion Zentrale Aufgaben des autorités policières de Koln) : https://www.giss-ev.de/filestorage/publikationen/160430_evaluationsbericht_skala.pdf



Fonctionnement du système

SKALA se concentre sur la prédiction des risques de criminalité en utilisant des données spatiales pour chaque quartier résidentiel. Cette approche garantit que les prévisions de la criminalité sont faites pour l'ensemble de la ville et pas seulement pour une partie de celle-ci. SKALA utilise une approche basée sur des hypothèses qui s'appuient sur des théories criminologiques et socio-scientifiques susceptibles d'expliquer la distribution spatiale et temporelle des cambriolages domestiques (la théorie de la quasi-répétition/*near-repeat*, par exemple, qui stipule que les cambrioleurs reviennent dans les endroits où ils ont déjà eu du succès).

Les probabilités de cambriolages, de vols par effraction dans des locaux commerciaux et de délits liés aux véhicules ont été calculées sur la base de données spatiales pour des quartiers d'habitation dans des circonscriptions de police sélectionnées. SKALA analyse les données historiques sur la criminalité tirées de diverses sources, telles que les registres de police, les systèmes d'information géographique et les données démographiques. Les données fournies au système comprennent des informations sur les crimes passés (lieu, heure et type de délits), ainsi que des informations plus générales concernant l'environnement et les facteurs socio-économiques associés à différents quartiers, comme le type de construction, les caractéristiques des habitants (revenus, etc.), les indicateurs de mobilité et les connexions à l'infrastructure (À quel type de butin peut-on s'attendre dans un quartier donné ? L'accès à l'autoroute à proximité de la propriété permet-il de s'enfuir rapidement ? etc.)⁸⁴.

La méthode basée sur des hypothèses criminologiques et socio-scientifiques garantit que la modélisation et les prévisions reposent sur des théories scientifiques et des résultats de recherche solides. Cette approche se distingue de nombreuses autres méthodes de *Predictive policing*, qui se base souvent uniquement sur l'approche *near-repeat*⁸⁵.

Une fois par semaine, SKALA produit des prévisions de délits pour les autorités policières sous forme de cartes montrant les quartiers où la probabilité de cambriolage est élevée. La police utilise les prévisions pour décider, par exemple, où envoyer davantage de voitures de patrouille et s'il convient de conseiller personnellement la population des zones vulnérables sur la protection contre les cambriolages et autres vols :

Predictions by SKALA refer to a time span of seven days, Monday till Sunday, for residential and commercial burglaries and vehicle offenses. In principle, the methodology used allows shorter prediction periods. However, the following necessary steps, which include an individual rating by local crime analysts and the planning of appropriate interventions by the police, set a lower bound for the prediction period, which is currently one week⁸⁶.

84 SKALA – *Predictive Policing in North Rhine–Westphalia* (Seidensticker, 2021, pp. 52-54): <https://bulletin.cepol.europa.eu/index.php/bulletin/article/view/440>, *Report: Automating Injustice* (2021, p. 21): <https://www.fairtrials.org/articles/publications/automating-injustice/>

85 *Projekt SKALA: Predictive Policing in Nordrhein–Westfalen* (2018): https://www.bka.de/SharedDocs/Downloads/DE/Publikationen/ForumKI/ForumKI2018/kiforum2018BodeAbstract.pdf?__blob=publicationFile&v=6

86 SKALA – *Predictive Policing in North Rhine–Westphalia* (Seidensticker, 2021, p. 54): <https://bulletin.cepol.europa.eu/index.php/bulletin/article/view/440>

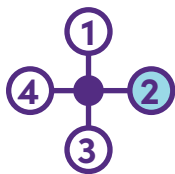
SKALA a été construit sur des composants logiciels compatibles avec les autres systèmes d'information existants du bureau de la police de l'État. Pour les modèles de prédiction, le système s'appuie sur des techniques d'exploration de données utilisant IBM SPSS Modeller⁸⁷. Les modèles de prédiction et les options de visualisation ont été développés en collaboration avec l'Université de Constance et ont été constamment adaptés au cours du projet⁸⁸. Le système utilise des algorithmes d'apprentissage automatique (*decision trees*, *spatio-temporal cluster analysis* (STCA), *random forests* et *regression models*) pour analyser les caractéristiques sélectionnées, construire un modèle prédictif et identifier les schémas et les corrélations qui indiquent une plus grande probabilité de criminalité⁸⁹. Ces prévisions sont présentées sous forme de cartes thermiques et autres représentations visuelles qui aident les forces de l'ordre à comprendre le risque de criminalité pour des zones et des périodes spécifiques.



Bénéficiaire direct du système

Les organisations policières sont souvent confrontées à des ressources limitées et à des contraintes budgétaires, d'où l'importance cruciale d'une allocation efficace de leurs ressources. SKALA bénéficie avant tout aux organisations policières en leur permettant de mieux cibler leurs efforts en identifiant les zones où les activités criminelles sont plus susceptibles de se produire. Cette meilleure allocation des ressources est propice à une plus grande efficacité organisationnelle. L'amélioration de l'efficacité contribue à rendre la communauté plus sûre, un objectif central des autorités policières.

Les employés des services de police bénéficient indirectement de SKALA en permettant de travailler de manière plus efficace suite au ciblage de zones propices à des activités criminelles. Les citoyens sont également des bénéficiaires indirects du système. En identifiant de manière proactive les zones à haut risque et en facilitant l'allocation des ressources de manière efficace, SKALA aide les forces de l'ordre à dissuader les activités criminelles, ce qui conduit à un environnement de vie plus sûr pour le grand public.



Degré d'autonomie du système

SKALA sert principalement d'outil d'aide à la décision aux organisations policières, en fournissant des analyses prédictives et des prévisions de risques. Le système n'a pas accès en temps réel à l'environnement concerné par sa tâche et la qualité de ses prévisions repose sur les données qui lui sont fournies. Le système ne prend pas de décisions ou d'actions de manière autonome, mais aide les forces de l'ordre à allouer des ressources et à déployer du personnel de manière plus efficace sur la base de ses prédictions. SKALA correspond au degré d'autonomie 2 (autonomie limitée).

87 *Beyond Effectiveness : Legitimising Predictive Policing in Germany* (Vepřek et al., 2020, p. 430): <https://www.kriminologie.de/index.php/krimoj/article/view/67/65>

88 *Predictive Policing in Germany* (Seidensticker, 2018, p. 5): https://www.researchgate.net/publication/332170526_Predictive_Policing_in_Germany

89 SKALA – *Predictive Policing in North Rhine-Westphalia* (Seidensticker, 2021, p. 54): <https://bulletin.cepol.europa.eu/index.php/bulletin/article/view/440>



Pistes de réflexions éthiques

Bénéfices et avantages du SIA

- Améliorer la sécurité. L'argument principal en faveur du déploiement de systèmes comme SKALA est l'amélioration de la sécurité. Avec une meilleure allocation des ressources policières, les zones à risque peuvent être mieux surveillées, ce qui pourrait potentiellement augmenter la sécurité pour l'ensemble de la population.
- Permettre une meilleure allocation des ressources policières. Le système permet l'identification des zones à risque élevé de délit, permettant ainsi une allocation plus efficace des ressources policières. En augmentant la présence policière dans ces zones, on peut potentiellement décourager les activités criminelles.
- Informer la population. Suivant les prédictions du système, les autorités peuvent informer les citoyens et les commerçants des zones à risque afin qu'ils puissent prendre les précautions nécessaires.

Inconvénients et risques du SIA

- Création d'un climat de méfiance. Les techniques de police prédictive peuvent entraîner des interventions policières excessives (Over-policing) dans certaines zones. Les policiers pourraient être amenés à intervenir de manière plus prompte et moins rationnelle dans les zones ciblées par le système, favorisant un climat de méfiance envers les forces de police auprès des populations qui se sentent surveillées. Les forces policières pourraient également être tentées d'utiliser le système pour d'autres types de crimes sans s'assurer de la robustesse et de la fiabilité des prédictions pour ces nouvelles infractions.
- Fiabilité et validité des prédictions. Ce type de système nécessite une énorme quantité de données pour l'entraînement des algorithmes d'apprentissage automatique, et il n'est pas garanti que le niveau de criminalité des zones touchées permette de générer suffisamment de données pour rendre possible une prédiction fiable de la criminalité.
- Iniquité dans l'allocation des ressources policières. En se concentrant sur un certain type de méfaits, le système peut limiter les ressources disponibles pour contrer d'autres problèmes potentiellement plus graves ou d'autres types de méfaits. De plus, le système pourrait simplement déplacer les concentrations de criminalité en intervenant dans les zones chaudes, déplaçant la criminalité vers des zones plus froides. Ceci renvoie à la question de savoir si le système est réellement efficace pour diminuer ou prévenir la criminalité dans son ensemble.
- Menace aux droits de la personne. Le système soulève des enjeux importants en matière de droits de la personne, comme le risque d'arrestation ou de détention arbitraire, de protection de la vie privée, de liberté d'expression et d'association, et le droit à l'égalité et à la non-discrimination.
- Orientation vers le contrôle plutôt que la prévention. Le système est utilisé principalement pour la prédiction et le contrôle des activités criminelles, sans s'intéresser suffisamment à l'intervention sociale et à la prévention. De plus, le système pourrait inciter à une augmentation des dispositifs de surveillance dans de nombreux quartiers.
- Rétroaction de la criminalisation et engagement dans un cercle vicieux. L'utilisation de données passées pour concentrer les effectifs policiers dans certaines zones peut augmenter la criminalisation dans ces zones, ce qui sera reflété dans les données utilisées pour les prédictions suivantes.
- Risque de profilage et de stigmatisation. Il y a un risque que des biais présents dans les données d'entraînement soient reproduits ou accentués par le système, entraînant un profilage et une stigmatisation. De plus, les modèles théoriques et les données sociodémographiques sur lesquels le système se base peuvent comporter leurs propres biais et problèmes. Il existe également une inquiétude quant à la manière dont les citoyens réagiront aux informations reçues concernant les prédictions de risque dans leur quartier. Cela pourrait encourager la stigmatisation et le profilage par les citoyens eux-mêmes.

Valeurs à prendre en considération pour le déploiement et l'utilisation de ce SIA (ou d'un système similaire)

- Confiance envers les forces de police. Une attention doit être portée à l'impact social potentiel de la sur-surveillance, y compris l'exacerbation des tensions entre la police et la population, la perte de confiance envers les forces de police, et l'impact sur les libertés d'association, de circulation et d'expression.
- Droits de la personne. S'assurer que le SIA respecte les droits de la personne, y compris le droit à l'égalité, à la non-discrimination, à la protection contre l'arrestation et la détention arbitraires, et le respect de la vie privée. De plus, les analyses et prédictions du système pourraient inciter à une augmentation des dispositifs de surveillance dans certains quartiers. Il est important de veiller à la protection de la vie privée des citoyens et notamment des populations plus vulnérables.
- Équité et non-discrimination. Il est important que le système ne contribue pas au profilage racial, à la discrimination ou à la stigmatisation des groupes vulnérables.
- Explicabilité. Puisque le système peut avoir un impact potentiellement élevé sur la vie des individus, il est essentiel de comprendre comment il fait ses prédictions.
- Justice sociale et prévention. Le système pourrait être utilisé pour des interventions plus constructives, moins axées sur la criminalisation et davantage sur l'intervention sociale et la prévention. Cela pourrait être une façon de s'assurer que le système soutient les valeurs de justice sociale.
- Prudence. La tentation d'appliquer le système à d'autres types de crimes doit être évitée sans une évaluation préalable de la robustesse et de la fiabilité des prédictions pour ces nouvelles applications.
- Sécurité et proportionnalité. SKALA peut améliorer la sécurité et accroître l'efficacité de l'allocation des ressources policières. Cependant, cela doit être fait de manière à ne pas mener à une sur-surveillance et à des interventions policières excessives.

Comment protéger ou restaurer les valeurs affectées par le SIA?

- Analyse d'impact. Il est nécessaire d'effectuer des analyses d'impact qui garantissent le respect des valeurs éthiques lors du déploiement et de l'utilisation de ces SIA. C'est particulièrement important pour un système comme SKALA qui a un niveau d'impact élevé puisqu'il est directement lié au système criminel et peut considérablement affecter la vie des individus. Compte tenu des enjeux importants liés à ce système, l'analyse d'impact devrait mettre l'accent sur les impacts sur les droits de la personne. Les risques d'arrestations et de détentions arbitraires, de violations de la vie privée, de restrictions de la liberté d'expression et d'association, et de violations du droit à l'égalité et à la non-discrimination sont tous des préoccupations potentielles qui doivent être considérées.
- Éviter la généralisation hâtive du système à d'autres usages. Il est souhaité que l'utilisation du système ne soit pas généralisée à d'autres crimes pour lesquels la quantité et la qualité des données pourraient être insuffisantes, ce qui rendrait cette extension très problématique.

Autres ressources

Seidensticker, K., & Schwarz, K. (2022). *Using Forecasting Methods on Crime Data: The SKALA Approach of the State Office for Criminal Investigation of North Rhine-Westphalia*. *ITISE 2022*, 39. <https://doi.org/10.3390/engproc2022018039>

Egbert, S., & Leese, M. (2020). *Criminal futures: predictive policing and everyday police work* (Ser. Routledge studies in policing and society). Routledge. <https://doi.org/10.4324/9780429328732>, pp. 178-182

Sommerer, L. M. (2017). *Geospatial Predictive Policing – Research Outlook & A Call For Legal Debate*. *Neue Kriminalpolitik*, 29(2), 147-164. <https://doi.org/10.5771/0934-9200-2017-2-147>

Algorithm Watch—*Automating society report* pour l'Allemagne : <https://automatingsociety.org/algorithmwatch/report2020/germany/>

Surtrac

Tableau 8. Description synthétique de *Surtrac*

Chronologie	2012 : programme pilote 2013 : expansion dans différents quartiers de la ville 2019 : <i>Surtrac</i> 2.0 (Modernisation du déploiement afin d'intégrer des extensions adaptées aux piétons et des avancées en matière de surveillance à distance) ⁹⁰
Pays	États-Unis
Organisation	Ville de Pittsburgh (<i>Department of Mobility and Infrastructure</i> - DOMI)
Développeur	Partenariat public-privé (<i>Department of Mobility and Infrastructure</i> - DOMI) de la ville de Pittsburgh et Carnegie Mellon University).
Bénéficiaire direct	Usagers (de la route)
Degré d'autonomie	3. Semi-autonome



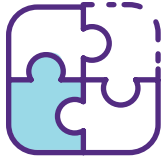
Objectif poursuivi

L'objectif du système *Surtrac* (*Scalable Urban Traffic Control*), utilisé à Pittsburgh, est d'optimiser la circulation dans les environnements urbains en gérant intelligemment les feux de circulation en temps réel. Ce système adaptatif de contrôle des feux de circulation a été mis au point en collaboration avec *Traffic21*, un groupe de recherche sur les transports de l'université Carnegie Mellon⁹¹. *Surtrac* vise à réduire les embouteillages, la durée des trajets et les émissions des véhicules, tout en augmentant l'efficacité globale du transport.

Le système utilise une combinaison d'intelligence artificielle, de théorie de la circulation et de données de circulation collectées en temps réel provenant de capteurs et de caméras pour ajuster les feux de circulation en fonction des conditions de circulation. En s'adaptant dynamiquement à l'évolution du trafic, l'objectif de *Surtrac* est de fournir aux usagers des voies publiques et à la population dans son ensemble un réseau de transport plus efficace et plus respectueux de l'environnement.

90 Metro21/Smart Cities Institute : https://www.cmu.edu/metro21/projects/safe-mobility/surtrac_2.0.html ; *Surtrac for the People: Upgrading the Surtrac Pittsburgh Deployment to Incorporate Pedestrian Friendly Extensions and Remote Monitoring Advances*, s. d.: <https://trid.trb.org/view/1681289>

91 *Delivering Artificial Intelligence in Government: Challenges and Opportunities* (Desouza, s. d., p. 12): <https://www.businessofgovernment.org/report/delivering-artificial-intelligence-government-challenges-and-opportunities>; *Surtrac: Scalable Urban Traffic Control* (Smith et al., 2013): https://www.ri.cmu.edu/pub_files/2013/1/13-0315.pdf; *The future of open city streets could start with smarter traffic lights* (Simek, 2022): https://www.cmu.edu/ambassadors/july-2022/pdf/ps_open-streets.pdf



Étapes de développement

Le système *Surtrac* a été développé à l'Institut de robotique de l'université Carnegie Mellon. La mise en œuvre de *Surtrac* à Pittsburgh a commencé par un programme pilote dans le quartier d'East Liberty, en juin 2012. Le déploiement initial a couvert neuf intersections et a démontré des améliorations significatives sur la durée de parcours, la réduction des embouteillages et la diminution des émissions des véhicules. En novembre 2013, le nombre d'intersections couvertes a été doublé.

Au fil des ans, le système a été étendu pour couvrir davantage d'intersections et de quartiers. Le succès de *Surtrac* à Pittsburgh suscite l'intérêt d'autres villes et régions, qui envisagent de mettre en œuvre des systèmes de contrôle du trafic adaptatif similaires pour résoudre leurs propres problèmes de circulation⁹². En 2019-20, *Surtrac 2.0* voit la modernisation du déploiement afin d'intégrer des extensions adaptées aux piétons et des avancées en matière de surveillance à distance⁹³.

Pour commercialiser *Surtrac* hors de Pittsburgh, les concepteurs fondent la société *Rapid Flow* en 2020. Cette société sera acquise en 2022 par *Miovision*⁹⁴, basé à Kitchener en Ontario (Canada)⁹⁵. *Miovision Surtrac* assure le maintien des systèmes déjà en place et poursuit l'élan de commercialisation commencé par *Rapid Flow* :

Surtrac has competition for its share of the grid from tech giants and startups alike. Fortunately for the 22 cities already using it – including Atlanta and several New England towns – it's relatively cheap, it's easy to install and maintain, and it taps existing infrastructure like streetlight cameras. The computer at each intersection is its own node that crunches a range of scheduling options and then adjusts signal timing every second while passing insights to neighboring lights. This decentralized approach makes Surtrac ideal for regulating unpredictable urban traffic and for gradual expansion⁹⁶.

En février 2023, Miovision fait l'acquisition de MicroTraffic⁹⁷, un fournisseur de systèmes d'analyse vidéo pour la sécurité routière, ainsi que *Traffop*⁹⁸. Ces acquisitions donnent une idée des orientations et des nouvelles capacités des systèmes modulaires pouvant se combiner à *Surtrac*, faisant usage des caméras et capteurs déjà en place :

What attracted us to MicroTraffic is their ability to accurately predict the likelihood of traffic crashes causing injuries or fatalities on roads and intersections. Like Miovision, this safety solution uses computer vision, video analytics, machine learning, and predictive modeling to analyze video traffic data. In fact, about 80% of the traffic video data they analyze comes from their customers that deploy Miovision Scout for traffic studies⁹⁹.

92 *Going Nowhere Fast? Smart Traffic Lights Can Help Ease Gridlock* | HowStuffWorks (2022): <https://science.howstuffworks.com/engineering/civil/smart-traffic-lights-news.htm>

93 *Surtrac for the People: Upgrading the Surtrac Pittsburgh Deployment to Incorporate Pedestrian Friendly Extensions and Remote Monitoring Advances*, s. d.: <https://trid.trb.org/view/1681289>; Metro21/Smart Cities Institute : https://www.cmu.edu/metro21/projects/safe-mobility/surtrac_2.0.html

94 *Smart Infrastructure for Future Urban Mobility*: <https://miovision.com/company>

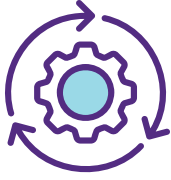
95 *RapidFlow Surtrac is now Miovision Surtrac*: <https://miovision.com/rapidflow>

96 *The future of open city streets could start with smarter traffic lights* (Simek, 2022): https://www.cmu.edu/ambassadors/july-2022/pdf/ps_open-streets.pdf

97 *Miovision welcomes MicroTraffic to help make streets safer*, 2023: <https://miovision.com/blog/miovision-welcomes-microtrafic-to-help-make-streets-safer>

98 *How to Gain Traffic Insights with Traffop ATSPMs*: <https://miovision.com/blog/gain-traffic-insights-with-traffop-atspm>

99 *Miovision welcomes MicroTraffic to help make streets safer*, 2023: <https://miovision.com/blog/miovision-welcomes-microtrafic-to-help-make-streets-safer>



Fonctionnement du système

Le système *Surtrac* utilise l'IA pour gérer dynamiquement les feux de circulation en temps réel en fonction des conditions de circulation. Les algorithmes d'intelligence artificielle utilisés par *Surtrac* se concentrent sur l'optimisation et la coordination, ce qui permet au système de s'adapter à l'évolution du trafic et de prendre des décisions intelligentes concernant la synchronisation des feux à chaque intersection¹⁰⁰.

Voici un aperçu du fonctionnement du système :

- Collecte de données. *Surtrac* utilise divers capteurs, tels que des caméras et des détecteurs à boucle inductive, pour recueillir des données sur le trafic en temps réel. Ces données comprennent des informations sur le nombre de véhicules, les files d'attente et le temps d'attente aux intersections.
- Optimisation locale. Chaque intersection est équipée de son propre agent d'intelligence artificielle, qui traite les données de trafic recueillies par ses capteurs. Ces informations sont utilisées pour déterminer les horaires de signalisation optimaux pour ce carrefour spécifique, en tenant compte de facteurs tels que le flux de circulation, les mouvements de piétons et la priorité accordée au transport en commun.
- Coordination. Les agents d'intelligence artificielle de chaque intersection communiquent avec leurs voisins pour assurer la fluidité du trafic sur l'ensemble du réseau. En partageant des informations sur les conditions de circulation actuelles et les horaires de signalisation prévus, les agents d'intelligence artificielle peuvent coordonner leurs actions, ce qui réduit les risques d'embouteillage et améliore l'efficacité globale du trafic.
- Adaptation continue. Le système met continuellement à jour les horaires des feux de signalisation en fonction des nouvelles données fournies par les capteurs. Cela permet au système de s'adapter à l'évolution du trafic tout au long de la journée, en veillant à ce que les feux de signalisation soient toujours optimisés en fonction des conditions réelles (et non aux conditions historiques pour une date ou une heure donnée).

Les algorithmes d'IA utilisés sont principalement basés sur des techniques d'apprentissage automatique, d'optimisation et de systèmes multi-agents. L'une des principales caractéristiques de *Surtrac* est sa nature décentralisée, où chaque intersection fonctionne de manière indépendante, mais communique avec ses voisins pour une meilleure coordination (Edge AI). Cette approche permet au système de s'adapter facilement et de maintenir son efficacité même lorsque le réseau s'agrandit.

L'analyse vidéo, l'apprentissage automatique et la modélisation prédictive pour analyser les données du trafic vidéo sont au cœur des techniques utilisées suite à l'acquisition de *MicroTraffic* par *Miovision Surtrac* :

It's the analytics that sets them apart. Based on their kinetic energy approach, which takes into account speeds, temporal separation, conflict angle and the vulnerability of the user, *MicroTraffic* can quantify how likely near misses – situations where two or more road users come in close proximity – might lead to collisions resulting in injuries or fatalities. According to a study by researchers at Toronto Metropolitan University, their approach can predict, based on near misses, the future number of fatal and injury crashes with 94% accuracy. Based on this information, road safety engineers can recommend safety plans for intersections that result in up to an 80% reduction in risk¹⁰¹.

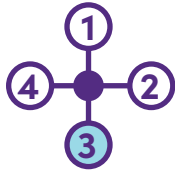
100 *Innovative Applications of AI: The Surtrac Application* | Interactive AI Magazine, s. d.: <https://interactiveaimag.org/ai-magazine-previews/ai-mag-previews-vol41-no1/>, *Smart Infrastructure for Future Urban Mobility* | US Department of Transportation, 2017. : <https://www.transportation.gov/utc/smart-infrastructure-future-urban-mobility>

101 *Incorporating Speed in a Traffic Conflict Severity Index to Estimate Left Turn Opposed Crashes at Signalized Intersections* (Anarkooli et al., 2021): <https://miovision.com/blog/miovision-welcomes-microtraffic-to-help-make-streets-safer>



Bénéficiaire direct du système

Le système *Surtrac* est mis en place et opéré par le *Department of Mobility and Infrastructure* (DOMI) de la ville de Pittsburgh. L'objectif du système (le travail qu'il accomplit) est de réduire les embouteillages, la durée des trajets et les émissions des véhicules, tout en augmentant l'efficacité globale des transports en ajustant les feux de circulation en temps réel en fonction des conditions de circulation. En ce sens, les usagers de la voie publique sont les bénéficiaires directs du système qui contribue à l'amélioration de la fluidité de la circulation et de leur vitesse de déplacement. L'organisation bénéficie indirectement du système, ainsi que d'autres acteurs tels les planificateurs de la circulation et les policiers.



Degré d'autonomie du système

Les capacités de détection du système et la synchronisation automatique des feux de circulation minimisent le besoin d'intervention humaine, automatisant le flot de la circulation en temps réel. Le système 'agit' de façon autonome à partir des données recueillies dans l'environnement concerné par sa tâche. Les opérateurs du système peuvent s'inspirer des 'actions' du système et utiliser les résultats des analyses prédictives des incidents de la route pour apporter des modifications à la planification urbaine. Le système *Surtrac* fonctionne avec un degré d'autonomie 3 (semi-autonome)¹⁰².



Pistes de réflexions éthiques

Bénéfices et avantages du SIA

- Fluidifie le trafic. L'optimisation de la synchronisation des feux de circulation est un objectif louable, améliorant la fluidité du trafic au bénéfice des usagers de la voie publique.
- Contribue à réduire les émissions de CO₂. La synchronisation optimisée des feux de circulation réduit le temps d'attente des véhicules aux intersections, contribuant à diminuer les émissions de CO₂.
- Maintient les règles du code de la signalisation aux intersections. Le système optimise la synchronisation des feux en maintenant les codes habituels de la signalisation en termes de progression des signaux lumineux, de périodes pour chaque couleur de lumière et de signalisation piétonne. Le déploiement de ce système ne génère pas, de prime abord, de risques supplémentaires en comparaison avec les systèmes actuels basés sur une programmation préétablie des feux selon le moment de la journée sur une base historique.
- Évaluer les risques d'accident et contribuer à la reconfiguration des intersections. Grâce à l'analyse des données collectées par les capteurs et les caméras, une fonctionnalité nouvellement développée de ce système permet de quantifier les risques d'accident aux différentes intersections. Les urbanistes et les responsables des infrastructures peuvent s'appuyer sur ces données pour envisager des modifications permettant d'améliorer la sécurité des usagers.

¹⁰² L'autonomie du système est ici liée à sa capacité d'intervenir sans supervision à partir de l'information reçue en temps réel par ses différents capteurs. Le système repose sur les réglementations relatives à la circulation, ainsi que sur différentes théories concernant la fluidité du trafic routier. Puisque la synchronisation des feux de circulation se fait de façon autonome, les responsables de la planification urbaine peuvent apprendre de leurs observations de ce que le système fait, et introduire ces éléments d'optimisation dans la théorie qui sera, rétroactivement, appliquée au système. Pour ces raisons, nous qualifions ce système de semi-autonome.

Inconvénients et risques du SIA

- Imputabilité et responsabilité. Qui serait responsable d'un accident résultant d'une inefficacité du système ou d'un dysfonctionnement ? La ville ou le développeur du système ? La réponse à cette question n'est pas évidente. D'autant plus que ce système fonctionne de façon autonome et 'décide' de la synchronisation des feux sans la présence d'un *humain dans la boucle*. De quelle manière un conducteur (ou tout autre usager de la voie publique) peut-il se plaindre du fonctionnement du système ou peut-il signaler un problème avec celui-ci?
- Arbitrage entre l'efficacité et la sécurité. La sécurité ne doit pas être compromise au profit d'une amélioration de la fluidité de la circulation. Est-il approprié de confier la gestion des feux de circulation en temps réel, et chaque décision que cela implique, à un système algorithmique? On sait que les systèmes d'IA utilisés dans les voitures autonomes ne sont pas tout à fait au point, en particulier pour la reconnaissance d'objets et pour l'anticipation des actions d'individus ou d'animaux présents sur la voie publique. Est-ce que le système est suffisamment efficace pour détecter tous les usagers de la voie publique, en particulier ceux qui ne se déplacent pas dans des véhicules (piétons, cyclistes, trottinettistes, personnes en chaises roulantes, etc.) ?
- Fiabilité et capacité du système de prendre en compte la diversité et les besoins spécifiques des usagers de la route. Si le système ne *reconnait* pas un individu comme étant un piéton, le signal pour piétons pourrait ne jamais s'activer, ce qui pose un risque potentiel pour la sécurité.
- Risque de discrimination ou d'iniquité si le système est optimisé pour la majorité des véhicules se déplaçant dans une direction particulière. Les personnes se déplaçant dans la direction opposée ou sur des voies secondaires pourraient subir des temps d'attente excessifs, ce qui peut créer des situations inéquitables. Cette perception d'une injustice causée par la machine (il n'y a pas d'agent de police avec lequel interagir pour communiquer son mécontentement) pourrait être source d'impatience et de témérité chez certains conducteurs (ou autres usagers de la voie publique), et causer des accidents.
- Complexité et risques associés au système disproportionnés par rapport aux bénéfices (par exemple, une réduction de quelques secondes du temps d'attente à un feu de circulation).
- Collecte de nombreuses informations dont certaines semblent superflues à l'objectif poursuivi. Ces données pourraient potentiellement être utilisées plus tard et servir à d'autres fins (identification des comportements et habitudes des individus, par exemple), posant des problèmes de confidentialité et de protection de la vie privée. L'utilisation des données recueillies par le système dans le cadre de son fonctionnement initial serait-elle appropriée pour un nouvel usage (dans le contexte d'une acquisition par une autre entreprise ou d'une nouvelle itération du système par exemple)?

Valeurs à prendre en considération pour le déploiement et l'utilisation de ce SIA (ou d'un système similaire)

- Confidentialité. La gestion des données collectées par le système doit respecter les droits à la vie privée des individus.
- Efficacité. Le système doit être efficace dans la synchronisation des feux de circulation et la détection de tous les usagers de la route. La sécurité des utilisateurs de la route doit être une priorité du déploiement de ce système, au-delà de la fluidité à laquelle il contribue.
- Équité. Le système doit être conçu de manière à assurer une circulation équitable pour tous les usagers de la voie publique, indépendamment de la direction ou du type de voie empruntée.
- Imputabilité et responsabilité. Il est important de clarifier qui serait responsable en cas de dysfonctionnement du système.
- Transparence. Les utilisateurs de la route doivent avoir accès à l'information concernant le fonctionnement du système et des recours possibles, notamment en cas de dysfonctionnement ou de temps d'attente excessif à une intersection.

Comment protéger ou restaurer les valeurs affectées par le SIA?

- Confidentialité. Il est important de clarifier la politique de confidentialité et de gestion des données/informations personnelles et collectives saisies par le système, y compris la durée et les méthodes de stockage.
- Efficacité. Le système pourrait être soumis à des tests et des évaluations pour s'assurer de sa performance en termes de détection de tous les usagers de la voie publique et de synchronisation juste et appropriée des feux de circulation.
- Équité. Des algorithmes plus sophistiqués pourraient être nécessaires pour assurer une gestion équitable du trafic, quel que soit le sens ou le type de voie emprunté par les usagers. Par exemple, une attention particulière devrait être portée aux piétons ou utilisateurs vulnérables de la voie publique, peu importe leur vitesse de déplacement et leur moyen de locomotion. Une synchronisation juste du signal piétonnier ne doit pas être délaissée au profit de l'optimisation de la circulation des véhicules.
- Imputabilité et responsabilité. Une clarification juridique et contractuelle pourrait être nécessaire pour déterminer qui serait responsable en cas de dysfonctionnement du système.
- Transparence. Rendre disponible aux utilisateurs de la voie publique des informations sur le fonctionnement du système, des moyens de signaler les dysfonctionnements et de communiquer leurs préoccupations.

Corti

Tableau 9. Description synthétique de Corti

Chronologie	Introduction : 2020
Pays	Suède
Organisation	SOS Alarm (entreprise à propriété partagée entre le gouvernement suédois et l'Association suédoise des autorités locales et régionales)
Développeur	Partenariat public-privé (SOS Alarm Sweden et Corti AI)
Bénéficiaire direct	Employés
Degré d'autonomie	3. Semi-autonome

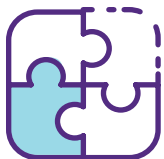


Objectif poursuivi

Corti est un système d'intelligence artificielle utilisée dans les services médicaux d'urgence. En Suède, il a été incorporé à SOS Alarm, l'organisation nationale chargée de traiter les appels d'urgence¹⁰³. L'objectif principal du système *Corti AI* est d'aider les opérateurs des services d'urgence à identifier avec précision la nature et la gravité de l'état de santé de l'appelant, ce qui leur permet de prendre rapidement des décisions plus éclairées.

Le système analyse les informations fournies par l'appelant et utilise sa vaste base de données de connaissances médicales pour identifier les problèmes de santé potentiels. Il propose des recommandations en temps réel aux opérateurs d'appel, ce qui garantit une réponse plus rapide et plus efficace. En améliorant la précision du triage, *Corti* contribue à réduire les temps de réponse et à améliorer les résultats pour les patients. En outre, il contribue à la formation des opérateurs en leur offrant un retour d'information continu et en tirant des enseignements de leurs interactions, améliorant ainsi la gestion globale des appels d'urgence.

¹⁰³ SOS Alarm Sverige AB is a Swedish publicly owned company that operates emergency number 112 in Sweden. The enterprise is owned equally by the Government of Sweden and the Swedish Association of Local Authorities and Regions. Source: Wikipedia.



Étapes de développement

Le système *Corti* a été développé par une société basée à Copenhague, *Corti Labs*, fondée en 2016. Le développement de la technologie a commencé par l'amélioration de la réponse médicale d'urgence en aidant les répartiteurs à prendre des décisions plus rapides et plus précises.

Avant de mettre en œuvre le système à grande échelle, *Corti Labs* a mené, à partir de 2017, différents projets pilotes en partenariat avec les services médicaux d'urgence de Copenhague au Danemark¹⁰⁴, le SAMU-SDIS 74 du département de Haute-Savoie en France¹⁰⁵ et le AREU Lombardie en Italie¹⁰⁶. L'étude visait à tester l'efficacité du système d'IA *Corti* dans l'identification des arrêts cardiaques extrahospitaliers (OHCA) lors des appels d'urgence¹⁰⁷. Les résultats ont démontré que *Corti* pouvait détecter les arrêts cardiaques avec une plus grande précision que les répartiteurs humains.

Au cours de l'étude danoise, *Corti* a été utilisée pour analyser plus de 100 000 appels d'urgence passés aux services médicaux d'urgence de Copenhague. Le SIA a 'écouté' les appels, traité les données et fourni des informations en temps réel pour aider les répartiteurs à prendre des décisions. L'étude s'est concentrée sur la capacité de l'IA à identifier les cas d'arrêt cardiaque, puisque leur détection et intervention précoce sont essentielles pour améliorer les chances de survie.

Les résultats de l'étude pilote étaient prometteurs. *Corti* a été capable de reconnaître les cas d'OHCA avec une précision de 93 %, contre 73 % pour les répartiteurs humains. L'étude a également montré que le SIA permettait un diagnostic plus rapide et plus précis. L'étude a non seulement mis en évidence la capacité de l'IA à contribuer à la détection de conditions médicales potentiellement mortelles, mais a également ouvert la voie à l'exploration de son utilisation par *Corti AI* dans d'autres aspects des services médicaux : détection d'autres problématiques cardiaques, assistance au diagnostic dans d'autres contextes de consultations à caractère médicaux (lignes Infosanté, etc.)¹⁰⁸, utilisation en contexte de pandémie de COVID-19¹⁰⁹, sans oublier le fort potentiel d'usage par des entreprises privées dans le domaine des assurances par exemple¹¹⁰.

104 Blomberg, S. N., Folke, F., Ersbøll, A. K., Christensen, H. C., Torp-Pedersen, C., Sayre, M. R., Counts, C. R., & Lippert, F. K. (2019). *Machine learning as a supportive tool to recognize cardiac arrest in emergency calls*. Resuscitation, 138, 322-329. <https://doi.org/10.1016/j.resuscitation.2019.01.015>

105 *What is Artificial Intelligence Doing on your Emergency Call in France?* | by Corti News | Medium, 2020: <https://medium.com/corti-ai/what-is-artificial-intelligence-doing-on-your-emergency-call-in-france-603f295e4672>

106 *AI that detects cardiac arrests during emergency calls will be tested across Europe this summer* - The Verge, 2018: <https://www.theverge.com/2018/4/25/17278994/ai-cardiac-arrest-corti-emergency-call-response>; *Artificial Intelligence & Emergency services: saving lives on the phone* - EENA, 2020: <https://eena.org/knowledge-hub/documents/detecting-out-of-hospital-cardiac-arrest-using-artificial-intelligence/>; EENA2019 - Artificial Intelligence - YouTube, 2019: <https://www.youtube.com/watch?v=V3LCe2tGCac>

107 *'Corti-s AI - Emergency medical services real-time speech analysis for cardiac predictions*: <https://ai-watch.github.io/AI-watch-T6-X/service/90016.html>

108 *AI for Ambulatory Clinics*: <https://www.corti.ai/use-cases/ambulatory-clinics>; Scholz, M. L., Collatz-Christensen, H., Blomberg, S. N. F., Boebel, S., Verhoeven, J., & Krafft, T. (2022). *Artificial intelligence in Emergency Medical Services dispatching: Assessing the potential impact of an automatic speech recognition software on stroke detection taking the Capital Region of Denmark as case in point*. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine, 30(1), 36 : <https://doi.org/10.1186/s13049-022-01020-6>

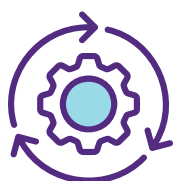
109 *Corti Solution Launches in Sweden to Support SOS Alarm Fight the Pandemic* | by Corti | Medium, 2020: <https://medium.com/corti-ai/corti-solution-launches-in-sweden-to-support-sos-alarm-fight-covid-19-f5fbd6d4c370>

110 *Corti · Health Insurance* : <https://www.corti.ai/use-cases/health-insurance>

4. Cas d'application de l'utilisation de l'IA dans le secteur public

C'est à la suite de ces résultats positifs que le système a été adopté et mis en place dans les centres d'appels d'urgence suédois opérés par SOS Alarm¹¹¹, entreprise à propriété partagée entre le gouvernement suédois et l'Association suédoise des autorités locales et régionales. Le processus de mise en œuvre a consisté à intégrer le système dans l'infrastructure existante des centres d'appels d'urgence et à former les répartiteurs à l'utilisation des informations fournies par l'IA. Au fur et à mesure de la mise en œuvre du système d'IA de Corti et de son utilisation dans des scénarios réels, les commentaires des répartiteurs et des autres parties prenantes ont été recueillis afin d'améliorer les performances du système. Ce processus itératif a permis à Corti Labs d'affiner son modèle d'IA, le rendant plus précis et plus efficace, et l'adaptant aux différents contextes des organisations desservies.

Depuis sa mise en œuvre, Corti a attiré l'attention sur son potentiel d'amélioration des interventions médicales d'urgence dans le monde entier¹¹². L'entreprise a conclu des partenariats avec plusieurs organisations, a étendu ses services aux États-Unis et ailleurs en Europe¹¹³.



Fonctionnement du système

Le système utilise une combinaison de techniques de traitement du langage naturel, de reconnaissance vocale et d'apprentissage automatique pour aider les répartiteurs médicaux lors des appels d'urgence. Voici un aperçu de son fonctionnement¹¹⁴:

- Le SIA commence par convertir les appels d'urgence audio en texte. Ce processus implique l'utilisation de la technologie de reconnaissance automatique de la parole (ASR), qui transcrit les mots prononcés en texte écrit, facilitant l'analyse du contenu de la conversation par l'IA.
- Une fois l'audio transcrit en texte, le système utilise des techniques de traitement du langage naturel pour comprendre et interpréter le contenu. Le *Natural language processing* (NLP) permet au système d'identifier les informations critiques telles que les symptômes, les antécédents déclarés par le patient au cours de l'appel et les indices contextuels issus de l'analyse du langage utilisé.
- Corti est entraîné sur un vaste ensemble de données (composé essentiellement d'un historique d'appels d'urgence reçus par une organisation), ce qui lui permet d'apprendre à reconnaître les schémas associés à différentes conditions médicales. Le système utilise des algorithmes d'apprentissage automatique pour identifier ces schémas en temps réel au fur et à mesure que de nouveaux appels sont reçus, ce qui augmente la précision de ses prédictions et de ses suggestions.

111 Sweden Turns to Artificial Intelligence During Medical Emergencies | by Corti News | Corti | Medium, 2020: <https://medium.com/corti-ai/sweden-turns-to-artificial-intelligence-during-medical-emergencies-67ef94af53a5> ; Health Tech Hub Copenhagen | Corti: <https://healthtechhub.org/members/corti/> ; SOS Alarm: *Increasing Quality and Safety at Sweden's National Medical Call Center*: <https://www.corti.ai/case-studies/sos-alarm-improved-triaging>

112 Corti.ai Raises \$27 Million in Series A Funding to Transform Patient Consultations With Artificial Intelligence | Business Wire, 2021: <https://www.businesswire.com/news/home/20210909005416/en/Corti.ai-Raises-27-Million-in-Series-A-Funding-to-Transform-Patient-Consultations-With-Artificial-Intelligence>

113 Priority Dispatch Announces Automatic Quality Assurance of Emergency Calls Through a Partnership With The Artificial Intelligence Company Corti.ai - Corti, 2023: <https://www.corti.ai/stories/priority-dispatch-announces-automatic-qualityassurance-of-emergency-calls-through-a-partnershipwith-the-artificial-intelligence-company-corti-ai>

114 Corti - AI for Patient Consultations : <https://www.corti.ai/> ; AI That Saves Lives: The Chatbot That Can Detect A Heart Attack Using Machine Learning | Bernard Marr: <https://bernardmarr.com/ai-that-saves-lives-the-chatbot-that-can-detect-a-heart-attack-using-machine-learning/> ; Having A Heart Attack? This AI Helps Emergency Dispatchers Find Out, 2018: <https://www.fastcompany.com/40515740/having-a-heart-attack-this-ai-helps-emergency-dispatchers-find-out>

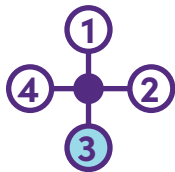
- Lorsque le système d'IA traite et analyse l'appel d'urgence, il fournit, en temps réel, des informations et des recommandations aux répartiteurs. Ces suggestions peuvent inclure des diagnostics possibles, des questions à poser à l'appelant ou des conseils pour hiérarchiser les ressources¹¹⁵. Cela aide les répartiteurs à prendre des décisions plus rapides et mieux éclairées dans les situations où le temps est compté.
- Le système d'IA de *Corti* apprend en permanence à partir des nouveaux appels d'urgence et des commentaires des répartiteurs, en affinant ses algorithmes et en améliorant ses performances au fil du temps. Ce processus d'apprentissage continu garantit que le système reste au fait des dernières connaissances médicales et des meilleures pratiques en matière d'intervention d'urgence.



Bénéficiaire direct du système

Les répartiteurs et le personnel médical d'urgence sont ceux qui bénéficient directement du travail du SIA, autant par leur interaction directe avec lui que par le besoin auquel le système répond. Le système aide les employés sous pression à identifier les informations critiques, hiérarchiser les ressources et prendre de meilleures décisions. Cette assistance réduit la charge cognitive des intervenants d'urgence et leur permet de se concentrer sur d'autres aspects importants de leur travail, tels que la coordination des ressources et le soutien émotionnel aux appelants.

En fournissant des diagnostics plus précis et en aidant les répartiteurs à prendre de meilleures décisions, le système optimise indirectement l'allocation des ressources. Cette efficacité accrue peut conduire à une meilleure performance globale de l'organisation, réduisant potentiellement les coûts et améliorant la confiance du public à l'égard des services d'urgence. Le patient bénéficie quant à lui d'une prise de décision plus rapide et plus précise lors des appels d'urgence, ce qui augmente ses chances de recevoir un traitement approprié en temps voulu.



Degré d'autonomie du système

Le SIA analyse la parole et les bruits de fond en temps réel et fournit des suggestions et des informations aux répartiteurs des services médicaux d'urgence. Les répartiteurs restent responsables de la prise de décision et de l'affectation des ressources. *Corti* est un outil supplémentaire qui améliore la capacité des répartiteurs à prendre des décisions éclairées, mais il ne les remplace pas et ne fonctionne pas sans leur intervention. Le SIA travaille en tandem avec les répartiteurs humains pour améliorer la rapidité, la précision et l'efficacité des interventions médicales d'urgence. Il s'agit donc d'un système de degré d'autonomie 3 (semi-autonome).

¹¹⁵ *This AI detects cardiac arrests during emergency calls* | World Economic Forum, 2018 : <https://www.weforum.org/agenda/2018/06/this-ai-detects-cardiac-arrests-during-emergency-calls/>



Pistes de réflexions éthiques

Bénéfices et avantages du SIA

- Accompagne les répartiteurs d'urgence. *Corti* suggère en temps réel des questions à poser aux appelants et des sujets supplémentaires à explorer. Il propose des probabilités concernant la présence de problématiques de santé chez l'appelant ou la personne pour qui l'appel est fait, guidant le répartiteur au cours de l'appel.
- Améliore l'efficacité des interventions du personnel de la santé. *Corti* peut aider les répartiteurs d'urgence, les infirmiers et les médecins dans leurs diagnostics et leurs recommandations de traitement. Cela peut conduire à un gain de temps considérable pour ces professionnels et permettre un traitement plus efficace des patients. De plus, dans des domaines comme la santé où la demande est énorme et la main-d'œuvre peut être difficile à trouver, les systèmes comme *Corti* peuvent fournir une assistance précieuse.
- Détecte des problèmes de santé. Le système a été initialement développé pour détecter les arrêts cardiaques lors des appels d'urgence, une tâche déjà complexe. Sa capacité à détecter d'autres problématiques de santé est potentiellement significative.
- Utilise les interactions passées et récentes entre le personnel de la santé et les patients afin d'améliorer son efficacité. Le système a été formé sur une grande quantité de données, ce qui lui permet de tirer des enseignements à partir d'un vaste ensemble d'expériences de santé et de scénarios d'urgence. Le système continue d'apprendre et de s'améliorer au fur et à mesure de son utilisation, augmentant potentiellement sa précision et son efficacité.

Inconvénients et risques du SIA

- Biais et discrimination. Comme le système a d'abord été entraîné sur une base de données en langue danoise, il pourrait y avoir des problèmes si une personne ne parle pas parfaitement la langue sur laquelle le système a été entraîné, ou si elle a un accent distinct ou une difficulté langagière.
- Manque de transparence. Le processus décisionnel de *Corti* ressemble à une *boîte noire*, ce qui pourrait rendre difficiles la compréhension et la validation des recommandations du système, ainsi que la détection d'erreurs éventuelles.
- Fiabilité.
 - Puisque le système *Corti* a initialement été conçu pour détecter les arrêts cardiaques, on peut se demander s'il est réellement efficace et fiable pour détecter d'autres problèmes de santé.
 - Risque de faux positifs. Cela pourrait amener à des recommandations de procédures inappropriées, comme une réanimation cardiaque inutile, ce qui entraîne des risques pour la santé du patient. De plus, les faux positifs peuvent conduire à une *fatigue des alertes* chez les répartiteurs de service d'urgence et à une perte de confiance à l'égard du système.
- Remplacement du personnel de santé. Il y a un risque que les systèmes comme *Corti* en viennent à remplacer une partie du personnel de santé, ce qui pourrait diminuer la qualité des soins prodigués et réduire l'apport du jugement critique de l'humain.

Valeurs à prendre en considération pour le déploiement et l'utilisation de ce SIA (ou d'un système similaire)

- Acceptabilité de la part du personnel. Le déploiement doit prendre en compte la formation et l'acceptation du personnel qui utilisera le système, et il doit y avoir des mécanismes en place pour résoudre les difficultés ou les réticences.
- Continuité du service public. Il est crucial d'avoir des alternatives et des solutions en cas de dysfonctionnement du système. Le personnel doit être préparé à accomplir ses tâches sans l'assistance du système afin d'assurer la continuité des services de santé.
- Fiabilité. Il est important de considérer comment le système a été développé et formé pour effectuer ses tâches, en particulier comment son utilisation a été généralisée pour détecter une variété de problèmes de santé. De plus, le système doit minimiser les faux positifs et les faux négatifs pour éviter les conséquences potentiellement néfastes pour les patients et pour éviter la *fatigue des alertes* chez les répartiteurs d'appels d'urgence.
- Inclusion et non-discrimination. Le système doit être en mesure de gérer une variété de langues, d'accents, et de capacités linguistiques différentes. Il faut aussi s'assurer que les différences de langues ou d'accents n'entraînent pas de biais dans l'utilisation du système.
- Perfectionnement. Prévoir des mécanismes pour évaluer, surveiller, et améliorer le système au fil du temps, et adapter l'utilisation du système en fonction des observations faites.
- Responsabilité. Il est important de clarifier qui est responsable lorsque le système fait une erreur ou lorsqu'un humain ignore un diagnostic correct émis par le système.
- Supervision humaine. Il faut garder un *humain dans la boucle* pour valider l'information et les suggestions du système, pour assurer le maintien de l'esprit critique, et pour éviter une dépendance excessive envers le système.
- Transparence et explicabilité.
 - Il est nécessaire de comprendre comment le système fait ses recommandations et d'avoir une certaine visibilité sur son processus de décision. L'opacité du processus peut créer des difficultés à valider et comprendre les suggestions du système, et à détecter ses éventuelles erreurs.
 - Il faut trouver un équilibre entre l'accélération du traitement de l'information permettant d'octroyer les services, et la transparence nécessaire à la contre-vérification de l'information fournie par le système.

Comment protéger ou restaurer les valeurs affectées par le SIA?

- Assurer la continuité du service. Il est nécessaire d'avoir des plans de contingence pour faire face à des situations où le système tomberait en panne ou deviendrait inopérant. La continuité des services même en l'absence du système est primordiale.
- Évaluation et révisabilité. L'évaluation du système ne doit pas être qu'un exercice ponctuel avant son déploiement, mais être un processus continu pendant l'intégration et l'utilisation du système. De plus, l'évaluation ne doit pas se concentrer uniquement sur le système lui-même, mais également sur la manière dont il est utilisé. L'insertion du système dans une structure existante, et son appropriation par les organisations, employés et patients, peuvent mener à l'émergence de problématiques non envisagées lors de l'évaluation initiale du système en lui-même, et de son utilisation idéale anticipée. Il est suggéré d'adopter des méthodes comme Z-Inspection (Zicari et al., 2021) pour évaluer, dans la pratique, à quel point un système d'IA est digne de confiance. Ceci pourrait permettre d'identifier et d'atténuer les risques associés au déploiement et à l'utilisation de SIA comme *Corti*.

4. Cas d'application de l'utilisation de l'IA dans le secteur public

- Maintenir un *humain dans la boucle*. Les participants soulignent l'importance d'avoir toujours un *humain dans la boucle* de décision. Cela permet de maintenir un élément d'esprit critique et de jugement humain dans le processus, et assure qu'il y a quelqu'un pour intervenir en cas d'erreur du système. La formation et la préparation des employés qui utiliseront Corti doit inclure des informations sur le fonctionnement du système, ses limites, la manière d'interpréter les informations qu'il fournit, et sur ce qu'il faut faire en cas de désaccord avec ses recommandations. Cela peut aider à réduire la résistance à l'adoption du système, à assurer son utilisation efficace et à identifier ses lacunes ou déficiences.
- Responsabilité professionnelle. Il est suggéré que les professionnels de la santé documentent leurs prises de décision et expliquent s'ils sont en accord ou en désaccord avec les recommandations du SIA. Cela pourrait aider à traiter des questions de responsabilité et à fournir des explications ou des précisions lors de situations délicates ou ambiguës.
- Transparence. Il est recommandé que les utilisateurs du système (médecins, infirmiers, répartiteurs d'urgence) comprennent comment le SIA arrive à ses conclusions. Cela pourrait aider à atténuer les problèmes de la *boîte noire* et à augmenter la confiance envers le système.

Autres ressources

Artificial Intelligence and EMS, 2017: <https://www.jems.com/operations/artificial-intelligence-and-ems/>

CTC Networks and Language Models: Prefix Beam Search Explained | by Lasse Borgholt | Corti | Medium, 2018: <https://medium.com/corti-ai/ctc-networks-and-language-models-prefix-beam-search-explained-c1d1ee23306>

Borgholt, L., Havtorn, J. D., Agić, Ž., Søgaard, A., Maaløe, L., & Igel, C. (2020). *Do End-to-End Speech Recognition Models Care About Context?* *Interspeech 2020*, 4352-4356. <https://doi.org/10.21437/Interspeech.2020-1750>

Capital Region of Denmark: Using AI to improve patient care on the national healthcare system: <https://www.corti.ai/stories/capital-region-of-denmark-using-ai-to-improve-patient-care-on-the-national-healthcare-system>

Borgholt, L., Tax, T. M. S., Havtorn, J. D., Maaløe, L., & Igel, C. (2021). *On Scaling Contrastive Representations for Low-Resource Speech Recognition* (arXiv:2102.00850). arXiv. <https://arxiv.org/abs/2102.00850>.

Corti: The importance of understanding a product (2022): <https://www.youtube.com/watch?v=hUuTHkR3agw>

Seattle Fire Medic One and Corti: Why it's important to focus on Quality Improvement (2022): <https://www.youtube.com/watch?v=MEIdD1Fy6lw>

Corti heart attack detection AI can now deploy on the edge with Scandinavian design (2018): <https://venturebeat.com/ai/cortis-heart-attack-detection-ai-can-now-deploy-on-the-edge-with-scandinavian-design/>

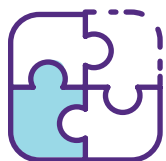
Tableau 10. Description synthétique de *Tengai*

Chronologie	2019 : développement du robot 2022 : introduction de l'avatar en ligne
Pays	Suède
Organisation	Municipalité d'Upplands-Bro
Développeur	Partenariat public-privé (Municipalité d'Upplands-Bro et l'entreprise Furhat Robotics)
Bénéficiaire direct	Organisation (Département des ressources humaines de la municipalité)
Degré d'autonomie	3. Semi-autonome



Objectif poursuivi

L'avatar *Tengai* a été créé pour résoudre le problème des préjugés inconscients dans le processus de recrutement. Ces préjugés inconscients peuvent avoir un impact significatif sur la sélection des candidats à un emploi, ce qui se traduit par une main-d'œuvre moins diversifiée et l'exclusion potentielle de candidatures hautement qualifiées. L'utilisation de l'IA dans les processus de recrutement vise à éliminer les préjugés inconscients et à faire en sorte que toutes les candidatures soient évaluées en fonction de leur mérite. L'avatar *Tengai* est considéré comme un complément au processus traditionnel de recrutement, car il a été conçu pour ne pas percevoir l'âge, le sexe, l'origine ou d'autres aspects des personnes interrogées (ex. style vestimentaire)¹¹⁶.



Étapes de développement

La municipalité suédoise d'Upplands-Bro a entamé une collaboration avec la société suédoise de conseil en intelligence artificielle Furhat Robotics en juin 2019. Le système d'IA nommé *Tengai* prend alors la forme d'un 'robot' (buste pouvant être déposé sur un bureau) et est l'un des premiers robots intervieweurs développés dans le but de rendre le processus de recrutement moins biaisé par rapport aux pratiques d'entretien traditionnelles. L'agence de recrutement et de dotation en personnel de la municipalité travaillait déjà depuis quelque temps à rendre ses processus de recrutement moins biaisés. Au cours des dix années précédant l'introduction du robot, l'agence avait cessé d'utiliser le curriculum vitae dans son processus de recrutement et s'était de plus en plus intéressée aux outils de recrutement basés sur les données. Par conséquent, l'utilisation d'un robot a été considérée comme une étape supplémentaire dans la prévention des préjugés lors des processus de recrutement. L'idée derrière l'utilisation du robot était d'éviter qu'une personne humaine soit présente au début du processus de recrutement afin de neutraliser tout préjugé psychosocial.

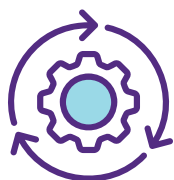
116 AI-X | *Tengai - Robot in recruitment processes*, s. d.: <https://ai-watch.github.io/AI-watch-T6-X/service/90130.html>

4. Cas d'application de l'utilisation de l'IA dans le secteur public

Un élément important du développement du SIA est la collaboration entre la société Furhat Robotics et *The Next Generation in recruitment* (TNG¹¹⁷), une agence suédoise de recrutement agréée. Les recherches de TNG sur les processus d'entretien structurés et non structurés sont le fruit d'une collaboration avec des experts en psychométrie et constituent la base de l'indicateur de performance du système d'exploitation de *Tengai* qui utilise les données issues des 15 années de développement de techniques d'entretien impartiales de TNG. Les résultats de la méthodologie d'entretien de l'étude de validation, combinés à la forme initialement physique de *Tengai*, sont des facteurs qui contribuent à l'expérience du candidat¹¹⁸.

Après les premiers essais du robot, il a été décidé, en l'espace d'une semaine, d'adopter ses services de manière permanente. Pour l'entreprise Furhat Robotics, il s'agissait également d'une première, l'agence municipale de recrutement et d'affectation étant son premier client dans le secteur public. Les premiers résultats obtenus après l'adoption du système ont été considérés comme une réussite et ont attiré l'attention des médias sur la municipalité en raison de son approche de recrutement innovante. Selon l'un des directeurs de la municipalité, le robot a rendu les processus de sélection et d'embauche plus rapides, moins coûteux et plus impartiaux, libérant ainsi des ressources cruciales qui peuvent être consacrées à d'autres tâches¹¹⁹.

À l'automne 2022, la municipalité déploie la nouvelle version de son système d'assistance au recrutement avec l'avatar *Tengai*, disponible en ligne 24/7 et évitant aux candidats de se déplacer pour leur entretien¹²⁰. D'autres municipalités suédoises, dont Nacka, intègrent alors cette version virtuelle du système à leur processus de recrutement¹²¹.



Fonctionnement du système

Suite au travail de collaboration avec le personnel chargé du recrutement visant à analyser et établir les compétences à évaluer, les candidats sont invités à un entretien avec le robot ou l'avatar *Tengai*, selon le cas. Le SIA mène ensuite les entretiens pour évaluer les candidats en analysant leur comportement, leur capacité à résoudre des problèmes et d'autres compétences. Si nécessaire, *Tengai* peut poser des questions complémentaires. Les entretiens sont ensuite analysés et combinés avec les scores de compétences, ce qui permet à *Tengai* d'effectuer un premier tri de candidatures prometteuses. Les recruteurs de la municipalité mèneront l'entretien final afin d'évaluer la motivation des candidats sélectionnés préalablement par le système¹²².

L'avatar *Tengai*¹²³ est conçu pour interagir avec les candidats de manière conversationnelle, en posant la même série de questions à tous les candidats et en appréciant leurs réponses de manière objective. Le système utilise également la technologie de reconnaissance faciale pour détecter les émotions et fournir un retour d'information aux recruteurs.

117 « We are an authorized recruitment and staffing agency specializing in unbiased and competency-based hiring. By challenging the traditional recruitment process, we find a diverse workforce and hidden talents, while contributing to increased diversity, higher innovation and better bottom line for its clients, and new career opportunities for job seekers ». ([About us | Specialists in Unbiased Recruitment | TNG](#))

118 *AI as Gatekeepers to the Job Market : A Critical Reading of, Performance, Bias, and Coded Gaze in Recruitment Chatbots*, 2021, pp. 22-25: <https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1572100&dsid=-6817>

119 AI-X | *Tengai - Robot in recruitment processes*, s. d.: <https://ai-watch.github.io/AI-watch-T6-X/service/90130.html>

120 *Tengai Announces World's First Human-Like Avatar for Unbiased Candidate Conversations* | Tengai Unbiased, s. d. : <https://www.mynewsdesk.com/tengai-interview-robot/pressreleases/tengai-announces-worlds-first-human-like-avatar-for-unbiased-candidate-conversations-3214835>

121 *Swedish Municipality Uses AI To Objectively Select and Engage Candidates In Recruitment* | Tengai Unbiased, s. d. : <https://www.mynewsdesk.com/tengai-interview-robot/pressreleases/swedish-municipality-uses-ai-to-objectively-select-and-engage-candidates-in-recruitment-3230595>

122 AI-X | *Tengai - Robot in recruitment processes*, s. d.: <https://ai-watch.github.io/AI-watch-T6-X/service/90130.html>

123 Nous nous concentrerons sur la version 'avatar en ligne' du système *Tengai*, qui semble être celui favorisé par les municipalités suédoises depuis l'automne 2022. Il est à noter que l'entreprise Furhat Robotics, partenaire initial dans ce projet, poursuit le développement de la version 'robot' du système pour utilisation commerciale dans des lieux physiques comme des banques, des aéroports, etc. Leur système propose l'intégration du robot et de l'avatar : « the groundbreaking combination of facial animation and replaceable masks make it easy and affordable to create expressive characters for any purpose ». Source: The Furhat Robot | Furhat Robotics, s. d.: <https://furhatrobotics.com/furhat-robot/>

Tengai utilise une combinaison d'algorithmes de traitement du langage naturel (NLP) et d'apprentissage automatique pour analyser les réponses des candidats.

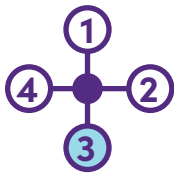
- Les algorithmes de traitement du langage naturel lui permettent d'analyser les réponses des candidats aux questions ouvertes et d'en extraire le sens. Le système analyse le langage, le ton et le contexte des réponses du candidat pour déterminer ses traits de personnalité.
- Les algorithmes d'apprentissage automatique lui permettent d'apprendre à partir des données qu'il recueille et d'améliorer sa précision au fil du temps. Le système est entraîné à partir d'un vaste ensemble d'entretiens et d'évaluations de la personnalité afin d'identifier des modèles et des corrélations entre l'utilisation de la langue et les traits de personnalité. Au fur et à mesure qu'il traite davantage de données, il devient plus apte à identifier des modèles et à faire des prédictions précises sur les traits de personnalité des candidats.

Pour ce qui est de l'évaluation des traits de personnalité, *Tengai* utilise l'évaluation standardisée appelée B5 (pour Big Five), qui est une variante du modèle à cinq facteurs (FFM) de la personnalité. L'intégration du modèle a été facilitée par la collaboration avec *Next Generation in recruitment* (TNG) dont il a été question plus tôt. Le modèle à cinq facteurs mesure cinq grandes dimensions de la personnalité : l'ouverture, la conscience, l'extraversion, l'amabilité et la stabilité émotionnelle¹²⁴. L'évaluation B5 de *Tengai* consiste en une série de questions conçues pour évaluer les traits de personnalité du candidat sur la base de ces dimensions. Les questions sont présentées sous forme de conversation et les algorithmes d'intelligence artificielle de *Tengai* analysent les réponses pour déterminer le profil de personnalité des candidats.



Bénéficiaire direct du système

Bien que la personne en recherche d'emploi utilise le système qui lui sert d'interface durant le processus de recrutement, c'est le département des ressources humaines en tant que responsables du recrutement qui en est le bénéficiaire direct. Les employés sont responsables du déploiement du système et de la prise de décisions d'embauche, mais l'organisation bénéficie directement de la présélection effectuée par le système qui lui permet d'orienter les ressources allouées au recrutement vers les candidats les plus prometteurs.



Degré d'autonomie du système

Tengai effectue des tâches de manière autonome comme la conduite d'entretiens et l'évaluation des candidats sur la base de leurs réponses. Le système établit la liste des candidats qui répondent le mieux aux critères préétablis et qui devront être rencontrés en priorité par les responsables du recrutement de la municipalité. Le système exerce un certain pouvoir discrétionnaire, priorisant des candidats pour la suite du processus d'embauche et en éliminant d'autres. Bien qu'il nécessite une intervention humaine pour l'établissement des critères à évaluer lors des entretiens et pour la décision finale d'embauche, le système prend part au processus de décision en interagissant en temps réel avec les candidats. Il facilite le travail de recrutement de la municipalité et permet de réduire l'impact des préjugés inconscients dans le processus d'embauche. Il s'agit d'un système de degré d'autonomie 3 (semi-autonome).

124 The science behind *Tengai's* validated candidate assessment, s. d. : <https://tengai.io/candidate-assessment/>



Pistes de réflexions éthiques

Bénéfices et avantages du SIA

- Allocation optimisée des ressources. Le système *Tengai* permet une meilleure allocation des ressources humaines et financières pour les organisations qui l'utilisent.
- Efficacité accrue du recrutement. *Tengai* permet d'augmenter l'efficacité du processus de recrutement en effectuant le premier tri des candidatures pour des postes à pourvoir. Il contribue ainsi à la sélection des meilleures candidatures et allège le travail des recruteurs humains.
- Réduction des biais de sélection. Le système a été conçu pour ne pas percevoir l'âge, le sexe, l'origine, ou d'autres aspects comme le style vestimentaire des personnes en recherche d'emploi. Cela a le potentiel de réduire certains biais de sélection et d'éviter les préjugés inconscients qui sont difficiles à éliminer lorsque des humains sont impliqués dans le processus de recrutement.
- Ajustement et apprentissage continu. Possibilité d'ajustement et d'affinement du système en s'appuyant sur les nouvelles données issues de nouveaux entretiens d'embauche, ou en modifiant les critères évalués par le système lors des entretiens.

Inconvénients et risques du SIA

- Accumulation de données. Le système recueille et gère un grand nombre de données biométriques et autres données personnelles, ce qui peut poser des problèmes de confidentialité et de protection des données. La nécessité et la pertinence de la collecte de ces données sont questionnables. De plus, le système utilise des technologies de reconnaissance faciale et de reconnaissance des émotions pour capter des données biométriques, ce qui peut être vu comme une intrusion dans la vie privée. De plus, la fiabilité de ces technologies est remise en question.
- Déshumanisation du processus d'embauche. L'usage d'un robot ou d'un avatar pour gérer les premières étapes du recrutement peut rendre le processus impersonnel et froid. De plus, le phénomène de *l'uncanny valley* (vallée de l'étrange)¹²⁵ fait en sorte que certaines personnes pourraient se sentir mal à l'aise ou inconfortables lors d'un entretien mené par un robot ou un avatar.
- Doute sur l'objectivité promise et risque de biais. L'idée que le système *Tengai* offre une objectivité supérieure par rapport aux humains peut être vue comme une forme de techno-utopisme, qui ignore les nombreux biais potentiels introduits par l'IA. Malgré la conception du système pour conduire des entretiens sans préjugés psychosociaux, certains participants aux ateliers de réflexion éthique estiment que d'autres types de biais ou de discriminations pourraient survenir. Par exemple, la façon dont les critères de sélection sont définis peut introduire des biais dans le processus de sélection, même si ces critères sont censés être objectifs. De plus, le système pourrait faire des associations imprévues entre des données intermédiaires et des critères officiels, ce qui pourrait mener à des *discriminations indirectes*. Des individus avec certaines caractéristiques (ex. : aveugle, ayant du mal à s'exprimer, ayant un fort accent) pourraient aussi être désavantagés par le système.

¹²⁵ La notion de *vallée dérangement* ou *vallée de l'étrange* a été présentée dans les années 1970, par l'ingénieur en robotique Masahiro Mori. Le scientifique japonais faisait l'hypothèse que plus les robots ont une apparence humaine, plus les observateurs humains font preuve d'empathie et d'émotions positives à l'égard de ces sujets. Cette pente ascendante finit toutefois par atteindre le haut de la vallée où les émotions changent : passé un certain stade de ressemblance, l'émotion prédominante est la répulsion. Masahiro Mori postulait qu'après ce pic, il y a la pente descendante de l'autre côté de la vallée : une fois ce stade franchi, si les robots deviennent réellement difficiles à distinguer des humains, les réponses émotionnelles redeviennent positives. Source : <https://www.numerama.com/tech/838037-quest-ce-que-la-vallée-dérangement-quand-on-regarde-un-robot-humanode.html>

- Iniquité et risques de préjudices. Ce système est davantage utile pour les recrutements où les bassins de candidatures sont importants. Il s'agit habituellement des processus de recrutement pour des postes moins élevés dans la hiérarchie et qui s'adressent à des personnes qui peuvent être plus vulnérables. Ces dernières risqueraient d'être davantage exposées au risque de préjudices de ces systèmes automatisés, car il y a probablement moins de surveillance et de contrôle humain pour ces processus d'embauche que pour ceux qui concernent des postes plus élevés dans la hiérarchie d'une organisation.
- Risque de création d'une chambre d'écho. Le système, en continuant d'apprendre sur les nouvelles données issues des entretiens d'embauche, pourrait devenir de plus en plus spécifique sur le type de candidat qu'il évalue positivement, ignorant ainsi les candidats atypiques qui pourraient convenir pour les postes. Un processus d'embauche étant très contextuel, certains aspects, tels que l'aisance à interagir avec une équipe ou la collégialité, pourraient être difficiles à évaluer par le système.
- Risque de manipulation des résultats. Avec l'avatar en ligne, il existe un risque que les résultats soient faussés, car certains candidats (ou des tiers) pourraient tricher avec l'avatar ou savoir quoi dire et à quel moment pour obtenir les meilleurs résultats possibles.

Valeurs à prendre en considération pour le déploiement et l'utilisation de ce SIA (ou d'un système similaire)

- Dignité humaine. Des préoccupations ont été exprimées quant à la déshumanisation du processus de recrutement avec l'utilisation d'un robot ou d'un avatar pour les entrevues. Certains intervenants estiment qu'un processus de recrutement doit nécessairement impliquer un humain, en particulier au début de celui-ci.
- Équité et impartialité. *Tengai* est conçu pour éliminer certains préjugés humains en ignorant des aspects comme l'âge, le sexe, l'origine et le style vestimentaire des personnes postulant à un emploi. Toutefois, l'utilisation du système pourrait avoir un impact sur les candidatures atypiques qui pourraient ne pas être évaluées de manière équitable, menant potentiellement à leur exclusion du processus de recrutement.
- Explicabilité et transparence. Les biais et les discriminations indirectes pourraient être difficiles à comprendre, répertorier et prendre en considération, surtout dans le contexte d'un système d'apprentissage automatique comme *Tengai*.
- Objectivité. Une évaluation attentive du système est nécessaire face à la promesse d'une objectivité parfaite, le système pouvant introduire de nouveaux biais non intentionnels.
- Légalité. Il faudrait évaluer si le fonctionnement du système *Tengai* est conforme à la législation en vigueur, notamment en matière de lutte contre la discrimination et de la protection des données.

Comment protéger ou restaurer les valeurs affectées par le SIA?

- Analyse de la pertinence et de la performance du système. Considérant les préoccupations soulevées par ce SIA, il est nécessaire d'évaluer la pertinence de son utilisation. Une vérification continue de sa performance sera essentielle tout au long de son déploiement et son utilisation. D'autre part, il pourrait être envisagé d'utiliser un tel système dans un rôle de *conseil* et éviter de lui confier l'évaluation et la hiérarchisation des candidats. La finalité du système serait alors différente et le degré de substitution du jugement humain beaucoup moins grand.
- Confidentialité et protection des données. Des politiques claires sur la collecte, l'utilisation et la protection des données devraient être mises en place. Il serait également nécessaire d'assurer la transparence sur qui gère ces données et comment elles sont utilisées (référence à la loi 25 qui énonce les principes de proportionnalité et de nécessité concernant la collecte de données¹²⁶).
- **Inclusion numérique.** Le système devrait être conçu de manière à permettre à toutes les personnes postulant à un emploi public d'avoir la possibilité de le faire. Cela pourrait nécessiter de permettre la réalisation de l'entretien avec un humain pour ceux qui préfèrent cette méthode.
- Implication humaine. Il faut éviter un déplacement vers des processus de recrutement entièrement automatisés. L'interaction humaine est importante dans ce processus, surtout pour le premier contact avec les candidats.
- Adapter le processus d'entrevue. Le système devrait être conçu de manière à être aussi convivial possible pour tous les candidats. Cela pourrait inclure l'option d'avoir un entretien avec un humain pour ceux qui préfèrent cette méthode.

126 En ligne : <https://www.quebec.ca/nouvelles/actualites/details/loi-25-nouvelles-dispositions-protégeant-la-vie-privee-des-quebecois-certaines-dispositions-entrent-en-vigueur-aujourd'hui-43212>

Note concernant les cas suivants

Il n'existe pas d'exemple concret de SIA correspondant à notre description du degré d'autonomie 4 (autonome), que ce soit dans le secteur privé ou public. Rappelons que notre catégorisation stipule que les systèmes d'IA autonomes :

- peuvent planifier, établir des priorités et accomplir des tâches sans intervention humaine,
- peuvent fixer leurs propres objectifs et s'adapter à de nouvelles situations en tirant les leçons de leurs expériences,
- peuvent analyser et optimiser leurs performances de manière indépendante, même dans des environnements complexes et incertains.

Pour les besoins de l'exercice proposé, nous présenterons deux cas d'application projetée ou en cours de développement, à partir des documents officiels des instances publiques qui les mettent de l'avant. Bien que ces exemples soient spéculatifs, ils sont ancrés dans la réalité des organisations qui les envisagent et dans la particularité des lieux visés par leur déploiement éventuel. Notre but est de susciter la discussion au sujet des objectifs visés, du fonctionnement envisagé et du degré d'autonomie décisionnelle projetée de systèmes qui nous semble excéder le degré d'autonomie 3 (semi-autonome) de notre catégorisation.

AuroraAI

Tableau 11. Description synthétique d'AuroraAI

Chronologie	<p>2018–2019 : projet pilote à petite échelle et processus de consultation¹²⁷</p> <p>2019 : décision de déployer le programme de 100 millions d'euros à plus grande échelle pour la période 2019–2023</p> <p>2020 : création d'un comité d'éthique¹²⁸</p> <p>2022 : mise en place du réseau <i>AuroraAI</i></p> <p>Le programme sera développé et déployé progressivement, par le biais d'expérimentations, sur une période de plusieurs décennies.</p>
Pays	Finlande
Organisation	Ministère des Finances et autres organisations publiques associées aux événements de vie des usagers
Développeur	Partenariat public envisagé (Finnish Digital Agency for the network)
Bénéficiaire direct	Usagers
Degré d'autonomie	4. Autonome (à confirmer en fonction de l'évolution du projet)

**Objectif poursuivi**

AuroraAI est un projet finlandais visant à construire une infrastructure numérique nationale où les structures de service actuelles de la société sont transformées en entités de service unifiées utilisant l'IA. L'objectif est d'améliorer la capacité des organisations à renforcer le bien-être des personnes en fournissant des services pour les différents *événements de vie* de manière intelligente, en collaboration entre les différents secteurs et fournisseurs de services. Le programme vise également à réduire les coûts des services et créer des opportunités d'intégration des services publics et privés¹²⁹.

127 *AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings* (Kuziemski & Misuraca, 2020, p. 9): <https://www.sciencedirect.com/science/article/pii/S0308596120300689>

128 *Governing Ethical AI Transformation: A Case Study of AuroraAI* (Leikas et al., 2022, pp. 6–7) : <https://www.frontiersin.org/articles/10.3389/frai.2022.836557/full>

129 *Ibid*, pp. 4–5

The world's best public administration enables a human-centred, proactive society, where people learn to understand their own wellbeing, and where services seek their way to people in a timely, information secured and ethically sustainable manner. The starting point of a human-centred society is the comprehensive wellbeing of citizens, companies and society. The human-centred approach finds its concrete form in life-events based thinking, which guides the operations of organisations, and the planning and implementation of services. This approach will create the best possible conditions for individuals to support their own wellbeing and the wellbeing of their loved ones in different life situations¹³⁰.

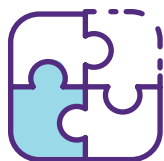
Sous la responsabilité du ministère des Finances et touchant plusieurs départements, le programme vise à organiser la production et la fourniture de services publics d'une manière plus individualisée, avec l'aide de l'apprentissage renforcé. Plutôt qu'un service unique, *AuroraAI* est une tentative de révolutionner les pratiques actuelles de gestion publique, qui ne sont plus axées sur l'efficacité et les résultats, mais qui placent les besoins des clients au centre de l'interaction. Ce *centrage sur l'humain* est envisagé comme un développement itératif des services publics avec des contributions significatives des citoyens et des entreprises, et à travers le prisme des *événements de vie*, c'est-à-dire des moments/situations qui nécessitent une interaction accrue entre le citoyen et l'État (ex. naissance, décès, etc.). *AuroraAI* se veut une plateforme ou un réseau de services, où l'opérateur public fixe des exigences spécifiques en matière de technologie et de processus, ainsi que des limites éthiques, permettant à quiconque de développer sa *proposition de valeur* au sein de la plateforme. La raison d'être du projet découle de l'écart croissant constaté entre la viabilité des finances publiques et la détérioration du taux de dépendance, ainsi que de l'espoir que de nouvelles chaînes de services personnalisés répondront mieux aux réalités changeantes du XXI^e siècle¹³¹.

Le programme *AuroraAI* vise à créer une société centrée sur l'humain et alimentée par l'IA, avec des capacités prédictives et un secteur public transformé. Une société dans laquelle le secteur public et d'autres fournisseurs de services sont conscients en temps réel des besoins, des défis et du bien-être des citoyens. Cette société utilisera l'IA et d'autres technologies de pointe pour permettre aux gens d'atteindre individuellement leurs objectifs de vie. Pour ce faire, elle offrira des services pertinents et des écosystèmes entiers, en temps réel ou de manière proactive, aux groupes de population et aux individus pour favoriser le bien-être collectif et individuel. Cette société orientera et planifiera la production de services, en combinant tous les secteurs, pour répondre aux besoins réels, sur la base de toutes les données disponibles au sujet des personnes et des différents segments de la population¹³².

130 *Leading the Way into the Age of Artificial Intelligence - Final Report of Finland's Artificial Intelligence Programme 2019*, 2019, p. 85: <https://julkaisut.valtioneuvosto.fi/handle/10024/161688>

131 *Governing Ethical AI Transformation: A Case Study of AuroraAI* (Leikas et al., 2022, p. 9): <https://www.frontiersin.org/articles/10.3389/frai.2022.836557/full>; *Hello, World: Artificial intelligence and its use in the public sector. Draft primer for public servants on the uses and considerations for AI in supporting public sector innovation and transformation* (Hello, World, 2019, pp. 116-120): <https://oecd-opsi.org/wp-content/uploads/2019/08/AI-Primer-Stakeholder-Feedback-Version.pdf>

132 *Ibid.*, p. 9



Étapes de développement

Le démarrage d'*AuroraAI* fait suite aux publications de deux documents d'orientation du gouvernement finlandais : 1) *Finland's Age of Artificial Intelligence*¹³³, publié en 2017 par le ministère des Affaires économiques et de l'Emploi, qui donnait un aperçu de l'ambition du pays et de la compréhension de son avantage concurrentiel concernant l'IA. 2) *Finnish AI Competences and How to Make Them Stronger*¹³⁴, publié en 2019 par le Centre d'Analyse, d'Évaluation et de Recherche du gouvernement, qui comparait les compétences en IA des différents domaines et secteurs du pays.

Ces documents montrent que les efforts de la Finlande en matière d'intelligence artificielle ont été menés d'une manière différente des autres stratégies nationales. Au-delà de la publication d'une stratégie globale, le gouvernement tente de modifier son propre mode de fonctionnement en utilisant des services d'IA. *AuroraAI* envisage ainsi l'utilisation de l'apprentissage par renforcement pour organiser la prestation de services publics de manière plus individualisée. Le programme débute en 2018 avec un projet pilote d'une durée de cinq mois, suivi d'un processus de consultation de trois mois qui s'est achevé en avril 2019¹³⁵.

Après la conclusion du projet pilote, il a été décidé de déployer le programme sur une plus grande échelle pour la période 2019–2023¹³⁶, avec un budget alloué de 100 millions d'euros. Au cours de la phase de consultation et de planification, les parties prenantes ont souligné l'ambition du programme et l'ampleur du changement culturel qu'il vise et ont fait remarquer qu'il n'était peut-être pas réaliste d'y parvenir avec les ressources et le calendrier actuels. Le programme a été salué pour son développement progressif et ses expérimentations¹³⁷.

AuroraAI repose d'abord sur la création d'un réseau permettant d'intégrer automatiquement les services les uns aux autres. Ce réseau, développé par l'Agence des services numériques et des données sur la population (l'Agence numérique finlandaise) utilise l'IA pour permettre l'interaction réciproque et autonome des services qui lui sont connectés. La mise en place de ce réseau a été complétée à la fin de 2022 en y intégrant un nombre limité de services¹³⁸.

Le concept de services dont il est ici question va bien au-delà du numérique : une piscine locale peut exister en tant que service, connecté au réseau *AuroraAI* de la même manière que les sources de données ou les logiciels. Ceci est rendu possible grâce à une interface de programmation d'application (API/Application Programming Interface) commune entre tous les services compatibles avec *AuroraAI* et conduit à une véritable numérisation : une combinaison de services numériques et non numériques interagissant les uns avec les autres¹³⁹.

133 *Finland's Age of Artificial Intelligence* (Finland Government, 2017) : https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf

134 *Finnish AI Competences and How to Make Them Stronger* (Finland Government 2019) : <https://tietokayttoon.fi/documents/113169639/113170760/3-2019-Finnish+AI+Competences+and+How+to+Make+Them+Stronger.pdf/e7448517-3d07-e52d-07ad-fb023572b8a2/3-2019-Finnish+AI+Competences+and+How+to+Make+Them+Stronger.pdf?version=1.0&t=1550127909000>

135 *AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings* (Kuziemski & Misuraca, 2020, p. 8-9) : <https://www.sciencedirect.com/science/article/pii/S0308596120300689>

136 *AuroraAI development and implementation plan 2019–2023* : <https://vm.fi/documents/10623/1464506/AuroraAI+development+and+implementation+plan+2019%E2%80%932023.pdf>

137 *Humans, Artificial Intelligence and Society – What Is AuroraAI?* : <https://www.eoppiva.fi/kurssit/humans-artificial-intelligence-and-society-what-is-auroraai/#/>, Theme 2 of 6.

138 *AuroraAI national artificial intelligence programme* : <https://digifinland.fi/en/our-operations/aurora-ai-national-artificial-intelligence-programme/> ; *The AuroraAI Programme* (Vanhanen & Paatero, 2019, p. 13) : https://www.europarl.europa.eu/cmsdata/240174/AIDA%20workshop_28052021_Alanko.pdf

139 *Governing Ethical AI Transformation: A Case Study of AuroraAI* (Leikas et al., 2022, p. 6) : <https://www.frontiersin.org/articles/10.3389/frai.2022.836557/full>

Dans la proposition initiale du projet, les utilisateurs étaient des acteurs totalement anonymes qui ne pouvaient être identifiés qu'à l'aide d'adresses IP. Un anonymat fort et protégé était considéré comme un facteur favorisant la confiance et évitant les risques d'atteinte à la vie privée et à la sécurité des données. Les notions de profil d'utilisateur et de gestion de profil sont cependant apparues, ainsi que le concept de compte *AuroraAI*. Il devint évident qu'avec l'addition de ces nouveaux éléments, les questions de protection de la vie privée et d'anonymat devaient être réexaminées puisque la notion d'anonymat automatiquement garanti était compromise.

L'addition des profils d'utilisateurs et/ou du compte *AuroraAI* découle d'un élément maintenant central du programme : l'amélioration du bien-être individuel et collectif basée sur les données. S'il n'existe aucun moyen de suivre les états de bien-être des utilisateurs après qu'ils ont reçu des recommandations de services basées sur leurs attributs donnés à un certain moment, les algorithmes d'apprentissage automatique ne pourront pas être utilisés pour générer des recommandations précises et, éventuellement, des prédictions. Sans cette capacité, le système ne sera qu'un service de recherche intelligent qui servira une personne à la fois, selon sa situation unique et temporaire, puis l'oubliera totalement. Sans le regroupement et la conservation des données pour chaque utilisateur distinct de la plateforme, il sera impossible pour les algorithmes du système d'apprendre comment les recommandations ont affecté le bien-être de l'utilisateur. Ceci est d'autant plus important qu'*AuroraAI* est ultimement envisagé comme un *guide de vie* personnel qui peut apprendre des attributs d'une personne et de l'historique des événements au fil du temps, et peut ainsi générer une orientation, une prestation de services et des recommandations prédictives/proactives très précises et pertinentes. Ceci nécessite de stocker les attributs, les événements et les commentaires de l'utilisateur dans un compte qui, s'il est anonyme, demeure spécifique à cet utilisateur et requiert une gestion de profil (dans le cadre d'*AuroraAI*, la solution DigiMe serait utilisée)¹⁴⁰.

De l'aveu même des initiateurs du programme, la mise en place complète et fonctionnelle du système se fera sur une longue période. Sa forme définitive ainsi que les technologies qui seront utilisées sont donc matière à spéculation. À la fin 2022, les parties prenantes au développement du système se concentraient sur 1) la finalisation de la première phase de développement technique de la plateforme *Aurora* et du moteur de recommandation de services, et 2) le développement de modèles opérationnels pour diriger la production de *services du futur, centrée sur l'homme et basée sur les données*, et la restructuration conséquente de toute la production de services par le secteur public.

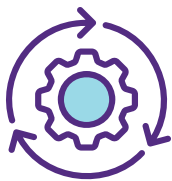
Les éléments techniques les plus ambitieux du programme s'inscrivent dans un avenir plus lointain et pourraient prendre de nombreuses années avant d'être complètement testés, analysés et intégrés à la plateforme. Cela s'applique en particulier au concept DigiMe¹⁴¹, aux recommandations de services automatiques et personnels, et à l'évaluation (et l'ajustement) aux effets de l'interaction humain-machine sur la production de services et la société en général¹⁴².

140 Ibid, pp. 8-9

141 The DigiMe, or a "holistic 360° profile," is essentially a digital profile or mirror, composed of certain data, attributes and their values of an individual and their situation in life. The DigiMe centers around a group of eight parameters based on the "The Stiglitz Model" eight-point list of wellbeing factors introduced by Stiglitz et al. (2009). These are: (i) Material living standards (income, consumption, and wealth); (ii) Health; (iii) Education; (iv) Personal activities including work; (v) Political voice and governance; (vi) Social connections and relationships; (vii) Environment (present and future conditions); and (viii) Insecurity, of an economic as well as a physical nature. Source: *Governing Ethical AI Transformation: A Case Study of AuroraAI*: <https://www.frontiersin.org/articles/10.3389/frai.2022.836557/full>

Voir aussi: *AuroraAI development and implementation plan 2019-2023* (p. 21): <https://vm.fi/documents/10623/1464506/AuroraAI+development+and+implementation+plan+2019%E2%80%932023.pdf>

142 *Governing Ethical AI Transformation: A Case Study of AuroraAI* (Leikas et al., 2022, p. 9): <https://www.frontiersin.org/articles/10.3389/frai.2022.836557/full>



Fonctionnement du système

Le programme national finlandais *AuroraAI* vise à créer une société dans laquelle les organisations publiques et privées coopèrent pour faire en sorte que les personnes puissent faire face à différents *événements de vie* de manière simple et pratique tout au long de l'existence. Les écosystèmes multipartites à interaction flexible utilisent des solutions d'IA pour développer des types de services entièrement nouveaux qui sont adaptés aux situations de vie des personnes et à ce que les entreprises peuvent offrir :

The program simultaneously intertwines governance and policy change, regulatory issues, technological innovation in a multi-vendor environment, the pooling of private and public sector interests, data-based modeling of individual and population situations, new procedures and actors' roles in the production and management of public services, as well as increasing the overall wellbeing of the individual¹⁴³.

Les services personnalisés sont basés à la fois sur les données personnelles que les citoyens fournissent (*MyData*¹⁴⁴ ; collectées par exemple grâce à des dispositifs de santé intelligents) et sur des données populationnelles. Les individus seront en mesure de produire eux-mêmes des données et d'y accéder dans un format numérique¹⁴⁵. *AuroraAI* répartit les personnes en différents groupes en fonction de leur *bien-être multidimensionnel*. À partir de ces regroupements, le système cible des ensembles de services appropriés et opportuns pour les événements et les situations de vie individuelles des personnes :

A life situation can be considered as a state of a state machine, which consists of a nearly infinite number of life situations and transitions between them (life events). In this context, a "state" covers all the information *Aurora* services have stored about the user in a distributed, anonymous manner. A change of this data yields a state transition. *AuroraAI* aims to facilitate these transitions by orchestrating optimal micro service combinations from an available pool to meet users' personal needs. People with similar data attributes are considered to reside, partially in similar life situations and therefore to benefit from similar service combinations¹⁴⁶.

Le réseau *AuroraAI* et ses services utilisent l'IA de quatre manières différentes¹⁴⁷ :

- Identifier la situation décrite par l'utilisateur. La situation décrite par l'utilisateur sera identifiée grâce à l'interaction entre l'utilisateur et le réseau *AuroraAI* sans utilisation de données personnelles identifiables.
- Regrouper les services. Plus l'évaluation de la situation de l'utilisateur est précise, plus l'adéquation entre l'offre de services et les besoins de l'utilisateur est grande. L'IA permet d'organiser les services en fonction des besoins identifiés.
- Apprentissage par renforcement. Le réseau *AuroraAI* apprendra à former des combinaisons de services et à les cibler plus précisément en fonction des besoins de l'utilisateur.
- Services d'IA. Les systèmes de référencement du réseau *AuroraAI*, telles que le chatbot *AuroraAI*, peuvent eux-mêmes être assistés par l'IA. Ils peuvent s'utiliser automatiquement et se compléter mutuellement.

143 Ibid, pp. 4-5

144 Finnish data entry: Your digital data can become "MyData": <https://finland.fi/business-innovation/finnish-data-entry-your-digital-data-can-become-mydata/> ; Finland to lead the way in MyData (2018): <https://lym.fi/en/-/finland-to-lead-the-way-in-mydata-980446>

145 Governing Ethical AI Transformation: A Case Study of *AuroraAI* (Leikas et al., 2022, p. 4): <https://www.frontiersin.org/articles/10.3389/frai.2022.836557/full>

146 Ibid, p. 5

147 *Humans, Artificial Intelligence and Society – What Is *AuroraAI*?*: <https://www.eoppiva.fi/kurssit/humans-artificial-intelligence-and-society-what-is-auroraai/#/> , Theme 2 of 6

Le système base la personnalisation à partir des données disponibles sur l'utilisateur et celles fournies par l'interaction avec l'utilisateur. Lorsque l'utilisateur emploie une interface compatible avec *AuroraAI* (chatbot, site web, etc.), il exprime ses besoins et ses souhaits en matière de services. En plus de ces données, la personnalisation utilise également des données anonymisées concernant d'autres utilisateurs.

Le système utilise les données pour cibler les services de trois manières¹⁴⁸ :

- Données sur l'état de bien-être. Plus la compréhension de la situation décrite par l'utilisateur est complète, plus il est possible de se faire une idée des besoins réels de l'utilisateur en matière de services.
- Données sur les services. Plus la compréhension du contenu des services est précise, mieux il est possible de les regrouper et de les identifier pour répondre aux besoins réels.
- Données sur l'impact des services sur le bien-être. Plus le retour d'information sur l'utilité des services est important, plus les ensembles de services qui pourront être créés à l'avenir seront pertinents.

AuroraAI peut être considérée comme une *vision utopique socio-technologique* dans laquelle l'intelligence artificielle libère les personnes, améliore leur bien-être, renforce l'impact des services et accroît leur productivité. Toutefois, rien de tout cela ne se produit soudainement et les étapes actuelles et les premiers pas du programme peuvent produire des résultats de niveau inférieur qui peuvent tout de même s'avérer utiles¹⁴⁹.

We do not know precisely what types of technological solutions will eventually be developed. We cannot as yet fully comprehend how we will ultimately achieve a human-centric approach. The *AuroraAI* programme is a journey on which we will learn through collaboration how to build a human-centric society in a sensible manner utilising the opportunities created by AI¹⁵⁰.



Bénéficiaire direct du système

Le programme *AuroraAI* crée une plateforme à laquelle diverses organisations de différents secteurs peuvent se joindre pour créer une offre de services aux citoyens dans différents événements de leur vie. *AuroraAI* accélère la mise en place d'un écosystème répondant aux besoins des citoyens, du commerce et de l'industrie, où les services intelligents fonctionnent de manière transparente, centrée sur l'humain et interactive. Le bénéficiaire direct est l'utilisateur.

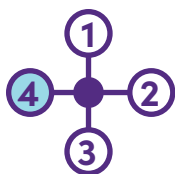
Le système bénéficie indirectement aux organisations publiques finlandaises en leur permettant de rester globalement informées des besoins des utilisateurs du réseau. Il leur permet aussi d'avoir une idée claire des différents prestataires de services et de ce qu'ils offrent¹⁵¹.

148 Ibid

149 *Governing Ethical AI Transformation: A Case Study of AuroraAI* (Leikas et al., 2022, pp. 9-10): <https://www.frontiersin.org/articles/10.3389/frai.2022.836557/full>

150 *Humans, Artificial Intelligence and Society – What Is AuroraAI?*: <https://www.eoppiva.fi/kurssit/humans-artificial-intelligence-and-society-what-is-auroraai/#/>, Theme 4 of 6.

151 *Leading the Way into the Age of Artificial Intelligence – Final Report of Finland's Artificial Intelligence Programme 2019*, 2019, p. 87-88: <https://julkaisut.valtioneuvosto.fi/handle/10024/161688>



Degré d'autonomie du système

Le degré d'autonomie d'*AuroraAI* dépendra de la forme concrète que prendra le système, des technologies qui seront utilisées et de la manière dont elles seront utilisées. Bien qu'indéterminées pour l'instant, les intentions mises de l'avant par les parties prenantes pointent vers un degré d'autonomie équivalent ou dépassant notre description du système semi-autonome. En effet, si l'on considère la tâche du système (le jumelage des usagers à une gamme de services) et acceptons la description proposée par les promoteurs du projet, le système pourra référer automatiquement des individus et des groupes d'individus à des services personnalisés selon leur situation propre en interagissant en temps réel avec l'*environnement* concerné par cette tâche. Cette interaction en temps réel prendra en compte la variation de la situation des individus et groupes d'individus et celle des services offerts par les secteurs publics et privés. Si l'ensemble du projet est réalisé conformément aux ambitions décrites, il s'agira d'un système de degré d'autonomie 4 (autonome).



Pistes de réflexions éthiques

Bénéfices et avantages du SIA

- Amélioration de la qualité de vie des citoyens. Le SIA pourrait anticiper les besoins des citoyens aux différentes étapes de leurs vies et y répondre de manière proactive, améliorant leur qualité de vie et les aidant à atteindre leurs objectifs personnels.
- Répartition efficace des ressources. Le système devrait aider l'administration finlandaise à identifier les besoins des citoyens et à coordonner efficacement la répartition des ressources sur le territoire. Cela pourrait améliorer l'efficacité de l'administration et la rendre plus réactive aux besoins des citoyens.
- Économie de temps pour les citoyens. Le SIA pourrait éviter aux citoyens de signaler eux-mêmes leurs changements de situation à l'État, et ajuster automatiquement les services qui leur sont proposés.
- Améliore potentiellement l'accès aux allocations. En détectant automatiquement les changements de situation des citoyens, le système pourrait aider à garantir que ceux qui ont droit à certaines allocations les reçoivent, même s'ils ne sont pas conscients de leur existence ou si les processus administratifs sont trop lourds pour eux. Ceci permettrait de réduire le phénomène du non-recours aux services qui affectent particulièrement les personnes vulnérables.
- Contribue au bien-être des individus. Le système pourrait être utilisé pour prévenir certains comportements nuisibles en ne proposant pas certaines activités ou publicités susceptibles de nuire à un individu en fonction de l'information détenue sur lui (par exemple, un individu ayant un problème de consommation ou de jeu ne recevrait pas d'information sur les loteries, jeux de hasard ou alcool).
- Contribue au développement de l'État. Le système fait partie d'un projet politique plus large en Finlande, où l'IA est promue comme un avantage et une force. Ce projet pourrait aider à faire avancer le développement de l'IA dans le pays, tout en favorisant la transformation de l'État.

Inconvénients et risques du SIA

- Collecte et confidentialité des données. Ce SIA soulève un risque généralisé d'atteinte à la vie privée des citoyens en lien avec l'utilisation, la sécurisation et le stockage de données personnelles et intimes. Les promoteurs du système semblent promettre l'anonymat, mais indiquent aussi qu'il sera nécessaire de relier les informations à des individus spécifiques pour assurer l'adéquation des services aux besoins individuels. Il est noté que tout ce qui a trait aux services publics et à l'infrastructure publique est considéré comme à haut risque par le projet de règlement sur l'IA (AI Act) de l'Union européenne. De plus, des préoccupations sont soulevées quant au caractère potentiellement invasif du système si un citoyen ne souhaite pas signaler un changement particulier à l'État. Il pourrait également y avoir une disproportion entre le type et la quantité de données recueillies sur les individus et les bénéfices que ces derniers retirent du système.
- Cybersécurité. Il existe un risque que les attaques informatiques ou toute autre forme de piratage puissent potentiellement déstabiliser l'offre de services publics.
- Difficultés pratiques d'implémentation. En raison des nombreux défis et préoccupations, le déploiement d'un tel système pourrait s'avérer difficile à réaliser.
- Iniquité potentielle en raison de la fracture numérique. Malgré l'éducation à la littératie numérique, certaines personnes pourraient ne pas être à l'aise avec la technologie et donc ne pas bénéficier d'un accès égal aux ressources et aux services publics. Ceci soulève une préoccupation plus générale concernant l'interaction humain-machine. Comment se fait cette interaction et comment les besoins sont-ils exprimés et compris par la machine?
- Détournement de la fonction du système. *AuroraAI*, présenté comme un outil visant à améliorer la qualité de vie des citoyens, pourrait être détourné et utilisé principalement comme un outil administratif, servant les intérêts du gouvernement plutôt que ceux des citoyens.
- Risque pour la démocratie. Le gouvernement pourrait utiliser les informations ciblées, région par région, sur les besoins et les désirs de la population pour proposer des politiques ou des projets qui attireraient le vote de certaines catégories de l'électorat. Les organismes publics ne sont pas à l'abri de ce genre d'utilisation partisane des données personnelles des citoyens.
- Risque de paternalisme, de manipulation et d'uniformisation de la société. Préoccupation quant au caractère paternaliste d'un système qui peut prendre des décisions sur ce qui est bon ou mauvais pour un citoyen. Ceci peut être perçu comme une atteinte à la liberté de choix de l'individu. De plus, il existe un risque que le système oriente ou modèle les besoins des citoyens au lieu de simplement répondre à leurs réels besoins et désirs. Cela soulève des inquiétudes particulières concernant l'implication du secteur privé comme acteur important dans le développement et le fonctionnement de ce système. Ceci peut conduire à des conflits d'intérêts et des manipulations de l'offre de services publics.
- Représentativité et rétroaction. Il y a un problème potentiel d'équité en ce qui concerne la représentation des différents groupes de la population dans les rétroactions vers le système. En effet, on observe une tendance à la sous-représentation de certains pans de la société dans les commentaires auto rapportés sur lesquels le système pourrait s'appuyer pour améliorer sa performance. Cette sous-représentation des commentaires et besoins de certains groupes pourrait mener à un traitement inéquitable de ces groupes par le système.

Valeurs à prendre en considération pour le déploiement et l'utilisation de ce SIA (ou d'un système similaire)

- Accessibilité et inclusion. Le système doit être conçu de manière à éviter la discrimination et à garantir une égalité d'accès aux services. Il doit être accessible et utilisable par tous, quelles que soient les compétences numériques des citoyens ou leurs caractéristiques personnelles.
- Bien-être et autonomie. La présentation du 'bien-être' comme un concept unique et peu défini peut être considérée comme une approche paternaliste qui ne respecte pas la diversité des conceptions du bien-être. Le SIA doit respecter le droit des individus à prendre leurs propres décisions sans influence ou manipulation induite, ce qui soulève des questions sur la fine ligne qui sépare la suggestion et la manipulation. En ce sens, il faudra veiller à ce que le système s'adapte aux besoins individuels changeants, et co-évolue avec la société, en respectant les besoins et les désirs des citoyens, sans les orienter ou les modeler outre mesure.
- Intégrité. Le système doit être conçu et utilisé de manière responsable, en évitant l'usage malveillant des données et en prévenant les conflits d'intérêts potentiels.
- Protection de la vie privée. Les données collectées, stockées et utilisées par le système pourraient être très personnelles. Il est donc crucial de veiller à leur confidentialité et de s'assurer qu'elles ne soient pas utilisées de manière inappropriée.
- Proportionnalité. Il doit y avoir un équilibre entre la quantité et la nature de données collectées, et le type de service fourni.
- Respect de la démocratie. Le SIA ne doit pas être utilisé de manière à influencer indûment les décisions politiques ou à favoriser certains groupes ou régions à des fins partisans ou politiques. Plus largement, *AuroraAI* s'inscrit dans un projet politique, ce qui soulève des questions sur l'usage de l'IA par le gouvernement et son impact sur le citoyen.
- Sécurité des données. Le SIA traitera un grand volume d'informations sensibles, et une forte sécurité est nécessaire pour protéger ces données des menaces potentielles comme le piratage ou les cyberattaques.
- Transparence et clarté. Il est important de communiquer clairement les intentions réelles de l'utilisation du système. Le langage vague et les discours promotionnels peuvent camoufler l'objectif réel et survaloriser les capacités du système. De plus, il est essentiel de faire en sorte que le SIA soit utilisé aux fins pour lequel il sera créé. Dans le cas d'*AuroraAI*, il semble qu'il y ait un risque que le système soit détourné de sa fonction officielle et devienne un outil d'administration plus qu'un outil pour le bien-être des citoyens.

Comment protéger ou restaurer les valeurs affectées par le SIA?

- Autonomie. Afin d'éviter un excès de paternalisme, les citoyens devraient conserver la possibilité de refuser certains services ou de choisir parmi plusieurs options.
- Équité. Pour éviter que le système n'exacerbe les inégalités existantes, des efforts devraient être faits pour rendre le système accessible à tous, indépendamment du niveau de compétences numériques des usagers.
- Éthique. L'application de principes éthiques tels que ceux énoncés dans la Déclaration de Montréal est recommandée pour répondre aux différentes préoccupations éthiques soulevées par un système comme *AuroraAI*.
- Partage de données. Les politiques concernant le partage de données avec des tiers, qu'ils soient privés ou publics, doivent être claires et précises.
- Sécurité. Le système doit être conçu avec des mesures de sécurité robustes pour le protéger du piratage et des cyberattaques en vue de garantir la continuité des services publics.
- Vie privée. Étant donné la nature sensible des données et la grande quantité d'informations personnelles que détiendrait ce système, des mesures de sécurité strictes devraient être mises en place pour garantir que les données personnelles sont protégées contre les accès non autorisés.

Autres ressources

AuroraAI – towards a human-centric, proactive society (Kopponen, 2019). Disponible en trois formats :

<https://vm.fi/documents/10623/1464506/>

[AuroraAI+development+and+implementation+plan+2019%E2%80%932023.pdf](https://vm.fi/documents/10623/1464506/AuroraAI+development+and+implementation+plan+2019%E2%80%932023.pdf) (Document)

https://joinup.ec.europa.eu/sites/default/files/event/attachment/2019-04/SEMIC_Webinar%20on%20AI%20and%20PA_Presentation_AuroraAI_Aleksi%20Kopponen%20%28FI%29_09-04-2019.pdf (Webinar)

https://www.youtube.com/watch?v=IZU_ptEr4eE (Vidéo)

AuroraAI: Finland's National Artificial Intelligence Program, 2020: <https://interestingengineering.com/innovation/auroraai-finlands-national-artificial-intelligence-program>

Define the Future with Finland: <https://www.youtube.com/watch?v=FVHHHeUidfY>

The AuroraAI national artificial intelligence programme begins – with the aim of using artificial intelligence to bring people and services together in a better way, 2020: <https://valtioneuvosto.fi/en/-/10623/the-auroraai-national-artificial-intelligence-programme-begins-with-the-aim-of-using-artificial-intelligence-to-bring-people-and-services-together-in-a-better-way>

The AuroraAI: A Human-Centric and Life-Event Based Public Sector Transformation - Observatory of Public Sector Innovation : <https://oecd-opsi.org/innovations/auroraai/>

AuroraAI development and implementation plan 2019–2023: <https://vm.fi/documents/10623/1464506/AuroraAI+development+and+implementation+plan+2019%E2%80%932023.pdf>

The well-being of young people and those looking for work can be supported with AuroraAI network | Digital and population data services agency, 2022: <https://dvv.fi/en/-/the-well-being-of-young-people-and-those-looking-for-work-can-be-supported-with-auroraai-network-network>

AuroraAI Network and Core Component Implementation – Open Source Repository : <https://github.com/vrk-kpa/auroraai-core-components>

SSOSA pour un parc éolien offshore

Tableau 12. Description synthétique de SSOSA¹⁵²

Chronologie	<p>2011 : <i>UK renewable energy roadmap</i></p> <p>2013 : création de <i>ORE Catapult</i></p> <p>2017 : création et financement de <i>HOME Offshore</i></p> <p>2017 : lancement du <i>Offshore Robotics for Certification of Assets (ORCA) Hub</i></p> <p>2022 : intégration de ORCA à <i>The National Robotarium</i></p>
Pays	Royaume-Uni
Organisation	UK Research and Innovation (UKRI), Department for Science, Innovation and Technology (DSIT) et Engineering and Physical Sciences Research Council (EPSRC) ¹⁵³
Développeur	Partenariat public-privé (Innovate UK, divers laboratoires de recherche et entreprises privées)
Bénéficiaire direct	Employés et Organisations
Degré d'autonomie	4. Autonome (à confirmer en fonction de l'évolution du projet)



Objectif poursuivi

Le Royaume-Uni est un leader en matière d'investissements stratégiques dans la production d'énergie éolienne en mer. En 2022, on y comptait 43 parcs éoliens offshore, 2 652 turbines en fonction pour une capacité installée de près de 13,7 GW¹⁵⁴. Cependant, pour accélérer sa transition vers les énergies renouvelables et réduire son empreinte carbone, l'État britannique a des projets ambitieux et prévoit multiplier par quatre l'électricité produite en mer d'ici 2030. Cette croissance continue de la production d'énergie éolienne offshore repose sur les progrès technologiques permettant la réduction des coûts d'exploitation et de maintenance (Operation and Maintenance/O&M) qui représente jusqu'à 30 % du coût total du cycle de vie d'un parc éolien en mer (Offshore Wind Farm/OWF)¹⁵⁵.

152 Symbiotic System of Systems approach

153 Engineering and Physical Sciences Research Council (EPSRC). En ligne : <https://www.ukri.org/councils/epsrc/>

154 Pour comparaison, l'ensemble du parc de production d'Hydro-Québec compte 61 centrales hydroélectriques et 24 centrales thermiques, ce qui représente une puissance installée de 37,2 GW. Source: Hydro-Québec <https://www.hydroquebec.com/production/>

155 *A review: Challenges and opportunities for artificial intelligence and robotics in the offshore wind sector* (Mitchell et al., 2022, p. 10): <https://www.sciencedirect.com/science/article/pii/S2666546822000088?via%3Dihub>

Pour garantir un approvisionnement énergétique fiable, abordable et résilient à partir des plateformes en mer, l'infrastructure nécessite un cycle d'ingénierie continu et complexe, associé à l'inspection, à la réparation, à la logistique, à la maintenance et à la mise hors service de différents sous-systèmes. Les principaux défis auxquels font face les exploitants de parcs éoliens peuvent être relevés grâce à la robotique et l'IA (Robotics and AI /RAI). Il s'agit notamment de :

Reduction of O&M costs – O&M expenditure of an OWF is a barrier which needs to be addressed to further develop the offshore wind sector. Financial risks which inhibit the development of O&M include turbine downtime, managing vessels and personnel, hazardous weather conditions, sea state and increasing distances to shore¹⁵⁶. To reduce costs, new standard operational procedures must be developed for robotics and Artificial Intelligence (RAI) deployment for asset inspection and maintenance. This enhancement in procedures allows for more frequent intervention and yields improvements in the frequency and quality of inspection and repair, also resulting in improved safety for personnel.

Removal of personnel from hazardous working environments – The need to develop RAI for offshore wind sector O&M is reflected by the concurrent need to reduce human presence from the offshore environment, minimizing exposure of the human workforce to dangerous weather conditions and sea states. This represents a non-intervention challenge for deployed robotic platforms.

Ecological issues in expanding offshore wind – The deployment of carbon intensive field support vessels including helicopters, crew transfer vessels, heavy jack up vessels and service operation vessels which are used for maintenance of OWFs and as a hub for engineers to live offshore. This leads to disruption and a potential for lasting detrimental effect on the habitats and species which inhabit these areas¹⁵⁷.

À ces défis s'ajoute le fait que la main-d'œuvre qualifiée dans le secteur offshore est vieillissante et que les nouvelles générations de diplômés rechignent à travailler dans des endroits à haut risque. Les opérateurs recherchent donc des méthodes plus rentables et plus sûres pour l'inspection, la réparation et la maintenance de leurs infrastructures. Une partie de cette solution consiste à déployer des robots et des systèmes autonomes dans l'air, sur les plateformes, à la surface de l'eau et sous la mer. Cela se traduira par une diminution du personnel en mer, une réduction des coûts et une amélioration de la sécurité. La vision à long terme de l'industrie est celle d'un parc énergétique offshore entièrement autonome, exploité, inspecté et entretenu depuis la côte¹⁵⁸.

156 For example, the Dogger Bank Offshore Development Zone is located between 125 and 290 km off the east coast of Yorkshire and extends over approximately 8,660 square kilometres at a water depths ranging from 18 to 63 metres. Source: *Offshore wind* – great.gov.uk international, s. d.: <https://www.great.gov.uk/international/content/investment/sectors/offshore-wind/>

157 *A review: Challenges and opportunities for artificial intelligence and robotics in the offshore wind sector* (Mitchell et al., 2022, p. 2): <https://www.sciencedirect.com/science/article/pii/S2666546822000882?via%3Dihub>

158 *The ORCA Hub: Explainable Offshore Robotics through Intelligent Interfaces* (Hastie et al., 2018, p. 1): <https://arxiv.org/abs/1803.02100>; *Technology Drivers in Windfarm Asset Management* (Barnes et al., 2018): <https://researchportal.hw.ac.uk/en/publications/technology-drivers-in-windfarm-asset-management>

4. Cas d'application de l'utilisation de l'IA dans le secteur public

Pour relever ces défis, de multiples projets de recherche ont été mis en place, avec la participation de divers acteurs des secteurs public et privé, afin de vérifier si des équipages robotisés peuvent être compétents dans la planification et l'exécution des opérations et de l'entretien, telles que les missions d'inspection et de réparation. Certains, parmi ces projets, visent l'intégration et le déploiement de systèmes robotisés pouvant fonctionner en tandem avec les installations offshore existantes et à venir, interagir en toute sécurité en mode autonome ou semi-autonome dans des environnements complexes et à haut risque, et coopérer avec des opérateurs à distance. Pour que cela soit possible, les opérateurs à distance devront avoir un niveau élevé d'information sur ce que font les systèmes autonomes et pourquoi. Cette transparence accrue des systèmes et de leurs processus décisionnels facilitera l'établissement d'une relation de confiance, ce qui est particulièrement important dans les situations dangereuses¹⁵⁹ :

To achieve the [...] goal of autonomous offshore infrastructure a host of different robotic and sensors systems are required to be able to operate independently and in conjunction with each other and humans. There are multiple different areas that are required to be sensorized for accurate monitoring on an offshore energy platform. For example, on a wind turbine the blades require inspection for corrosion, water ingress and internal damage, the internal monopile, housing and nacelle require an array of sensors monitoring for different environmental measurands such as structural vibration, machine vibration, weld strain, humidity, pressure and temperature, the underwater foundation of the turbine requires inspection for scour, corrosion in the foundation and damage to the outlet pipes¹⁶⁰.

Une évaluation conjointe des possibilités des systèmes et des besoins en matière d'opérations et maintenance des parcs éoliens permet d'envisager une architecture numérique symbiotique qui inclut la planification, les besoins fonctionnels, opérationnels et de sécurité pour assurer la prévention des avaries et la résilience des parcs éoliens. Avec un tel éventail de systèmes robotiques et de capteurs développés pour être déployés ensemble et fonctionner pendant de longues périodes avec une interaction humaine limitée, voire inexistante, l'ensemble du système robotique doit être soigneusement conçu et géré. La méthode proposée pour y parvenir efficacement est la création d'une *Symbiotic System of Systems approach (SSOSA)*. Des projets de recherche misent sur cette approche et réfléchissent au développement d'un écosystème symbiotique entre les personnes, les plateformes robotiques, l'infrastructure, l'environnement et les systèmes¹⁶¹.

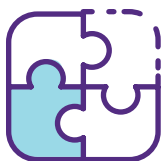
In common with best practice in other sectors, the long term industry vision for the offshore renewable energy sector is advanced predictive maintenance, based on data driven approaches, using improved sensing and digital models that can accurately forecast the future state of health and therefore maintenance requirements of these complex assets¹⁶².

159 Ibid.

160 *The Future Workplace : A Symbiotic System of Systems Environment*. Dans T. Semwal & F. Iqbal, *Cyber-Physical Systems* (1re éd., p. 259-329). CRC Press. (McConnell et al., 2021, pp. 274-275): <https://doi.org/10.1201/9781003186380-18>

161 Ibid. p. 276

162 *Technology Drivers in Windfarm Asset Management* (Barnes et al., 2018): <https://researchportal.hw.ac.uk/en/publications/technology-drivers-in-windfarm-asset-management>



Étapes de développement

Des laboratoires universitaires et industriels s'intéressent de plus en plus aux technologies robotiques pour le secteur des énergies renouvelables. La sécurité et la prise de décision des systèmes robotisés sont au cœur de plusieurs de ces initiatives. Plusieurs défis sont encore à relever dans la recherche, notamment en matière d'IA, de *Machine Learning* (ML), d'*Underwater Autonomous Vessel* (AUV), d'*Autonomous Surface Vessel* (ASV), de *Beyond Visual Line of Sight* (BVLOS) et de *Digital Twin* (DT). Les avancées dans ces secteurs de recherche et l'intégration de ces technologies dans une architecture de système unique constituent la *Symbiotic System of Systems approach* (SSOSA), une approche qui rendra possible le déploiement de *flottes multi-robots résidentes entièrement autonomes*¹⁶³.

L'adoption de la SSOSA pour le secteur éolien offshore s'impose progressivement à la fois par la priorisation de l'éolien comme source d'énergie renouvelable par le gouvernement britannique et par la création de filières technologiques innovatrices d'abord orientées vers l'industrie du pétrole et du gaz et l'industrie aérospatiale¹⁶⁴. L'orientation de ces filières technologiques vers l'éolien offshore se fait donc conjointement aux projets des autres secteurs industriels.

Les principaux éléments des développements associés aux parcs éoliens en haute mer se déroulent en plusieurs étapes.

- Avec le *UK renewable energy roadmap* de 2011, le gouvernement britannique reconnaît le potentiel important des parcs éoliens en mer pour l'atteinte de ses objectifs de développement des énergies renouvelables. Le rapport insiste sur la nécessité d'investir dans l'innovation technologique afin de réduire les coûts d'installation, d'opération et de maintenance pour atteindre les cibles d'expansion des parcs éoliens¹⁶⁵.
- Mise sur pied d'*ORE Catapult* en 2013¹⁶⁶. *ORE* (Offshore Renewable Energy) fait partie du réseau *Catapult* mis en place par *Innovate UK* pour accélérer la création et la croissance d'entreprises britanniques dans des secteurs à forte croissance. Les projets d'*ORE Catapult* concernent l'éolien offshore en général, mais une partie importante des efforts se concentre sur le développement et l'intégration de la robotique et des systèmes autonomes¹⁶⁷.
- *The Industrial Digitisation Review* mis en place dans le *UK Industrial Strategy Green Paper* de 2017 rapporte qu'une variété de nouvelles technologies présentent un potentiel important pour l'éolien en mer. Ces technologies incluent les données massives (*big data*) et les capteurs, la simulation numérique avancée et l'utilisation de la robotique pour les services d'exploitation et de maintenance¹⁶⁸.

163 A review: *Challenges and opportunities for artificial intelligence and robotics in the offshore wind sector* (Mitchell et al., 2022, pp. 29-30): <https://www.sciencedirect.com/science/article/pii/S266654682200088?via%3Dihub>

164 Global Expert Mission—USA Robotics & AI in Extreme Environments 2019: https://iuk.ktn-uk.org/wp-content/uploads/2020/12/B2_11_KTN_USA-Robotics--AI-in-Extreme-Environments.pdf

165 *UK renewable energy roadmap: 2011* - GOV.UK, 2011, pp. 42-57 : <https://www.gov.uk/government/publications/renewable-energy-roadmap>

166 home page - ORE, s. d.: <https://ore.catapult.org.uk/>

167 CATAPULT - Robotics & Autonomous Systems, s. d.: <https://ore.catapult.org.uk/what-we-do/innovation/robotics-and-autonomous-systems/>

168 *Made Smarter Review* - GOV.UK, 2017: <https://www.gov.uk/government/publications/made-smarter-review>

4. Cas d'application de l'utilisation de l'IA dans le secteur public

- Création de *HOME Offshore* en 2017¹⁶⁹, financé par l'*Engineering and Physical Sciences Research Council (EPSRC)* du Royaume-Uni. Il s'agit d'un partenariat avec cinq grandes universités britanniques portant sur la recherche et le développement de techniques de détection avancées, la robotique, les modèles de réalité virtuelle et l'intelligence artificielle, afin de réduire les coûts d'opération et d'entretien des parcs éoliens en haute mer. La mission du projet *HOME Offshore* est d'intégrer les éléments suivants dans un cadre unique afin de réduire les coûts d'exploitation tout en améliorant l'efficacité des cycles de maintenance¹⁷⁰ :
 - Multi-physics model of the dynamics of OWFs, where the novelty lies in linking the aerodynamics, hydrodynamics, control dynamics, and structural dynamics with the thermo and electro-magnetic dynamics aspects.
 - Drone technology with improved run-time duration and levels of sophistication.
 - High-fidelity data streams from advanced sensing techniques, which are low cost.
 - Post processing data analytics of large volumes of data or 'big data'.
 - Emergence of improved robotic platforms and capabilities in the form of improved robot manipulation and self-certification.
- Le projet *Multi-Platform Inspection, Maintenance and Repair in Extreme Environments (MIMREE)*, financé par *Innovate UK* et associé à *ORE Catapult*, est un projet d'innovation intersectoriel dont l'objectif est de révolutionner les opérations éoliennes offshore en utilisant la robotique et l'IA (RAI). Le projet a débuté en 2019 et vise à démontrer comment les bateaux mères autonomes (autonomous motherships) et les équipages robotisés peuvent être compétents dans la planification et l'exécution des opérations et de l'entretien telles que les missions d'inspection et de réparation, sans nécessiter de présence humaine en mer. La réalisation de ces objectifs conduirait à la première plateforme robotisée d'inspection et de réparation entièrement autonome au monde et *ORE Catapult* a évalué que cela pourrait permettre d'économiser 26 millions de livres sterling sur la durée de vie d'un parc éolien en haute mer¹⁷¹.
- The *UK Robotics and Artificial Intelligence Hub for Offshore Robotics for Certification of Assets (ORCA Hub)*, établi en 2017 et financée par l'*Engineering and Physical Sciences Research Council (EPSRC)*, est l'une des plus importantes collaborations universitaires et industrielles britanniques dans le domaine de la robotique en environnements extrêmes. Le *ORCA Hub* est un projet associant plusieurs organisations¹⁷² dont l'objectif principal est d'utiliser des équipes de robots et des systèmes intelligents autonomes (AIS) sur des parcs de production énergétique éloignés afin de réaliser des opérations plus rentables, sûres et efficaces.

La création d'une plateforme offshore entièrement autonome, gérée et inspectée depuis le continent, est l'objectif ultime de l'*ORCA Hub*. Le système de systèmes au centre du projet du *ORCA Hub* comprend la plateforme en haute mer, les actifs connexes, les systèmes de surveillance et des systèmes robotiques divers (robots aériens, de surface et sous-marins) nécessaires pour assurer le fonctionnement autonome et continu de la plateforme¹⁷³.

169 HOME-Offshore: *Holistic Operation and Maintenance for Energy from Offshore Wind Farms/Details of Grant*, s. d.: <https://gow.epsrc.ukri.org/NGBOViewGrant.aspx?GrantRef=EP/P009743/1>

170 A review: *Challenges and opportunities for artificial intelligence and robotics in the offshore wind sector* (Mitchell et al., 2022, p. 20): <https://www.sciencedirect.com/science/article/pii/S2666546822000088?via%3Dihub>

171 Ibid.

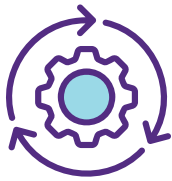
172 A consortium of five universities, working with 35 industrial and innovation partners. The largest academic centre in the world for research into robotics technology for offshore energy infrastructure. Source : *Offshore energy infrastructure research unveiled* - Energy Sustainability Solutions, 2019: <https://essmag.co.uk/offshore-energy-infrastructure-research-unveiled/>

173 *The Future Workplace: A Symbiotic System of Systems Environment*. Dans T. Semwal & F. Iqbal, *Cyber-Physical Systems* (1re éd., p. 259-329). CRC Press. (McConnell et al., 2021, pp. 274-275): <https://doi.org/10.1201/9781003186380-18>

Un des volets importants de la recherche effectuée par le *ORCA Hub* concerne l'auto-certification (*self-certification*) à la fois des systèmes composant le parc éolien dans son ensemble et des équipements robotisés spécifiquement attirés aux opérations et à la maintenance du parc. Dans le contexte de ces recherches, les plateformes robotisées sont considérées comme faisant partie intégrante du *système de systèmes* plus large qu'est le parc éolien en haute mer. L'auto-certification constitue une étape primordiale à l'établissement de la SSOSA puisque chaque élément d'un système de systèmes doit être à même d'établir son propre état pour évaluer sa propension à effectuer sa part du travail dans une mission multi-systèmes¹⁷⁴.

Un autre élément clé de la mise en œuvre de systèmes robotiques dans les zones reculées est la manière dont ils interagissent avec les humains qui les utilisent à la fois à distance et sur le terrain. Ce domaine est connu sous le nom de *Human-Robot Interaction* (HRI). Au sein du *ORCA Hub*, le système MIRIAM (Multimodal Intelligent Interaction for Autonomous systems) a été créé pour permettre aux opérateurs d'interroger les robots en temps réel en utilisant le langage naturel. La confiance entre l'humain et le robot est essentielle lorsqu'il s'agit d'accomplir des tâches complexes et dangereuses. La principale méthode consiste donc à utiliser des interactions en langage naturel pour que les décisions du robot soient communiquées de manière transparente et compréhensible¹⁷⁵.

Il a été annoncé en 2022 que la technologie appliquée et les partenariats industriels développés par le *ORCA Hub* se poursuivraient au *National Robotarium*, le centre britannique pour la robotique et l'intelligence artificielle¹⁷⁶. Située à Édimbourg, l'installation dispose de technologies et de laboratoires à la fine pointe pour développer et tester des solutions robotiques et d'intelligence artificielle dans trois domaines distincts : la robotique et les systèmes autonomes (RAS), l'interaction humain-robot (HRI) et les applications laser de précision (PLA).



Fonctionnement du système

L'exploitation et l'entretien automatisé d'un parc éolien offshore requièrent une combinaison de capteurs déployés en permanence et d'autres méthodes d'inspection pouvant être utilisées périodiquement et rapidement en cas d'urgence. Le capteur *Limpet*¹⁷⁷ est un exemple d'une plateforme multi-capteurs créée pour être déployée par des humains ou des robots partout où cela est nécessaire. La manière de déployer et d'utiliser ces capteurs dans des environnements extrêmes est un élément clé de la mission d'*ORCA Hub*. Plusieurs robots peuvent être utilisés : des robots à roues ou des robots quadrupèdes pour les besoins terrestres, des drones pour les inspections aériennes et divers véhicules sous-marins autonomes (Autonomous Underwater Vessel/AUV) pour les inspections sous-marines¹⁷⁸.

174 *ORCA Robotics - Robot and Asset Self-Certification*, s. d.: <https://orcahub.org/innovation/focus-areas/robot-and-asset-self-certification>

175 *The Future Workplace: A Symbiotic System of Systems Environment*. Dans T. Semwal & F. Iqbal, *Cyber-Physical Systems* (1re éd., p. 259-329). CRC Press. <https://doi.org/10.1201/9781003186380-18> (McConnell et al., 2021, p. 275)

176 *Applied research by ORCA Hub to continue at The National Robotarium* - YouTube, s. d.: <https://www.youtube.com/watch?v=dkH0RjbO4Ho>; *Edinburgh's Orca Hub secures £2.5m to power up robotics research* | The Scotsman, s. d.: <https://www.scotsman.com/business/edinburghs-orca-hub-secures-ps25m-to-power-up-robotics-research-3275939>; *Research - The National Robotarium*, s. d.: <https://thenationalrobotarium.com/research/>

177 *The Limpet: A ROS-Enabled Multi-Sensing Platform for the ORCA Hub* (M. Sayed et al., 2018): <http://www.mdpi.com/1424-8220/18/10/3487>; *Limpet II: A Modular, Untethered Soft Robot* (M. E. Sayed et al., 2021): <https://www.liebertpub.com/doi/10.1089/soro.2019.0161>

178 *The Future Workplace: A Symbiotic System of Systems Environment*. Dans T. Semwal & F. Iqbal, *Cyber-Physical Systems* (1re éd., pp. 259-329). CRC Press. (McConnell et al., 2021, p. 275): <https://doi.org/10.1201/9781003186380-18>; *A review: Challenges and opportunities for artificial intelligence and robotics in the offshore wind sector* (Mitchell et al., 2022, pp. 13-16): <https://www.sciencedirect.com/science/article/pii/S2666546822000088?via%3DIihub>

4. Cas d'application de l'utilisation de l'IA dans le secteur public

Pour éviter que les robots aient à être entreposés puis déployés manuellement par des humains, on envisage des *robots résidents* présents en permanence à proximité des éoliennes dans les parcs éoliens. Ces robots peuvent être autonomes, mais ils peuvent également être commandés à distance depuis la côte. Seule une interaction humaine minimale avec le parc éolien sera nécessaire, sauf si une action corrective ne peut être menée à bien par un robot ou une équipe de robots.

La SSOSA permettra l'interconnexion des flottes de robots multiples afin de permettre aux robots de tirer parti de leurs capacités respectives pour assurer l'exécution des missions de façon autonome. Ces systèmes permettront une vision et une approche holistiques des problèmes qui se posent dans un parc éolien en haute mer. L'infrastructure numérique assurera le déploiement rapide de *flottes robotiques résidentes* interconnectées par le biais d'une architecture numérique symbiotique, ce qui permettra le partage des connaissances. L'un des défis du déploiement futur de la RAI est de s'assurer que les flottes multi-robots ont la capacité de collaborer, coopérer et corroborer (C3) afin d'établir une *gouvernance collaborative* (*gouvernance C3*) entre l'infrastructure, les personnes, les systèmes autonomes et l'environnement¹⁷⁹ :

A SSOSA represents the next steps required to seamlessly integrate human-machine partnerships and multiple robotic platforms for mutualistic collaboration. This is reinforced by bidirectional communications, encouraging knowledge sharing across cyber physical systems. [...] AI supports the improvement of accuracy and decision making, wherein a system can make its own decisions relative to the scenario or environment. Improved RAI supports an increase in automated BVLOS [Beyond Visual Line of Sight] missions wherein a robot can be assigned a mission and deployed remotely by an operator. The combination of RAI and a SSOSA leads to advanced decision making for autonomous systems wherein a robot can 'Adapt and Thrive' when faced with unforeseen challenges. This provides an improved operational overview for a remote operator due to increased accessibility of deployed robot properties such as mission status, inspection status and data collected all within a highly accessible DT [Digital Twin] which allows for human readable data¹⁸⁰.

En tant que système de systèmes, la SSOSA fait appel à un large éventail de technologies et capacités liées à l'IA : robotique, vision par ordinateur, apprentissage automatique, interaction humain-robot et robot-robot, détection et perception, manipulation robotique et algorithmes pour la planification et la prise de décision. L'amélioration des performances des différentes composantes technologiques intégrées à la SSOSA permet d'envisager un niveau d'automatisation avancé des plateformes offshore grâce à des systèmes robotisés autonomes ayant accès à une *carte de connaissances distribuées* (map of distributed knowledge). Outre l'automatisation des opérations et de la maintenance, cela permet au système de faire des recommandations aux humains qui demeurent dans la boucle (human in the loop) concernant les nouvelles priorités de missions à objectifs multiples. Ces nouvelles informations/recommandations sont aussi partagées par les différents CPS (cyber-physical systems) de manière autonome, de façon à être intégrées aux opérations au bénéfice de l'ensemble du système de systèmes¹⁸¹.

L'objectif est ainsi le développement d'une *SSOSA prospère* où une plateforme robotique autonome peut surmonter des événements imprévus nécessitant une déviation de l'objectif initial de sa mission, de façon à réaliser un profil de mission optimisé. Un tel système pourra suggérer de manière autonome des solutions aux problèmes et menaces qu'il détecte afin d'assurer la permanence et la continuité de sa mission.

179 Ibid. p. 28 ; *The Future Workplace : A Symbiotic System of Systems Environment*. Dans T. Semwal & F. Iqbal, *Cyber-Physical Systems* (1re éd., p. 259-329). CRC Press. (McConnell et al., 2021, p. 280): <https://doi.org/10.1201/9781003186380-18>

180 *A review: Challenges and opportunities for artificial intelligence and robotics in the offshore wind sector* (Mitchell et al., 2022, p. 30): <https://www.sciencedirect.com/science/article/pii/S2666546822000088?via%3Dihub>

181 *The Future Workplace: A Symbiotic System of Systems Environment*. Dans T. Semwal & F. Iqbal, *Cyber-Physical Systems* (1re éd., p. 259-329). CRC Press. (McConnell et al., 2021, p. 283): <https://doi.org/10.1201/9781003186380-18>

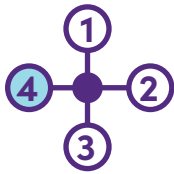


Bénéficiaire direct du système

L'automatisation de l'opération et la maintenance des parcs éoliens offshore bénéficie directement aux opérateurs de ces infrastructures (organisations), ainsi qu'aux employés responsables des tâches d'entretien et de réparation.

En tant qu'organisations, les opérateurs de parc éolien en haute mer bénéficient de modes opératoires plus efficaces, efficaces et écologiques, en plus d'avoir une vision d'ensemble, en temps réel, de l'état d'un parc éolien donné, sans avoir à quitter la côte. La capacité prédictive du système permet d'anticiper l'usure et de prévoir des travaux d'entretien avant que des bris ne se produisent afin de prolonger la durée de vie de l'actif.

Les employés bénéficient directement du système qui diminue ou élimine leur exposition aux dangers associés au travail sur les OWFs. L'information que fournit le système sur l'état de l'infrastructure leur permet d'effectuer avec plus de précision et d'efficacité leur travail d'entretien à partir de la côte. Ceci contribue également à améliorer la qualité de vie des employés qui n'ont plus à se rendre sur les parcs offshore pour de longues périodes (rappelez-vous que certains OWFs sont à plus de 200 km de la côte). De plus, lorsqu'une réparation ou un entretien nécessite la présence d'employés sur un parc éolien offshore, ceux-ci peuvent compter sur l'aide de la plateforme robotique en place, ainsi que sur l'information globale générée par le système.



Degré d'autonomie du système

Le degré d'autonomie du système dépendra de la forme concrète qu'il prendra, des technologies qui seront utilisées et de la manière dont elles seront utilisées. Bien qu'indéterminé pour l'instant, les projets mis de l'avant par les parties prenantes pointent vers un degré d'autonomie dépassant notre description du système semi-autonome. En effet, si l'on considère la tâche du système (l'autonomie quasi complète de l'opération et la maintenance de parcs éoliens offshore) ce système de systèmes pourra effectuer des missions d'entretien des actifs offshore en interagissant en temps réel avec l'*environnement* concerné par cette tâche et en anticipant le statut et les besoins futurs des infrastructures des parcs éoliens et des actifs robotiques résidents responsables des opérations et de l'entretien. La capacité pour le système de déroger de la mission établie (à la suite d'un constat autonome sur le changement des circonstances en temps réel) et de la replanifier afin d'en atteindre tout de même les objectifs, nous semble indiquer un niveau d'autonomie décisionnelle particulièrement élevé. Si l'ensemble du projet est réalisé conformément aux ambitions décrites, il s'agira d'un système de degré d'autonomie 4 (autonome).



Pistes de réflexions éthiques

Bénéfices et avantages du SIA

- Gestion autonome du fonctionnement et de l'entretien des parcs éoliens en mer. La SSOSA prévoit le déploiement de flottes multi-robots entièrement autonomes. Ce déploiement permettrait d'améliorer l'efficacité des opérations d'entretien des parcs éoliens offshore et de réduire le temps de réaction pour répondre à un bris ou une urgence.
- Sécurité des travailleurs. La présence humaine sur les plateformes éoliennes offshore serait réduite, minimisant l'exposition des employés aux dangers en mer. Les employés pourraient accomplir leur travail depuis la côte en supervisant ou contrôlant, à distance, certains équipements offshore.
- Diminution des impacts environnementaux. En minimisant la présence humaine sur les sites d'éoliennes offshore, le système peut potentiellement réduire l'impact environnemental du transport des employés, du matériel et des biens nécessaires pour permettre leur présence sur ces sites pendant la durée de leur séjour.

Inconvénients et risques du SIA

- Perte d'emplois. Les travailleurs actuellement chargés de l'entretien des parcs éoliens offshore risquent de perdre leur emploi à cause de l'automatisation. L'expertise et l'expérience des employés actuels ne sont pas directement transposables à un travail de supervision ou de contrôle à distance.
- Sécurité des travailleurs. Si le nombre d'employés humains sur les sites diminue avec l'automatisation, la sécurité des travailleurs restants pourrait être négligée. D'autant plus que ces travailleurs s'inséreront dans la plateforme automatisée ce qui présente de nouveaux risques pour leur sécurité.
- Impact environnemental. Il n'est pas clair que les bénéfices environnementaux de l'automatisation l'emportent sur les coûts environnementaux associés à la création et à la maintenance des robots et de l'infrastructure informatique nécessaire (utilisation de batteries, bande passante, serveurs informatiques, déchets électroniques, etc.).
- Risques liés à l'autonomie élevée du système.
 - Avec un tel degré d'autonomie, des problèmes d'alignement de l'IA avec les valeurs humaines pourraient survenir. Par exemple, le système pourrait devoir choisir entre assurer la sécurité d'un employé humain et préserver l'intégrité de la plateforme elle-même. Quel sera le choix du système dans sa version la plus autonome?
 - Une trop grande dépendance au système pourrait conduire à une perte de contrôle des actifs offshore due à des bogues techniques, des conditions climatiques extrêmes ou d'autres incidents divers. S'il y a dysfonctionnement ou défaillance du système, il peut ne pas y avoir de solutions de rechange adéquates pour garantir la sécurité des employés et la continuité de l'approvisionnement énergétique.
 - La nature autonome et interconnectée de la plateforme robotique au cœur de ce système le rend particulièrement vulnérable à des cyberattaques qui pourraient perturber le fonctionnement des parcs éoliens et menacer l'approvisionnement énergétique du pays.

Valeurs à prendre en considération pour le déploiement et l'utilisation de ce SIA (ou d'un système similaire)

- Alignement du SIA avec les valeurs humaines (AI Alignment) et maintien de l'humain dans la boucle décisionnelle. Il sera important de maintenir un contrôle humain sur le système pour s'assurer que les décisions prises par l'IA seront en accord avec les valeurs et les priorités humaines. Il convient, par exemple, de s'assurer que la sécurité des personnes ne devienne pas secondaire face à l'objectif plus large de la pérennité du parc éolien.
- Bénéfices environnementaux. L'un des objectifs du déploiement de la SSOSA est de réduire l'impact environnemental des opérations éoliennes. Cependant, il est important de quantifier correctement les bénéfices environnementaux, en prenant en compte tous les facteurs, y compris l'empreinte environnementale des robots, des serveurs et des batteries.
- Continuité du service public. Les infrastructures éoliennes offshore étant d'importance stratégique, elles doivent être protégées contre les vulnérabilités et défaillances potentielles des systèmes autogérés. Les organisations responsables de leur fonctionnement doivent être à même de continuer à remplir leur mission (approvisionnement énergétique de la population) même si le système autonome n'est plus disponible.
- Impact sur l'emploi. Le passage à un système largement autonome peut entraîner une perte d'emploi pour ceux qui travaillent actuellement sur les plateformes éoliennes. Il est important de considérer comment reconvertir ces employés ou intégrer leur expertise dans le nouveau système.
- Sécurité des employés. La sécurité des employés qui seront toujours nécessaires sur les plateformes éoliennes offshore est primordiale.

Comment protéger ou restaurer les valeurs affectées par le SIA?

- Sécurité des travailleurs. La SSOSA devrait être conçue avec des dispositifs de sécurité pour protéger les travailleurs sur le site.
- Impact environnemental. Il est important de surveiller de près l'impact environnemental global de la SSOSA. Une analyse d'impact environnemental comparative entre la SSOSA proposée et le fonctionnement actuel serait utile pour disposer d'une information précise.
- Protection des infrastructures. La mise en place de normes de sécurité élevées est suggérée pour protéger les infrastructures essentielles contre les menaces extérieures et les dérives potentielles de la SSOSA. Concernant les cyberattaques il est suggéré que les communications entre les robots, le système informatique et les humains sur la côte soient sécurisées au maximum.
- Anticipation d'une défaillance du système. Des solutions alternatives doivent être prévues en cas de dysfonctionnement ou de défaillance du système autonome. Les employés et les organisations doivent être préparés à accomplir leurs tâches sans l'assistance du SIA afin d'assurer la continuité des services à la population.

Autres ressources

Towards the future of the energy industry with intelligent robots | AI FOR GOOD WEBINARS - YouTube, 2022: <https://www.youtube.com/watch?v=wS87eGyPI1c>

Robots primed to revolutionise offshore energy industry safety, 2021: <https://www.energyvoice.com/oilandgas/north-sea/332959/robots-revolutionise-offshore-energy-industry-safety/>

Green energy for more than one million UK homes: Siemens Gamesa to supply flagship turbines for world's second-largest offshore wind power plant, 2023: <https://www.siemens-gamesa.com/newsroom/2023/03/032923-siemens-gamesa-press-release-offshore-wind-power-plant-anglia-united-kingdom>

Unique underwater drone carries out first ever autonomous inspection of wind-farms - The National Robotarium, 2022: <https://thenationalrobotarium.com/unique-underwater-drone-carries-out-first-ever-autonomous-inspection-of-windfarms/>

Robotics and Artificial Intelligence Hubs | Autonomy and Verification Network, s. d.: <https://autonomy-and-verification.github.io/projects/rai-hubs> (liens vers une grande quantité d'articles techniques)

Papers-ORI/Oxford Robotics Institute, s. d.: <https://www.ori.ox.ac.uk/publications/papers/>

5. Conclusion

5. Conclusion

Le déploiement et l'utilisation de l'intelligence artificielle dans le secteur public sont en plein essor au Québec et ailleurs dans le monde. De plus en plus d'organisations publiques utilisent des solutions d'IA ou se questionnent sur la manière de recourir à cette technologie. La tendance devrait s'accélérer au cours des prochaines années, car l'IA offre de nombreux avantages aux organisations cherchant à améliorer l'efficacité et la qualité des services aux citoyens. La variété des cas d'utilisation présentés dans ce rapport démontre que tous les domaines d'action publique peuvent potentiellement tirer profit des technologies d'IA.

Les impacts de l'IA sont variables et souvent difficiles à anticiper. Comme nous l'observons dans les cas d'application étudiés, les bénéfices et les inconvénients associés aux SIA dépendent du contexte et des conditions dans lesquelles ils sont développés et intégrés au sein des organisations publiques. Par conséquent, l'encadrement de leur déploiement et utilisation devra tenir compte de ces facteurs contextuels (Dean et al., 2021).

De nombreux exemples présentés dans ce rapport mettent en évidence le caractère évolutif des SIA dont les performances s'améliorent au rythme des progrès technologiques. Ces évolutions peuvent également affecter le fonctionnement des SIA et les actions qu'ils entreprennent comme nous l'avons vu avec l'extension des types de délits pris en compte par *SKALA*, ou avec les fonctionnalités additionnelles en développement dans les cas d'*ABLE* et *Surtrac*. Puisque les SIA tendent à évoluer rapidement, des usages non anticipés peuvent apparaître, ou des usages a priori anodins peuvent ensuite devenir problématiques. Comme l'avance *AlgorithmWatch (Artificial Intelligence Act, 2023)*, l'applicabilité de règles spécifiques ne devrait pas dépendre du type de technologie d'IA utilisée, mais de l'impact que la technologie a sur les individus et la société. Ceci implique la nécessité d'une évaluation continue des impacts du SIA à chaque étape de son cycle de vie.

Cette recommandation a été partagée au cours des ateliers de réflexion éthique portant sur les cas d'application que nous avons documentés. Ces ateliers regroupaient des intervenants de domaines et d'expertises divers. D'une manière générale, les participants aux ateliers reconnaissent des avantages incontestables aux différents SIA, tels que l'optimisation des processus, l'amélioration de l'efficacité et de la prestation des services aux citoyens. Les participants mettent cependant en garde les organisations publiques qui souhaitent déployer et utiliser l'IA dans leurs services.

Un consensus s'établit par exemple sur la nécessité d'avoir accès à l'information permettant de comprendre et d'évaluer les différents SIA, en particulier lorsqu'ils sont utilisés dans le secteur public. Cela implique un degré suffisant d'explicabilité et de transparence quant au fonctionnement de chaque système et en ce qui concerne l'utilisation des données, en particulier celles relatives à la vie privée des citoyens. Cette préoccupation centrale est à la source de toute possibilité d'amenuiser ou de prévenir les répercussions potentiellement négatives de cette technologie novatrice qui, comme les cas présentés l'indiquent, se manifeste sous différentes formes. De plus, les experts ont maintes fois répété que la possibilité de comprendre et d'analyser un SIA donné doit se maintenir dans le temps, à travers les différentes étapes de conception, de déploiement et d'utilisation du système.

Un autre thème récurrent est le maintien de l'autonomie, tant pour les individus (citoyens, employés) que pour les organisations. Dans un monde où les systèmes automatisés jouent un rôle croissant dans la prise de décision, la nécessité de conserver un *humain dans la boucle* est une recommandation récurrente des participants aux groupes de discussion. Le maintien de l'autonomie humaine est par ailleurs crucial pour assurer la continuité des services publics en cas de dysfonctionnement ou de cyberattaque, un autre défi régulièrement évoqué lors des ateliers de réflexion éthique.

Plusieurs participants se sont également inquiétés du rôle joué par les entreprises privées dans le développement ou le fonctionnement de SIA utilisés dans le secteur public. En effet, les processus d'acquisition, de conception et de déploiement de systèmes d'IA impliquent souvent une variété d'acteurs issus des domaines public et privé, ce qui les rend complexes, divers et difficiles à appréhender. L'intrication de différents fournisseurs de composantes et de services, que ce soit pour le *software*, le *hardware*, pour les différents serveurs virtuels – entreposage des données et leur traitement (calcul) – ou pour l'entretien et l'optimisation, fait de chaque instance de déploiement de SIA un processus unique. Il est toutefois essentiel de noter que pour les systèmes conçus à l'interne par des organismes publics comme pour les systèmes acquis auprès d'entreprises privées, la nature connectée des SIA fait en sorte qu'ils reposent tous, à une étape ou une autre, sur une infrastructure numérique impliquant un nombre limité d'acteurs privés (Amazon Web Services, Microsoft Azure, Google Cloud, etc.). L'essor et la démocratisation de l'utilisation de l'IA dans les organisations publiques laissent envisager une situation similaire en ce qui a trait à l'offre des solutions d'IA en général. Il faut donc s'attendre à une concentration de l'offre de services auprès de quelques grands acteurs privés, à la fois pour des raisons d'efficacité économique et de marchandisation, que pour assurer la sécurité et la stabilité de ces technologies (World Economic Forum, 2023 ; Sanchez-Graells, 2022).

Portée et limitations de la recherche

Notre recherche avait pour objectif d'explorer les implications de l'intégration de l'IA dans des organisations publiques ayant des similarités avec les ministères et organismes du gouvernement du Québec. Pour parvenir à cet objectif, nous avons documenté des cas concrets de déploiement de solutions d'IA dans des organisations publiques de pays de l'OCDE. Les organisations concernées opéraient à différents niveaux gouvernementaux (municipal, régional et national) et dans divers secteurs de politique publique (finance, infrastructure, ressources humaines, transport, etc.). Les ateliers de discussion regroupant des spécialistes de disciplines diverses nous ont permis d'amorcer une réflexion sur les enjeux éthiques soulevés par chacun des cas.

La documentation des cas a été effectuée en utilisant des informations publiquement disponibles, ce qui a limité la profondeur de notre analyse en fonction de la transparence et de la publicité de chaque initiative. Certains cas étaient bien documentés grâce à des articles scientifiques et des détails techniques, tandis que d'autres dépendaient des informations fournies par les entreprises fournisseuses de solutions d'IA ou les instances gouvernementales responsables du déploiement des systèmes. Pour ces derniers, nous avons porté une attention particulière à la distinction entre l'information factuelle et celle liée à l'activité promotionnelle. Malgré ces limitations, quantité d'autres exemples auraient pu être documentés. La sélection des cas s'est basée sur la pertinence de chacun et sur la variété de l'ensemble et relève, dans une large part, de la subjectivité des auteurs. Cette recherche n'avait pas la prétention d'être exhaustive, mais plutôt exploratoire.

Dans ce rapport, nous avons opté pour une présentation la plus objective possible de chacun des cas, laissant le soin aux spécialistes en éthique de l'IA, participants aux ateliers que nous avons organisés, d'analyser plus en détail les avantages, inconvénients, valeurs et préoccupations soulevés par les différents cas d'application. Les SIA s'immiscent dans les organisations qui les déploient et les utilisent, et il est parfois difficile d'isoler complètement les impacts de ces technologies des impacts inhérents à la structure organisationnelle préexistante. En ce sens, la question du partage de pouvoir discrétionnaire entre humains et système d'IA nous apparaît prometteuse pour approfondir la dimension relationnelle de l'introduction d'artefacts techniques dans des organisations publiques¹⁸².

182 Voir, entre autres: *Managing Algorithms for Public Value* (Selten & Meijer, 2021) ; *Administration by Algorithm? Public Management meets Public Sector Machine Learning* (Veale & Brass, 2020) ; *Artificial Intelligence, Discretion, and Bureaucracy* (Bullock, 2019) ; *Discretion and the Quest for Controlled Freedom* (Tony Evans & Peter Hupe, 2020), en particulier le chapitre 20 (*Automated Discretion*) ; *The agency of algorithms: Understanding human-algorithm interaction in administrative decision-making* (Peeters, 2020) ; *From Street-Level to System-Level Bureaucracies: How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control* (Bovens & Zouridis, 2002) ; *Artificial Discretion as a Tool of Governance: A Framework for Understanding the Impact of Artificial Intelligence on Public Administration* (Young et al., 2019).

Le déploiement et l'utilisation des solutions d'IA constituent des dynamiques d'innovation continues nourries par les avancées technologiques, mais aussi par des facteurs contextuels tels que le financement et la proactivité de la recherche, la transformation des organisations publiques, l'acceptabilité sociale et la rétroaction des utilisateurs. Les systèmes, une fois déployés, tendent en effet à évoluer en termes de capacité, d'efficacité et d'impacts, au gré de leur usage et de leur popularité. Cette mutation continue, résultat de l'introduction de systèmes aux caractéristiques techniques évolutives dans des contextes organisationnels adaptatifs, a constitué un défi à la documentation des cas. La présentation que nous avons faite de chacun d'eux n'est en cela qu'une « photo » d'un moment précis d'un processus de déploiement de systèmes sociotechniques en transformation.

Les trois ateliers de discussion, regroupant chacun de 5 à 7 experts de différents domaines, ont permis d'identifier les préoccupations éthiques à prendre en compte dans l'éventualité du déploiement, au Québec, de SIA similaire aux cas présents à l'échelle internationale. Il est toutefois important de noter que les enjeux identifiés lors de ces ateliers sont eux aussi limités par l'information rendue disponible aux participants à travers notre présentation des cas et qu'un travail d'analyse plus approfondie sera nécessaire pour obtenir une vue plus complète des enjeux relatifs à chacune des applications.

Perspectives pour de futures recherches

Un approfondissement des cas présentés, ou d'autres cas pertinents, en documentant plus avant leur cheminement organisationnel et social, ainsi que les contextes culturels et réglementaires spécifiques à leur juridiction, nous semble une approche prometteuse afin d'éclairer davantage les différentes mécaniques d'implémentation de SIA dans le secteur public. Un tel approfondissement, en enrichissant la contextualisation, permettrait également d'identifier les éléments qui méritent davantage d'attention, d'approfondir la réflexion éthique et de proposer des recommandations plus spécifiques à chacune des applications d'IA. Parallèlement à cet approfondissement, une analyse des mécanismes de gouvernance existants permettrait d'identifier les forces, les faiblesses et les lacunes de différentes approches contribuant à l'établissement, au Québec, d'une gouvernance efficace, proactive et sécuritaire de ces technologies émergentes potentiellement prometteuses.

À l'image des travaux législatifs récents du Parlement européen (Artificial Intelligence Act, 2023), les discussions entamées devraient se poursuivre à différents niveaux afin de définir et clarifier les termes, les risques et les potentiels de l'utilisation de l'IA dans le secteur public. Notre exploration de cas et l'amorce de réflexion éthique proposée dans ce rapport nous indiquent que cette discussion doit être intégrée à chaque étape du déploiement de systèmes d'IA.

L'intégration de l'IA dans les administrations publiques est une tendance croissante avec des implications significatives. Notre recherche a permis d'explorer divers cas de déploiement de ces technologies et d'amorcer une réflexion éthique sur ces enjeux. D'autres recherches devraient approfondir notre compréhension des implications du développement de l'IA dans les organisations publiques québécoises, notamment à travers la perspective des usagers des services publics, laquelle a été brièvement évoquée dans cette étude. Il ne faut pas perdre de vue que l'IA représente un défi sociétal majeur qui requiert une évaluation de l'acceptabilité sociale de la technologie et de son utilisation grandissante au sein des organisations publiques.

6. Bibliographie

6. Bibliographie

AI as Gatekeepers to the Job Market : A Critical Reading of; Performance, Bias, and Coded Gaze in Recruitment Chatbots. (2021). <https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1572100&dswid=-6817>

AI for Ambulatory Clinics. (s. d.). Consulté 21 avril 2023, à l'adresse <https://www.corti.ai/use-cases/ambulatory-clinics>

AI that detects cardiac arrests during emergency calls will be tested across Europe this summer—The Verge. (2018, avril 25). <https://www.theverge.com/2018/4/25/17278994/ai-cardiac-arrest-corti-emergency-call-response>

AI That Saves Lives : The Chatbot That Can Detect A Heart Attack Using Machine Learning | Bernard Marr. (s. d.). Consulté 21 avril 2023, à l'adresse <https://bernardmarr.com/ai-that-saves-lives-the-chatbot-that-can-detect-a-heart-attack-using-machine-learning/>

AI-X | Tengai—Robot in recruitment processes. (s. d.). Consulté 16 mars 2023, à l'adresse <https://ai-watch.github.io/AI-watch-T6-X/service/90130.html>

Anarkooli, A. J., Persaud, B., Milligan, C., Penner, J., & Saleem, T. (2021). Incorporating Speed in a Traffic Conflict Severity Index to Estimate Left Turn Opposed Crashes at Signalized Intersections. *Transportation Research Record: Journal of the Transportation Research Board*, 2675(5), 214–225. <https://doi.org/10.1177/0361198120986167>

Applied research by ORCA Hub to continue at The National Robotarium—YouTube. (s. d.). Consulté 17 mai 2023, à l'adresse <https://www.youtube.com/watch?v=dkH0RjbO4Ho>

Arowolo, O., Salemme, V., & Suvorov, A. (2022). Towards Whole Health Toxicology : In-Silico Prediction of Diseases Sensitive to Multi-Chemical Exposures. *Toxics*, 10(12), 764. <https://doi.org/10.3390/toxics10120764>

Artificial Intelligence & Emergency services : Saving lives on the phone—EENA. (2020, janvier). <https://eena.org/knowledge-hub/press-releases/artificial-intelligence-eena-corti-project>

Artificial intelligence act. (2023).

Artificial Intelligence and EMS. (2017, janvier).

AuroraAI: Finland's National Artificial Intelligence Program. (2020). <https://interestingengineering.com/innovation/auroraai-finlands-national-artificial-intelligence-program>

Automated bus lane traffic enforcement begins October 7 with new camera technology on MTA buses | Mass Transit. (2019, septembre 24). Mass Transit.

Barnes, M., Brown, K., Carmona, J., Cevasco, D., Collu, M., Crabtree, C., Crowther, W., Djurovic, S., Flynn, D., Green, P. R., Heggo, M., Kababbe, K., Kazemtabrizi, B., Keane, J., Lane, D., Lin, Z., Mawby, P., Mohammed, A., Nenadic, G., ... Watson, S. (2018). *Technology Drivers in Windfarm Asset Management.* <https://doi.org/10.17861/20180718>

Bovens, M., & Zouridis, S. (2002). From Street-Level to System-Level Bureaucracies : How Information and Communication Technology is Transforming Administrative Discretion and Constitutional Control. *Public Administration Review*, 62(2), 174–184. <https://doi.org/10.1111/0033-3352.00168>

Bullock, J. B. (2019). Artificial Intelligence, Discretion, and Bureaucracy. *The American Review of Public Administration*, 49(7), 751–761. <https://doi.org/10.1177/0275074019856123>

Campion, A., Hernandez, M.-G., Mikhaylov Jankin, S., & Esteve, M. (2020). Managing Artificial Intelligence Deployment in the Public Sector. *Computer*, 53(10), 28–37. <https://doi.org/10.1109/MC.2020.2995644>

CATAPULT - Robotics & Autonomous Systems. (s. d.). Consulté 22 mai 2023, à l'adresse <https://ore.catapult.org.uk/what-we-do/innovation/robotics-and-autonomous-systems/e>

Chatila, R., Dignum, V., Fisher, M., Giannotti, F., Morik, K., Russell, S., & Yeung, K. (2021). Trustworthy AI. Dans B. Braunschweig & M. Ghallab (Éds.), *Reflections on Artificial Intelligence for Humanity* (Vol. 12600, p. 13-39). Springer International Publishing. https://doi.org/10.1007/978-3-030-69128-8_2

Coenen-Huther, J. (2003). Le type idéal comme instrument de la recherche sociologique: *Revue française de sociologie*, Vol. 44(3), 531-547. <https://doi.org/10.3917/rfs.443.0531>

Coglianesi, C., & Lehr, D. (2017). Regulating by Robot : Administrative Decision Making in the Machine-Learning Era. *THE GEORGETOWN LAW JOURNAL*, 105.

Conduent, Hayden AI Team Up to Automate Bus Lane Enforcement. (2021, avril 7). Government Technology (Govtech.com). <https://www.govtech.com/biz/conduent-hayden-ai-team-up-to-automate-bus-lane-enforcement.html>

Conduent Transportation and Hayden AI Announce Technology. (2021, avril 6). GlobeNewswire. <https://www.globenewswire.com/news-release/2021/04/06/2205051/0/en/Conduent-Transportation-and-Hayden-AI-Announce-Technology-Partnership-to-Improve-Bus-Lane-Performance-and-Traffic-Safety.html>

Corti · Healthcare DENMARK. (s. d.). Consulté 21 avril 2023, à l'adresse <https://www.healthcaredenmark.dk/toolbox/danish-solutions/corti/>

(Corti Solution Launches in Sweden to Support SOS Alarm Fight the Pandemic | by Corti | Corti | Medium. (2020, juin 16). <https://medium.com/corti-ai/corti-solution-launches-in-sweden-to-support-sos-alarm-fight-covid-19-f5fbd6d4c370>

Corti—AI for Patient Consultations. (s. d.). Consulté 21 avril 2023, à l'adresse <https://www.corti.ai/>

Corti.ai Raises \$27 Million in Series A Funding to Transform Patient Consultations With Artificial Intelligence | Business Wire. (2021, septembre). <https://www.businesswire.com/news/home/20210909005416/en/Corti-ai-Raises-27-Million-in-Series-A-Funding-to-Transform-Patient-Consultations-With-Artificial-Intelligence>

Critch, A. (2020). *AI Research Considerations for Human Existential Safety* (ARCHES).

CTC Networks and Language Models : Prefix Beam Search Explained | by Lasse Borgholt | Corti | Medium. (2018, janvier 8). <https://medium.com/corti-ai/ctc-networks-and-language-models-prefix-beam-search-explained-c11d1ee23306>

Cugurullo, F. (2020). Urban Artificial Intelligence : From Automation to Autonomy in the Smart City. *Frontiers in Sustainable Cities*, 2, 38. <https://doi.org/10.3389/frsc.2020.00038>

Dean, S., Gilbert, T. K., Lambert, N., & Zick, T. (2021). Axes for Sociotechnical Inquiry in AI Research. *IEEE Transactions on Technology and Society*, 2(2), 62-70. <https://doi.org/10.1109/TTS.2021.3074097>

Déclaration de Montréal pour un développement responsable de l'IA. (2018). <https://www.declarationmonreal-iaresponsable.com/>

Defacqz, S., & Dupuy, C. (2021). A transformative change through a coordination process and a steering agency. The case of the financial information system of the French central state. *International Review of Administrative Sciences*, 0020852321105888. <https://doi.org/10.1177/00208523211058859>

Desouza, K. C. (s. d.). *Delivering Artificial Intelligence in Government : Challenges and Opportunities*.

Desouza, K. C., Dawson, G. S., & Chenok, D. (2020). Designing, developing, and deploying artificial intelligence systems : Lessons from and for the public sector. *Business Horizons*, 63(2), 205-213. <https://doi.org/10.1016/j.bushor.2019.11.004>

Dingli, A., Haddod, F., & Klüver, C. (Éds.). (2021). *Artificial Intelligence in Industry 4.0 : A Collection of Innovative Research Case-studies that are Reworking the Way We Look at Industry 4.0 Thanks to Artificial Intelligence* (Vol. 928). Springer International Publishing. <https://doi.org/10.1007/978-3-030-61045-6>

Diyani, R. (2017). Le système d'information Chorus : Développements et perspectives. *Gestion & Finances Publiques*, 1, 102-113. <https://doi.org/10.3166/gfp.2017.00015>

Edinburgh's Orca Hub secures £2.5m to power up robotics research | The Scotsman. (s. d.). Consulté 17 mai 2023, à l'adresse <https://www.scotsman.com/business/edinburghs-orca-hub-secures-ps25m-to-power-up-robotics-research-3275939>

EENA2019—Artificial Intelligence—YouTube. (2019). <https://www.youtube.com/watch?v=V3LCe2tGCac>

Egis in Ireland. (s. d.). Consulté 18 avril 2023, à l'adresse <https://www.egis-group.com/locations/europe-central-asia/ireland>

Etscheid, J. (2019). Artificial Intelligence in Public Administration : A Possible Framework for Partial and Full Automation. Dans I. Lindgren, M. Janssen, H. Lee, A. Polini, M. P. Rodríguez Bolívar, H. J. Scholl, & E. Tambouris (Éds.), *Electronic Government* (Vol. 11685, p. 248-261). Springer International Publishing. https://doi.org/10.1007/978-3-030-27325-5_19

European Commission. Joint Research Centre. (2020). *AI Watch, artificial intelligence in public services : Overview of the use and impact of AI in public services in the EU*. Publications Office. <https://data.europa.eu/doi/10.2760/039619>

Fast Forward : The Plan to Modernize New York City Transit. (s. d.). NYC Transit. Consulté 24 mars 2023, à l'adresse https://static1.squarespace.com/static/5afef986c3c16a2dc6705929/t/5b072571f950b7a5e621a4ff/1527194994914/Fast+Forward+Plan_05-24-2018_3.15PM.pdf

Filer, D. L., Kothiya, P., Setzer, R. W., Judson, R. S., & Martin, M. T. (2017). tcpl : The ToxCast pipeline for high-throughput screening data. *Bioinformatics*, 33(4), 618-620. <https://doi.org/10.1093/bioinformatics/btw680>

First-ever mobile bus lane enforcement solution in New York | Short News | Siemens Mobility USA. (s. d.). Consulté 22 mars 2023, à l'adresse <https://press.siemens.com/global/en/pressrelease/siemens-mobility-launches-first-ever-mobile-bus-lane-enforcement-solution-new-york>

gov.ie—Road Safety Strategy 2013-2020. (s. d.). Consulté 18 avril 2023, à l'adresse <https://www.gov.ie/en/publication/8ed7fd-road-safety-strategy-2013-2020/>

Green energy for more than one million UK homes : Siemens Gamesa to supply flagship turbines for world's second-largest offshore wind power plant. (2023). <https://www.siemensgamesa.com/newsroom/2023/03/032923-siemens-gamesa-press-release-offshore-wind-power-plant-anglia-united-kingdom>

Guglielmi, G. J. (2014). Organisation. Dans *Dictionnaire d'administration publique* (p. 358-359). Presses universitaires de Grenoble; Cairn.info. <https://www.cairn.info/dictionnaire-d-administration-publique--9782706121371-p-358.htm>

Harrison, T. M., & Luna-Reyes, L. F. (2022). Cultivating Trustworthy Artificial Intelligence in Digital Government. *Social Science Computer Review*, 40(2), 494-511. <https://doi.org/10.1177/0894439320980122>

Hastie, H., Lohan, K., Chantler, M., Robb, D. A., Ramamoorthy, S., Petrick, R., Vijayakumar, S., & Lane, D. (2018). *The ORCA Hub : Explainable Offshore Robotics through Intelligent Interfaces* (arXiv:1803.02100). arXiv. <http://arxiv.org/abs/1803.02100>

Having A Heart Attack ? This AI Helps Emergency Dispatchers Find Out. (2018, novembre). <https://www.fastcompany.com/40515740/having-a-heart-attack-this-ai-helps-emergency-dispatchers-find-out>

Hayden AI Company Profile : Valuation & Investors | PitchBook. (s. d.). Consulté 27 mars 2023, à l'adresse <https://pitchbook.com/profiles/company/268302-52>

Hayden AI Granted Patent for Traffic Rule Management Technology. (2022, juin 28). CISION PR Newswire. <https://www.prnewswire.com/news-releases/hayden-ai-granted-patent-for-traffic-rule-management-technology-301576449.html>

Haynes, R. C. (2010). ToxCast on Target : In Vitro Assays and Computer Modeling Show Promise for Screening Chemicals. *Environmental Health Perspectives*, 118(4), a. <https://doi.org/10.1289/ehp.118-a172a>

Health Insurance. (s. d.). Consulté 21 avril 2023, à l'adresse <https://www.corti.ai/use-cases/health-insurance>

Health Tech Hub Copenhagen | Corti. (s. d.). Consulté 21 avril 2023, à l'adresse <https://healthtechhub.org/members/corti/>

Hello, World : Artificial intelligence and its use in the public sector (OECD Working Papers on Public Governance 36; OECD Working Papers on Public Governance, Vol. 36). (2019). <https://doi.org/10.1787/726fd39d-en>

Henman, P. (2020). Improving public services using artificial intelligence : Possibilities, pitfalls, governance. *Asia Pacific Journal of Public Administration*, 42(4), 209–221. <https://doi.org/10.1080/23276665.2020.1816188>

home page—ORE. (s. d.). Consulté 17 mai 2023, à l'adresse <https://ore.catapult.org.uk/>

HOME-Offshore : Holistic Operation and Maintenance for Energy from Offshore Wind Farms/Details of Grant. (s. d.). Consulté 17 mai 2023, à l'adresse <https://gow.epsrc.ukri.org/NGBOViewGrant.aspx?GrantRef=EP/P009743/1>

How to manage AI procurement in public administration | World Economic Forum. (2023). <https://www.weforum.org/agenda/2023/07/how-to-manage-ai-procurement-in-public-administration/>

Implementation of the national AuroraAI programme—Valtiovarainministeriö. (s. d.). Consulté 20 février 2023, à l'adresse <https://vm.fi/en/auroraai-en>

In vitro screening of environmental chemicals for targeted testing prioritization : The ToxCast project—PubMed. (2010). <https://pubmed.ncbi.nlm.nih.gov/20368123/>

Innovative Applications of AI: The Surtrac Application | Interactive AI Magazine. (s. d.). Consulté 13 avril 2023, à l'adresse <https://interactiveaimag.org/ai-magazine-previews/ai-mag-previews-vol41-no1/>

Ireland's new Road Safety Authority – and strategy 2007–2012 | CIECA. (s. d.). Consulté 18 avril 2023, à l'adresse <https://www.cieca.eu/news/125>

Jacob, S., & Souissi, S. (2022). *L'intelligence artificielle dans l'administration publique au Québec*. Cahiers de recherche sur l'administration publique à l'ère numérique [# 05]. <https://www.administration-numerique.chaire.ulaval.ca/publications/cahiers-de-recherche>

Johnson, C. D., Miller, M. E., Rusnock, C. F., & Jacques, D. R. (2017). A framework for understanding automation in terms of levels of human control abstraction. *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1145–1150. <https://doi.org/10.1109/SMC.2017.8122766>

Joint Research Centre. (2022). *AI Watch : European landscape on the use of artificial intelligence by the public sector*. Publications Office. <https://data.europa.eu/doi/10.2760/39336>

Kaber, D. B. (2018). Issues in Human–Automation Interaction Modeling : Presumptive Aspects of Frameworks of Types and Levels of Automation. *Journal of Cognitive Engineering and Decision Making*, 12(1), 7–24. <https://doi.org/10.1177/1555343417737203>

Kim, D., Yu, C., Xie, S., Chen, Y., Kim, J.-Y., Kim, B., Kulkarni, J. P., & Kim, T. T.-H. (2022). An Overview of Processing-in-Memory Circuits for Artificial Intelligence and Machine Learning. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 12(2), 338–353. <https://doi.org/10.1109/JETCAS.2022.3160455>

Kopponen, A. (2019). *AuroraAI – towards a human-centric, proactive society*.

Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector : Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy*, 44(6), 101976. <https://doi.org/10.1016/j.telpol.2020.101976>

Latzko-Toth, G. (2009). *L'étude de cas en sociologie des sciences et des techniques*. CIRST. https://numerique.banq.qc.ca/patrimoine/details/52327/2532581?docref=hilFlhRoLMiF3L-il_VBg

Leading the way into the age of artificial intelligence—Final report of Finland's Artificial Intelligence Programme 2019. (2019).

Lehner, O. M., & Knoll, C. (2022). *Artificial Intelligence in Accounting : Organisational and Ethical Implications* (1re éd.). Routledge. <https://doi.org/10.4324/9781003198123>

Leikas, J., Johri, A., Latvanen, M., Wessberg, N., & Hahto, A. (2022). Governing Ethical AI Transformation : A Case Study of AuroraAI. *Frontiers in Artificial Intelligence*, 5, 836557. <https://doi.org/10.3389/frai.2022.836557>

Leyer, M., Oberländer, A. M., Dootson, P., & Kowalkiewicz, M. (2020). *Decision-making with artificial intelligence : Towards a novel conceptualization of patterns*.

Made Smarter Review—GOV.UK. (2017). <https://www.gov.uk/government/publications/made-smarter-review>

« *Making Cities More Efficient with Computer Vision and Edge AI*, » a Presentation from Hayden AI - Edge AI and Vision Alliance. (2022, octobre 3). Edge AI and Vision Alliance. <https://www.edge-ai-vision.com/2022/10/making-cities-more-efficient-with-computer-vision-and-edge-ai-a-presentation-from-hayden-ai/>

Mayr, A., Klambauer, G., Unterthiner, T., & Hochreiter, S. (2016). DeepTox : Toxicity Prediction using Deep Learning. *Frontiers in Environmental Science*, 3. <https://doi.org/10.3389/fenvs.2015.00080>

McConnell, A., Mitchell, D., Donaldson, K., Harper, S., Blanche, J., Lim, T., Flynn, D., & Stokes, A. (2021). The Future Workplace : A Symbiotic System of Systems Environment. Dans T. Semwal & F. Iqbal, *Cyber-Physical Systems* (1re éd., p. 259–329). CRC Press. <https://doi.org/10.1201/9781003186380-18>

Medaglia, R., Gil-Garcia, J. R., & Pardo, T. A. (2023). Artificial Intelligence in Government : Taking Stock and Moving Forward. *Social Science Computer Review*, 41(1), 123–140. <https://doi.org/10.1177/08944393211034087>

Memory in Artificial Intelligence : The Key to Advanced Learning—Cow-Shed Startup. (2023). <https://www.cow-shed.com/blog/memory-in-artificial-intelligence-the-key-to-advanced-learning>

Miovision welcomes MicroTraffic to help make streets safer. (2023, février 28). <https://miovision.com/blog/miovision-welcomes-microtraffic-to-help-make-streets-safer/>

Mitchell, D., Blanche, J., Harper, S., Lim, T., Gupta, R., Zaki, O., Tang, W., Robu, V., Watson, S., & Flynn, D. (2022). A review : Challenges and opportunities for artificial intelligence and robotics in the offshore wind sector. *Energy and AI*, 8, 100146. <https://doi.org/10.1016/j.egyai.2022.100146>

Morio, C. (2014a). Fonctionnaire. Dans *Dictionnaire d'administration publique* (p. 234–235). Presses universitaires de Grenoble; Cairn.info. <https://www.cairn.info/dictionnaire-d-administration-publique--9782706121371-p-234.htm>

Morio, C. (2014b). Usager. Dans *Dictionnaire d'administration publique* (p. 515–516). Presses universitaires de Grenoble; Cairn.info. <https://www.cairn.info/dictionnaire-d-administration-publique--9782706121371-p-515.htm>

MTA to Install ABLE Cameras on 300 Additional Buses. (2022, octobre 3). Metropolitan Transportation Authority (MTA). <https://new.mta.info/press-release/mta-install-able-cameras-300-additional-buses-across-nine-routes-end-of-2022>

Mucchielli, A. (2007). *Les processus intellectuels fondamentaux sous-jacents aux techniques et méthodes qualitatives*.

Mykhailychenko, R. (2019). The 4th industrial revolution : Responding to the impact of artificial intelligence on business. *Foresight*, 21(2), 318–319. <https://doi.org/10.1108/FS-04-2019-109>

New York City Deploys Video Bus Lane Enforcement. (2022, août 18). Government Technology ([Govtech.com](https://www.govtech.com)). <https://www.govtech.com/fs/new-york-city-deploys-video-bus-lane-enforcement>

New York City to install Siemens camera system for bus lane enforcement | Traffic Technology Today. (2019, mars 29). Traffic Technology International. <https://www.traffictechnologytoday.com/news/enforcement/new-york-city-to-install-siemens-camera-system-for-bus-lane-enforcement.html>

New York City Transit expands bus lane enforcement program—Bus—Metro Magazine. (2019, novembre 1). METRO Magazine. <https://www.metro-magazine.com/10030873/new-york-city-transit-expands-bus-lane-enforcement-program>

New York City Uses Mobile Camera System to Catch Vehicles in Bus Lanes | Vision Systems Design. (2022, septembre 12). VISION SYSTEMS DESIGN. <https://www.vision-systems.com/embedded/article/14282607/new-york-city-uses-mobile-camera-system-to-catch-vehicles-in-bus-lanes>

Ng, C., & Alarcon, J. (2020). *Artificial Intelligence in Accounting : Practical Applications* (1re éd.). Routledge. <https://doi.org/10.4324/9781003003342>

Nickles, M., Rovatsos, M., & Weiss, G. (Éds.). (2004). *Agents and computational autonomy : Potential, risks, and solutions.* Springer.

Offshore energy infrastructure research unveiled—Energy Sustainability Solutions. (2019). <https://essmag.co.uk/offshore-energy-infrastructure-research-unveiled/>

Offshore wind—Great.gov.uk international. (s. d.). Consulté 17 mai 2023, à l'adresse <https://www.great.gov.uk/international/content/investment/sectors/offshore-wind/>

ORCA Robotics—Robot and Asset Self-Certification. (s. d.). Consulté 17 mai 2023, à l'adresse <https://orcahub.org/innovation/focus-areas/robot-and-asset-self-certification>

Osman, C.-C. (2019). Robotic Process Automation : Lessons Learned from Case Studies. *Informatica Economica*, 23(4/2019), 66-71. <https://doi.org/10.12948/issn14531305/23.4.2019.06>

Our Cameras | GoSafe Welsh Road Casualty Reduction Partnership. (s. d.). Consulté 18 avril 2023, à l'adresse <https://www.gosafe.org/learn/how-we-enforce/our-cameras/>

Papers-ORI/Oxford Robotics Institute. (s. d.). Consulté 23 mai 2023, à l'adresse <https://www.ori.ox.ac.uk/publications/papers/>

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 30(3). WorldCat.org. <https://doi.org/10.1109/3468.844354>

Peeters, R. (2020). The agency of algorithms : Understanding human–algorithm interaction in administrative decision-making. *Information Polity*, 25(4), 507-522. <https://doi.org/10.3233/IP-200253>

Pereira, L. M., & Lopes, A. B. (2020). *Machine Ethics : From Machine Morals to the Machinery of Morality* (Vol. 53). Springer International Publishing. <https://doi.org/10.1007/978-3-030-39630-5>

Priority Dispatch Announces Automatic Quality Assurance of Emergency Calls Through a Partnership With The Artificial Intelligence Company Corti.ai—Corti. (2023, février 17). <https://www.corti.ai/stories/priority-dispatch-announces-automatic-qualityassurance-of-emergency-calls-through-a-partnershipwith-the-artificial-intelligence-company-corti-ai>

Rao, A., & Verweij, G. (2017). *PwC's Global Artificial Intelligence Study : Sizing the prize.* PwC. <https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html>

Redflex Holdings Limited Annual Report 2012. (2012). *ANNUAL REPORT.*

Redflex Traffic Systems—Verdict Traffic. (s. d.). Consulté 18 avril 2023, à l'adresse https://www.roadtraffic-technology.com/contractors/traffic_man/redflex-traffic-systems/

Research—The National Robotarium. (s. d.). Consulté 19 mai 2023, à l'adresse <https://thenationalrobotarium.com/research/>

Robotics and Artificial Intelligence Hubs | Autonomy and Verification Network. (s. d.). Consulté 23 mai 2023, à l'adresse <https://autonomy-and-verification.github.io/projects/rai-hubs>

Robots primed to revolutionise offshore energy industry safety. (2021). <https://www.energyvoice.com/oilandgas/north-sea/332959/robots-revolutionise-offshore-energy-industry-safety/>

Russell, S. J., Norvig, P., & Davis, E. (2010). *Artificial intelligence : A modern approach* (3rd ed). Prentice Hall.

Safety Camera Expansion 10th February 2020. (s. d.). Consulté 17 avril 2023, à l'adresse <https://www.rsa.ie/news-events/news/details/2020/02/10/safety-camera-expansion-10th-february-2020>

Safety Cameras—Garda. (s. d.). Consulté 16 février 2023, à l'adresse <https://www.garda.ie/en/roads-policing/safety-cameras/>

Sanchez-Graells, A. (2022). Identifying Emerging Risks in Digital Procurement Governance. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4254931>

Sayed, M. E., Roberts, J. O., McKenzie, R. M., Aracri, S., Buchoux, A., & Stokes, A. A. (2021). Limpet II : A Modular, Untethered Soft Robot. *Soft Robotics*, 8(3), 319-339. <https://doi.org/10.1089/soro.2019.0161>

Sayed, M., Nemitz, M., Aracri, S., McConnell, A., McKenzie, R., & Stokes, A. (2018). The Limpet : A ROS-Enabled Multi-Sensing Platform for the ORCA Hub. *Sensors*, 18(10), 3487. <https://doi.org/10.3390/s18103487>

Seidensticker, K. (2018). *Predictive Policing in Germany.*

Seidensticker, K. (2021). *SKALA - Predictive Policing in North Rhine-Westphalia.*

Selten, F., & Meijer, A. (2021). Managing Algorithms for Public Value: *International Journal of Public Administration in the Digital Age*, 8(1), 1-16. <https://doi.org/10.4018/IJPADA.20210101.0a9>

Simek, P. (2022). *The future of open city streets could start with smarter traffic lights.*

Smart Infrastructure for Future Urban Mobility | US Department of Transportation. (2017). <https://www.transportation.gov/utc/smart-infrastructure-future-urban-mobility>

Smith, S. F., Barlow, G. J., Xie, X.-F., & Rubinstein, Z. B. (2013). *Surtrac: Scalable Urban Traffic Control.*

SOS Alarm : Increasing Quality and Safety at Sweden's National Medical Call Center. (s. d.). Consulté 21 avril 2023, à l'adresse <https://www.corti.ai/case-studies/sos-alarm-improved-triaging>

Speed cameras in Ireland : How they work and types | Carzone Advice. (2023, février). <https://www.carzone.ie/motoring-advice/speed-cameras-in-ireland-how-they-work-and-types/2925>

Stratégie d'intégration de l'intelligence artificielle dans l'administration publique 2021-2026. (2021). *administration publique*. <https://www.quebec.ca/gouvernement/politiques-orientations/vitrine-numeriqc/strategie-integration-ia-administration-publique-2021-2026/mettre-a-profit-les-avancees-de-ia-au-service-du-secteur-public>

Study shows Irish speed cameras provide five-fold benefit | ITS International. (s. d.). Consulté 18 avril 2023, à l'adresse <https://www.itsinternational.com/its2/feature/study-shows-irish-speed-cameras-provide-five-fold-benefit>

Surtrac for the People : Upgrading the Surtrac Pittsburgh Deployment to Incorporate Pedestrian Friendly Extensions and Remote Monitoring Advances. (s. d.). Consulté 13 avril 2023, à l'adresse <https://trid.trb.org/view/1681289>

Sweden Turns to Artificial Intelligence During Medical Emergencies | by Corti News | Corti | Medium. (2020, février). <https://medium.com/corti-ai/sweden-turns-to-artificial-intelligence-during-medical-emergencies-67ef94af53a5>

Swedish Municipality Uses AI To Objectively Select and Engage Candidates In Recruitment | Tengai Unbiased. (s. d.). Consulté 14 février 2023, à l'adresse <https://www.mynewsdesk.com/tengai-interview-robot/pressreleases/swedish-municipality-uses-ai-to-objectively-select-and-engage-candidates-in-recruitment-3230595>

Tengai Announces World's First Human-Like Avatar for Unbiased Candidate Conversations | Tengai Unbiased. (s. d.). Consulté 14 février 2023, à l'adresse <https://www.mynewsdesk.com/tengai-interview-robot/pressreleases/tengai-announces-worlds-first-human-like-avatar-for-unbiased-candidate-conversations-3214835>

The AuroraAI national artificial intelligence programme begins – with the aim of using artificial intelligence to bring people and services together in a better way. (2020). <https://valtioneuvosto.fi/en/-/10623/the-auroraai-national-artificial-intelligence-programme-begins-with-the-aim-of-using-artificial-intelligence-to-bring-people-and-services-together-in-a-better-way>

The Furhat Robot | Furhat Robotics. (s. d.). Consulté 21 mars 2023, à l'adresse <https://furhatrobotics.com/furhat-robot/>

The science behind Tengai's validated candidate assessment. (s. d.). Consulté 20 mars 2023, à l'adresse <https://tengai.io/candidate-assessment/>

The well-being of young people and those looking for work can be supported with AuroraAI network network | Digital and population data services agency. (2022). <https://dvv.fi/en/-/the-well-being-of-young-people-and-those-looking-for-work-can-be-supported-with-auroraai-network-network>

This AI detects cardiac arrests during emergency calls | World Economic Forum. (2018, juin). <https://www.weforum.org/agenda/2018/06/this-ai-detects-cardiac-arrests-during-emergency-calls/>

Thomas, R. S., Bahadori, T., Buckley, T. J., Cowden, J., Deisenroth, C., Dionisio, K. L., Frithsen, J. B., Grulke, C. M., Gwinn, M. R., Harrill, J. A., Higuchi, M., Houck, K. A., Hughes, M. F., Hunter, E. S., Isaacs, K. K., Judson, R. S., Knudsen, T. B., Lambert, J. C., Linnenbrink, M., ... Williams, A. J. (2019). The Next Generation Blueprint of Computational Toxicology at the U.S. Environmental Protection Agency. *Toxicological Sciences*, 169(2), 317-332. <https://doi.org/10.1093/toxsci/kfz058>

Toll, D., Lindgren, I., Melin, U., & Madsen, C. Ø. (2019). Artificial Intelligence in Swedish Policies : Values, Benefits, Considerations and Risks. Dans I. Lindgren, M. Janssen, H. Lee, A. Polini, M. P. Rodríguez Bolívar, H. J. Scholl, & E. Tambouris (Éds.), *Electronic Government* (Vol. 11685, p. 301-310). Springer International Publishing. https://doi.org/10.1007/978-3-030-27325-5_23

Tony Evans & Peter Hupe. (2020). *Discretion and the Quest for Controlled Freedom*. Palgrave Macmillan; eBook Collection (EBSCOhost). <https://aces.bibl.ulaval.ca/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=2236659&lang=fr&site=ehost-live>

Towards the future of the energy industry with intelligent robots | AI FOR GOOD WEBINARS - YouTube. (2022). <https://www.youtube.com/watch?v=wS87eGyPl1c>

Toxicity Testing in the 21st Century : A Vision and a Strategy | The National Academies Press. (2007). <https://nap.nationalacademies.org/catalog/11970/toxicity-testing-in-the-21st-century-a-vision-and-a>

Transit Buses—Hayden AI. (s. d.). Hayden AI official website. Consulté 16 février 2023, à l'adresse <https://www.hayden.ai/applications/transit-buses>

UK renewable energy roadmap : 2011—GOV.UK. (2011). <https://www.gov.uk/government/publications/renewable-energy-roadmap>

Unique underwater drone carries out first ever autonomous inspection of windfarms—The National Robotarium. (2022). <https://thenationalrobotarium.com/unique-underwater-drone-carries-out-first-ever-autonomous-inspection-of-windfarms/>

US Patent for Systems and methods for managing traffic rules using multiple mapping layers with traffic management semantics Patent (Patent # 11,322,017 issued May 3, 2022)—Justia Patents Search. (s. d.). Consulté 27 mars 2023, à l'adresse <https://patents.justia.com/patent/11322017>

Vanhanen, M. M., & Paatero, S. (2019). *The AuroraAI Programme*.

van Noordt, C., & Misuraca, G. (2022). Artificial intelligence for the public sector : Results of landscaping the use of AI in government across the European Union. *Government Information Quarterly*, 39(3), 101714. <https://doi.org/10.1016/j.giq.2022.101714>

Vassileva, B., & Zwilling, M. (Éds.). (2021). *Responsible AI and Ethical Issues for Businesses and Governments*: IGI Global. <https://doi.org/10.4018/978-1-7998-4285-9>

Veale, M., & Brass, I. (2020). *Administration by Algorithm ? Public Management meets Public Sector Machine Learning*. *Algorithmic Regulation*. <https://doi.org/10.1093/oso/9780198838494.001.0001>

Vepřek, L. H., Sibert, L., Sehn, L., Köpp, L., & Friedrich, D. (2020). Beyond Effectiveness : Legitimising Predictive Policing in Germany. *Kriminologie - Das Online-Journal | Criminology - The Online Journal*, 423-443 Seiten. <https://doi.org/10.18716/OJS/KRIMOJ/2020.3.3>

Wang, P. (2008). *What Do You Mean by "AI"?*

What is Artificial Intelligence Doing on your Emergency Call in France ? | by Corti News | Corti | Medium. (2020, février 5). <https://medium.com/corti-ai/what-is-artificial-intelligence-doing-on-your-emergency-call-in-france-603f295e4672>

White House launches artificial intelligence information portal. (2021). <https://ourpublicservice.org/blog/white-house-launches-artificial-intelligence-information-portal/>

White, J., & Searle, R. (Éds.). (2015). *Rethinking Machine Ethics in the Age of Ubiquitous Technology*: IGI Global. <https://doi.org/10.4018/978-1-4666-8592-5>

Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial Intelligence and the Public Sector—Applications and Challenges. *International Journal of Public Administration*, 42(7), 596-615. <https://doi.org/10.1080/01900692.2018.1498103>

Young, M. M., Bullock, J. B., & Lecy, J. D. (2019). Artificial Discretion as a Tool of Governance : A Framework for Understanding the Impact of Artificial Intelligence on Public Administration. *Perspectives on Public Management and Governance*, gvz014. <https://doi.org/10.1093/ppmgov/gvz014>

Zerilli, J. (2021). *A Citizen's Guide to Artificial Intelligence*. The MIT Press. <https://doi.org/10.7551/mitpress/12518.001.0001>

Zhang, L., Tan, J., Han, D., & Zhu, H. (2017). From machine learning to deep learning : Progress in machine intelligence for rational drug discovery. *Drug Discovery Today*, 22(11), 1680-1685. <https://doi.org/10.1016/j.drudis.2017.08.010>

Zicari, R. V., Brodersen, J., Brusseau, J., Dudder, B., Eichhorn, T., Ivanov, T., Kararigas, G., Kringen, P., McCullough, M., Moslein, F., Mushtaq, N., Roig, G., Sturtz, N., Tolle, K., Tithi, J. J., Van Halem, I., & Westerlund, M. (2021). Z-Inspection © : A Process to Assess Trustworthy AI. *IEEE Transactions on Technology and Society*, 2(2), 83-97. <https://doi.org/10.1109/TTS.2021.3066209>



obvia

obvia.ca