

*Mémoire  
présenté à l'École nationale d'administration publique  
dans le cadre du programme de Maîtrise en administration publique  
pour l'obtention du grade de Maître ès science (M. Sc.)*

Mémoire intitulé  
**Revue de portée sur l'utilisation des applications d'intelligence artificielle par  
les autorités fiscales dans la détection des comportements fiscaux frauduleux**

Présenté par  
**Adama Dao**

**Juillet, 2025**

Le mémoire intitulé

**Revue de portée sur l'utilisation des applications d'intelligence artificielle par les autorités fiscales dans la détection des comportements fiscaux frauduleux**

Présenté par

**Adama Dao**

Est évalué par les membres du jury de mémoire suivants :

Marie-Soleil Tremblay, Professeur titulaire et présidente  
Justin Lawarée, Professeur adjoint et directeur de mémoire  
Guillaume Beaumier, Professeur adjoint et évaluateur

« **À mon épouse**, Ini Augustine TIOYÉ, merci infiniment pour ta patience et ta détermination à m'accompagner tout au long de cette aventure. Ta présence à mes côtés à des heures tardives a été un catalyseur.

**À mes enfants Nousso et Bahoudé DAO**, merci d'avoir supporté et accepté mes absences répétées. »

## Remerciements

---

Toute œuvre intellectuelle est le fruit d'un travail acharné rendu possible grâce au concours de personnes inspirantes.

Je remercie particulièrement mon directeur de mémoire, le professeur Justin Lawarée qui a cru en mon projet dès les premiers instants où j'ai pris attache avec lui pour m'accompagner dans cette aventure, alors qu'il ne m'avait jamais rencontré auparavant en tant qu'étudiant. Le Pr Lawarée a toujours été présent pour me prodiguer des conseils avisés avec rigueur et professionnalisme. Grâce à son concours et son appui constant, j'ai obtenu la bourse d'études supérieures du Canada au niveau de la maîtrise (BESC-M) CRSH-CRSNG-IRSC à travers ce mémoire.

J'adresse également mes remerciements aux membres du jury, les professeurs Marie-Soleil Tremblay et Guillaume Beaumier pour leur précieuse contribution à la finalisation de mon projet de mémoire.

Finalement, j'adresse mes vifs remerciements au professeur Séni Mahamadou Ouédraogo, agrégé des facultés de droit à l'université Thomas Sankara du Burkina Faso, déclencheur de ma passion pour le numérique et qui m'a toujours encouragé pour la recherche et couvert d'encouragements et de motivation.

## Résumé

---

Avec la mondialisation et la numérisation de l'économie, les entreprises et les individus ont développé des stratagèmes de fraude fiscale plus sophistiqués. En réponse, les autorités fiscales (AF) nationales ont intégré des systèmes d'intelligence artificielle (SIA) dans la détection de la fraude, l'évasion et l'évitement fiscal. Afin d'analyser les retombées de ces nouvelles pratiques, cette recherche explore les caractéristiques des SIA utilisés.

Une revue de portée (RP) a été conduite grâce au protocole *PRISMA* et l'approche *SPIDER*. Ces deux méthodes, couplées au logiciel DistillerSR, ont permis d'identifier 12 articles pertinents évalués par les pairs entre 2017 et 2024. En outre, une analyse thématique déductive des données a été mise en œuvre afin de faciliter l'interprétation des thèmes identifiables.

Les AF ont déployé divers algorithmes supervisés, non supervisés et hybrides pour détecter divers comportements frauduleux liés aux impôts directs et indirects. L'utilisation des SIA a également mis en évidence la qualité du service, le développement de stratégies fiscales novatrices et une amélioration des opérations de vérification.

La RP a démontré un apport significatif des SIA dans la détection des comportements fiscaux frauduleux. Cependant, plusieurs limites ont été relevées comme l'absence de preuves sur les divers changements annoncés sur le plan organisationnel et stratégique, le manque d'impact économique chiffré. Des recommandations ont été proposées pour des études à long terme et non ponctuelles sur les impacts économiques ainsi que la satisfaction des contribuables.

## Abstract<sup>1</sup>

---

With the globalization and digitalization of the economy, businesses and individuals have developed increasingly sophisticated tax fraud schemes. In response, national tax authorities (TAs) have integrated artificial intelligence systems (AIS) to detect fraud, evasion, and avoidance. To analyze the outcomes of these new practices, this research explores the characteristics of the AIS being used.

A scoping review (SR) was conducted using the PRISMA protocol and the SPIDER approach. These two methods, combined with the use of DistillerSR software, enabled the identification of 12 peer-reviewed articles published between 2017 and 2024. Additionally, a deductive thematic analysis was implemented to facilitate the interpretation of identifiable themes.

TAs have implemented various supervised, unsupervised, and hybrid algorithms to detect fraudulent behavior related to both direct and indirect taxes. The use of AIS has also highlighted improvements in service quality, tax strategy optimization, and audit operations.

The SR revealed a significant contribution of AIS in detecting fraudulent tax behavior. However, several limitations were noted, including the lack of evidence regarding the organizational and strategic changes announced, and the absence of quantified economic impacts. Recommendations have been made for conducting long-term studies, rather than one-off analyses, focusing on economic impacts and taxpayer satisfaction.

---

<sup>1</sup> La traduction a été faite par ChatGPT

## Table des matières

---

Remerciements .....	ii
Résumé .....	iii
Liste des tableaux .....	vii
Liste des figures .....	viii
Liste des abréviations, sigles et acronymes .....	ix
INTRODUCTION .....	10
<b>CHAPITRE 1 : FONDEMENTS DE LA RECHERCHE</b> .....	13
1.1 Problématisation .....	13
1.1 Objectifs de la recherche .....	16
<b>CHAPITRE 2 : CADRE THÉORIQUE ET REVUE DE LA LITTÉRATURE</b> .....	17
2.1 La typologie des autorités fiscales (AF) .....	17
2.2 L'utilisation des SIA pour la détection des comportements fiscaux frauduleux .....	20
2.2.1 L'intelligence artificielle .....	21
2.2.2 Les systèmes d'intelligence artificielle (SIA) .....	21
2.3 Les comportements fiscaux qualifiés de frauduleux .....	26
2.3.1 La fraude fiscale .....	26
2.3.2 L'évasion fiscale .....	27
2.3.3 L'évitement fiscal et la planification fiscale abusive .....	28
2.3.4 Analyse comparée des concepts de fraude en matière fiscale .....	29
2.4 État de la question .....	32
2.4.1 L'intelligence artificielle .....	33
2.4.2 Les méthodes d'apprentissage automatique .....	33
2.4.2.1 L'apprentissage supervisé .....	33
2.4.2.2 L'apprentissage non supervisé .....	35
2.4.2.3 L'apprentissage semi-supervisé (hybride) .....	35
2.4.2.4 L'apprentissage par renforcement .....	36
<b>CHAPITRE 3 : DEVIS MÉTHODOLOGIQUE</b> .....	37
3.1 Processus de la revue de portée .....	37
3.1.1 Le protocole <i>PRISMA</i> .....	37
3.2 Les critères d'éligibilité : la méthode <i>SPIDER</i> .....	38
3.3 Sources d'information .....	41
3.3.1 La détermination des sources des références .....	41
3.3.2 Stratégie de recherche .....	42
3.4 Processus de sélection des références avec le logiciel <i>DistillerSR</i> .....	42
3.4.2 Première phase : Élimination des références sur la base des types de référence .....	43
3.4.3 Deuxième phase : Les critères d'éligibilité contenus dans le titre et le résumé .....	44
3.4.4 Troisième phase : Critères liés au contenu des articles .....	45
3.5 Analyse de données .....	47
3.5.1 Analyse thématique déductive des données .....	48
<b>CHAPITRE 4 : RÉSULTATS</b> .....	51
4.1 Caractéristiques générales de l'échantillon .....	51

4.2 Analyse des données et processus d'évaluation des SIA.....	52
4.3 Résultats liés aux comportements fiscaux frauduleux, aux interventions des SIA, aux effets et impacts.....	53
4.3.1 Intervention des systèmes d'intelligence artificielle sur les comportements fiscaux frauduleux identifiés .....	53
4.3.1.1 Apprentissage supervisé.....	54
4.3.1.2 Apprentissage non supervisé.....	55
4.3.1.3 Apprentissage hybride ou mixte .....	55
4.3.2 Effets et impacts organisationnels .....	58
4.3.2.1 Renforcement de la performance des opérations fiscales .....	58
4.3.2.2 Optimisation de la gestion stratégique et du ciblage .....	58
4.3.2.3 Bonification de la qualité du service aux contribuables.....	59
<b>CHAPITRE 5 : DISCUSSIONS</b> .....	60
5.1 Synthèse des principaux résultats et apports de la revue de portée .....	60
5.2 Examen critique des méthodologies de l'échantillon d'étude.....	61
5.3 Analyse critique de la relation entre les SIA utilisés et les comportements fiscaux frauduleux.....	62
5.4 Analyse des effets et des impacts organisationnels .....	65
5.4.1 Effets et impacts stratégiques .....	65
5.4.2 Effets et impacts organisationnels .....	66
5.4.3 Effets et impacts sur la qualité du service rendu aux contribuables .....	68
5.5 Limites des études analysées et de la revue de portée .....	69
5.6 Pistes de recherche futures .....	70
<b>CONCLUSION</b> .....	72
<b>BIBLIOGRAPHIE ET RÉFÉRENCES</b> .....	74
<b>ANNEXES</b> .....	83

## Liste des tableaux

---

Tableau 1 : Caractéristiques générales des AF

Tableau 2 : Panorama des SIA

Tableau 3 : Synthèse des principaux comportements fiscaux frauduleux

Tableau 4 : Synthèse des critères d'éligibilité selon l'outil *SPIDER*

Tableau 5 : Inclusion et exclusion des types de références

Tableau 6 : Nombre de références incluses et exclues à la première phase

Tableau 7 : Nombre de références incluses et exclues à la deuxième phase

## Liste des figures

---

Figure 1 : Diagramme de flux PRISMA 2020

Figure 2 : Exemple d'arborescence générée par l'analyse thématique déductive

Figure 3 : Répartition des articles par pays et par année de publication

Figure 4 : Modèles d'apprentissage automatique, comportements fiscaux frauduleux et leurs caractéristiques

## Liste des abréviations, sigles et acronymes

---

AF	Autorité fiscale
ARC	Agence du revenu du Canada
BD	Bases de données
IA	Intelligence artificielle
<i>IRS</i>	<i>Internal Revenue Service</i>
LIR	Loi de l'impôt sur le revenu
LTA	Loi sur la taxe d'accise
OCDE	Organisation de coopération et de développement économiques
<i>OECD</i>	<i>Organisation for Economic Co-operation and Development</i>
<i>PRISMA</i>	<i>Preferred Reporting Items for Systematic reviews and Meta-Analyses</i>
RP	Revue de portée
SIA	Systèmes d'intelligence artificielle
<i>SPIDER</i>	<i>Sample - Phenomenon of Interest – Design - Evaluation et Research type</i>
TPS	Taxe sur les produits et services
TVA	Taxe sur la valeur ajoutée

## INTRODUCTION

---

La détection et la prévention de la fraude en fiscalité sont un enjeu majeur pour les autorités fiscales des États développés ou en développement. L'avancée des technologies informationnelles et numériques a transformé profondément le fonctionnement des institutions publiques et plus particulièrement les administrations responsables de la collecte des impôts. Avec l'introduction du gouvernement numérique dans les années 1980 (St-Amant, 2005), les services offerts par les autorités fiscales se sont digitalisés rapidement (Bassey et al., 2022) pour non seulement servir les citoyens, s'adapter à l'évolution du commerce numérique, mais aussi pour faire face aux enjeux de la fraude et de l'évasion fiscale. (Lagodiienko et al., 2022)

Aujourd'hui, à travers l'intégration des Systèmes d'intelligence artificielle (SIA), les administrations fiscales tirent parti des renseignements produits par les organisations publiques et privées ainsi que les individus afin de combattre la non-conformité et la fraude fiscales. Cependant, plusieurs défis comme la complexité des transactions internationales, des techniques de fraude et l'accroissement des données financières rendent les audits fiscaux et la collecte des impôts plus difficiles à cerner.

En effet, la mondialisation de l'économie a favorisé la délocalisation des entreprises multinationales, la multiplication des échanges commerciaux et l'augmentation accrue des données financières. Couplée aux nouvelles technologies, la mondialisation a offert aux particuliers et aux entreprises la possibilité d'échapper à l'impôt (Alstadsaeter et al., 2023). En 2016, le Consortium international des journalistes d'investigation (ICIJ) a rendu publics des millions de données financières provenant de paradis fiscaux (Alvarardo et al., 2019). Ces informations financières concernaient 213 634 entités financières *offshore* créées par la firme d'avocats Mossack-Fonseca.

Plusieurs auteurs ont soulevé les comportements frauduleux des entreprises multinationales. Parmi eux, Zheng et al. (2023) indiquent que les multinationales peuvent acquérir de nombreuses sociétés-écrans dans diverses juridictions en investissant dans plusieurs secteurs économiques. Elles transfèrent ensuite leurs profits dans des paradis fiscaux.

Quant aux données, elles sont massives et complexes. D'abord, Zheng et al. (2023) mentionnent que les stratagèmes fiscaux contiennent d'énormes volumes de données provenant de sources

diverses à travers le monde. En outre, les avancées technologiques et la délocalisation des entreprises ont contribué à la diversification des lieux de stockage des données. Ainsi, selon l'Organisation de coopération et de développement économiques (OCDE, 2022), les Administrations fiscales n'ont pas souvent accès à toutes les données de leurs contribuables, car celles intéressant l'autorité fiscale d'un pays peuvent être détenues dans une autre juridiction.

Au vu de ces différents défis, les méthodes traditionnelles de contrôle et de collecte des impôts deviennent moins efficaces.

Avec le développement rapide de l'économie numérique, les États sont confrontés à la fraude liée à la fiscalité internationale par les entreprises. Selon Raikov (2021), il est difficile de déterminer comment et où les revenus sont générés, d'identifier le pays dans lequel les serveurs hébergent le domaine web, où la direction opère, où l'entreprise traite les commandes, etc.

En matière de pertes fiscales, le rapport 2021 de l'organisme britannique *Tax Justice Network* (2021) montre que les États, objets de l'étude, enregistrent des pertes en recettes fiscales directes d'environ 312 milliards de dollars US par an, dues à la fraude et l'évasion fiscales des multinationales. Par ailleurs, les pertes fiscales des gouvernements liées aux activités illicites des particuliers fortunés qui transfèrent leur argent à l'étranger s'élèvent à 171 milliards de dollars US par an.

Malgré les efforts consentis par les gouvernements à travers les diverses ententes d'échanges de renseignements financiers entre pays, le rapport *Tax Justice Network* (2024) indique une augmentation des pertes fiscales dues aux entreprises. Selon le rapport de 2023, tous les États, objets de l'étude, ont enregistré des pertes fiscales d'environ 480 milliards de dollars US imputables à l'exploitation abusive des failles du système fiscal et à l'évasion fiscale par les entreprises et les particuliers. Plus particulièrement, les études du rapport de 2023 mentionnent que le Canada, par exemple, a perdu des recettes fiscales d'environ 3,65 milliards de dollars US et inflige des pertes fiscales de 11,56 milliards de dollars US à d'autres pays. La perte fiscale mondiale a progressé de 35 % entre 2021 et 2023.

Au vu de ces enjeux, il apparaît que les méthodes traditionnelles de lutte contre les comportements frauduleux connaissent des limites, car les fraudeurs fiscaux adaptent leurs stratégies en fonction

des mesures prises par les gouvernements. Depuis plus d'une décennie, les États intègrent de plus en plus les SIA dans leur stratégie de lutte contre les différents stratagèmes de fraude fiscale.

Afin d'analyser les retombées de ces nouvelles pratiques, une **revue de portée** (RP) est réalisée afin de mettre en lumière les caractéristiques des SIA ainsi que leurs applications aux enjeux liés à la fraude fiscale. Cette étude sera organisée de la façon suivante : le chapitre 1 traitera des fondements de la recherche. Ce stade de l'étude sera consacré à la description de la problématisation ainsi qu'aux objectifs de l'étude. Il sera suivi d'une revue de la littérature et du cadre théorique en développant les principaux concepts utilisés (chapitre 2). Il sera défini ensuite le cadre méthodologique de l'étude (chapitre 3) qui sera suivi des résultats (chapitre 4) qui seront discutés au chapitre 5.

## CHAPITRE 1 : FONDEMENTS DE LA RECHERCHE

---

L'objet de ce chapitre est de poser les fondements conceptuels et scientifiques de cette étude. Cela s'articule autour de deux aspects fondamentaux que sont la problématique et les objectifs de recherche. La problématisation met en évidence les enjeux soulevés par l'usage des SIA par les AF. La section suivante présente les objectifs de la recherche ainsi que les questions spécifiques qui la guideront.

### 1.1 Problématisation

La numérisation de l'administration fiscale désigne l'ensemble des mesures appuyées sur les technologies de l'information visant à digitaliser les différentes étapes de la relation entre les contribuables et les autorités fiscales, du premier enregistrement jusqu'au paiement des impôts (Lagodiienko et al., 2022). Afin de tirer profit de ces technologies modernes, les États ont amélioré leurs infrastructures pour des raisons diverses : établir une meilleure relation avec les contribuables, permettre le paiement des impôts, établir des facturations numériques, collecter et analyser les données fiscales (Huong et al., 2023), améliorer les systèmes d'évaluation des risques, détecter les comportements frauduleux (Lagodiienko et al., 2022). Par exemple, l'AF du Bangladesh (le *National Board of Revenue*) utilisait plusieurs technologies numériques comme « l'enregistrement, la déclaration, le paiement et les retenues à la source électroniques pour lutter contre l'évasion fiscale depuis 2014 » (Sarker et Ahmed, 2022, p. 245).

D'ailleurs, dans sa réponse aux recommandations du Comité consultatif sur l'observation à l'étranger, l'Agence du revenu du Canada (ARC) a intégré un nouveau système d'évaluation intégrée des risques. Ce système procède à une analyse des renseignements sur les contribuables dans les bases de données de l'ARC en appliquant des algorithmes sophistiqués pour évaluer le risque d'inobservation fiscale des grandes entreprises (Agence du revenu du Canada, s. d.). Par ailleurs, dans son Plan ministériel 2018-2019, l'ARC indique que les SIA permettraient d'améliorer la détection des stratagèmes fiscaux complexes, difficiles à repérer et impliquant diverses parties, entreprises comme particuliers. Cependant, les rapports sur les résultats ministériels de l'ARC ne fournissent pas des résultats probants sur les opportunités qu'offrent les SIA dans la détection de la fraude fiscale.

Au niveau international, OCDE a développé plusieurs instruments juridiques en vue de lutter contre la fraude et l'évasion fiscales auxquelles participent plus de 150 juridictions. Les plus importants de ces instruments sont : la Convention concernant l'assistance administrative mutuelle en matière fiscale (OCDE, s.d.), le Forum mondial sur la transparence et l'échange de renseignements à des fins fiscales (OCDE, s.d.) et les échanges automatiques de renseignements (OCDE, s.d.). La collecte et l'utilisation massive des données financières, résultant des conventions internationales, constituent une opportunité pour les AF, car ces renseignements sont susceptibles d'alimenter les SIA. Ainsi, la disponibilité des informations encourage les AF à investir davantage dans les nouvelles technologies fiscales plus avancées (Ihnatišínová, 2021) comme les SIA.

D'autre part, le *Forum on Tax administration* (OCDE, s.d.) a produit un inventaire des initiatives en matière de technologie fiscale avancée. Le rapport indique que plus de 50 % des juridictions ayant participé au sondage utilisent les SIA pour la détection de la fraude et de l'évasion fiscales (OCDE, 2022). L'AF de l'Inde a introduit les SIA pour la première fois en vue d'effectuer des traitements répétitifs de documents, de collecter des données et de décider de l'authenticité ou non des déclarations de revenus des contribuables (Rathi et al., 2021).

Dans des communiqués de presse en 2023 et 2024, l'autorité fiscale des États-Unis (*Internal Revenue Service*) a annoncé l'utilisation de SIA de pointe afin d'identifier les risques de non-conformité dans le secteur de la fiscalité des sociétés, de personnes, de l'impôt général sur le revenu et de la fiscalité internationale. Dans le même communiqué de presse, l'IRS indique avoir recouvré plusieurs millions de dollars grâce à l'utilisation des SIA dans des dossiers frauduleux (*U.S. Department of the Treasury, 2023*).

En Suède, l'autorité fiscale suédoise (*Skatteverket*) a créé le projet *TAIS* en recourant aux SIA afin d'améliorer le processus d'identification des segments de contribuables auteurs de fraudes (OCDE, 2022) pour ainsi lutter contre la fraude et l'évasion fiscales. Quant au Brésil, l'AF utilise la géolocalisation et les SIA pour repérer les habitants d'un même quartier qui ont généralement des revenus, actifs et dépenses similaires (OCDE, 2022). En République populaire de Chine, le *State Tax Administration (STA)* combine l'apprentissage automatique à d'autres technologies pour sélectionner les dossiers de vérification fiscale électronique à haut risque (OCDE, 2022). Cette

revue indique que plusieurs AF utilisent des SIA pour prévenir et contrer les comportements fiscaux frauduleux et évaluer les risques fiscaux.

Toujours en Europe, la Direction générale des finances publiques a mis en place le système « foncier innovant » en 2021, alimenté par les SIA et les données géographiques. En 2023, ce système a détecté près de 140 000 piscines non déclarées. Ces découvertes de pratiques frauduleuses ont engendré des recettes de plus de 40 millions de taxes foncières supplémentaires (HuffPost. 2024), augmentant ainsi les impôts directs recouvrés. En outre, la commission des finances, de l'économie générale et du contrôle budgétaire du Parlement français a produit un rapport d'information sur l'usage des nouvelles technologies dans le contrôle fiscal le 6 juin 2024. Ce rapport consacre l'utilisation du SIA dans la détection de la fraude fiscale. Cependant, le rapport soulève l'enjeu de l'évaluation des apports et des limites du recours aux SIA pour lutter contre la fraude fiscale (Assemblée nationale du Québec, 2024).

Dans cette perspective, plusieurs recherches scientifiques ont mis en évidence le besoin d'étudier davantage les effets et les impacts réels des SIA au sujet de la détection de la fraude fiscale par les AF. Selon Zhou (2019), les AF chinoises devraient estimer l'impact des nouvelles technologies et des nouveaux produits sur le système fiscal. Pour Kumar et al. (2023), les SIA constituent un avantage pour la mise en œuvre d'outils pour la détection de la fraude liée à la taxe sur les produits et services (TPS), mais cela nécessite des réflexions supplémentaires approfondies afin de répondre aux exigences des régulateurs et des entreprises. Aussi, Pérez López et al. (2019) pensent que la technologie qu'ils ont proposée pourrait avoir un impact sur d'autres types de taxes. Cette avenue devrait être analysée. Finalement, Rahimikia et al. (2017) estiment que l'objectif premier du SIA est de fournir une plus grande précision dans la détection de la fraude fiscale au niveau national. Pour les auteurs, l'analyse des effets doit s'effectuer dans le cadre d'une autre recherche.

Bien qu'il existe plusieurs études sur l'utilisation des SIA pour la détection des comportements fiscaux frauduleux, les études scientifiques en la matière sont disparates. L'accent est souvent mis sur des cas spécifiques d'impôts directs ou indirects (la TPS et les impôts sur le revenu des particuliers et des entreprises) et de modèles particuliers de SIA comme les réseaux neuronaux artificiels, les régressions aléatoires, etc. En effet, il a été constaté une lacune de synthèse générale

menant à une exploration et une cartographie structurée des différents SIA en usage dans les AF, leurs performances ainsi que les résultats potentiels.

Malgré l'implémentation des SIA de façon croissante par les AF dans leur processus opérationnels, peu d'études de revue de la littérature ont été menées afin de fournir une comparaison des modèles applicables aux types de fraudes et une évaluation des effets et des impacts. Avec la croissance des stratagèmes fiscaux, il est essentiel de connaître davantage les SIA utilisés pour détecter la fraude. Cela permettrait aux gouvernements de mieux cerner les caractéristiques des données fiscales et financières utilisées, les types de SIA applicables aux comportements fiscaux frauduleux, la mesure de la performance des SIA. Ainsi, les AF pourront mieux appréhender les effets et les impacts sur les contrôles fiscaux et le recouvrement fiscal.

### **1.1 Objectifs de la recherche**

Pour pallier le manque de cartographie de l'utilisation des SIA par les AF pour la détection des comportements fiscaux frauduleux, cette recherche a été initiée pour répondre à cette problématique au moyen d'une RP qui paraît plus appropriée. Selon Pollock et al. (2024), contrairement à la revue systématique, la RP fournit une identification, une clarification, une cartographie et un compte rendu de l'étendue des preuves disponibles sur un sujet, un domaine, un concept ou une question particulière. Cette RP vise à identifier les principales tendances, les lacunes existantes et une exploration des recherches futures.

Dans cette optique, l'objectif de ce travail de recherche est de combler les lacunes en matière de connaissances relatives à l'utilisation des SIA par les AF en vue de détecter la fraude, l'évasion fiscale ainsi que l'évitement fiscal. La question de recherche a deux volets : une question principale et une secondaire.

Question principale : **Quelle est l'étendue des applications des systèmes d'intelligence artificielle dans la détection des comportements fiscaux frauduleux par les AF ?** Question secondaire : **Quels sont les effets et les impacts organisationnels de l'utilisation des SIA par les AF dans la détection des comportements fiscaux frauduleux ?**

## **CHAPITRE 2 : CADRE THÉORIQUE ET REVUE DE LA LITTÉRATURE**

---

Ce chapitre présente les principaux concepts qui orienteront cette étude ainsi qu'une revue de la littérature en lien avec le cadre théorique. D'abord, la population étudiée dans cette recherche est relative aux AF dont la typologie sera analysée. Ensuite, un état des lieux des méthodes de SIA utilisés en général dans la littérature sera fait. Cependant, l'objectif de cette étude n'est pas de présenter une liste exhaustive des SIA existants. La dernière partie de ce chapitre sera consacrée à l'état de la question qui présentera les SIA utilisés pour la détection des comportements fiscaux frauduleux, objet de l'intervention.

### **2.1 La typologie des autorités fiscales (AF)**

Selon Faúndez-Ugalde et al. (2020, p. 1), les AF ont pour « fonction principale de gérer la conformité fiscale afin de détecter et de prévenir les comportements criminels et de fournir des services et une éducation aux contribuables ». La conformité fiscale fait référence à l'observation des obligations fiscales. Pour Franzoni (1998, p. 55), « la conformité fiscale consiste en la déclaration exacte de la base d'imposition, au calcul correct de la taxe à payer, au dépôt de la déclaration dans les délais et au paiement des montants dus dans les délais ».

En ce qui concerne les entités publiques chargées de l'administration de l'impôt, l'OCDE (2023) propose la typologie suivante : les administrations supranationales (les pays de l'UE uniquement), les administrations centrales ou nationales (ARC ou l'*Internal Revenue Service* des États-Unis d'Amérique, *IRS*), les administrations d'États fédérés/Régions (l'Agence du revenu du Québec par exemple), les administrations locales (les collectivités régionales ou provinciales et les collectivités locales) et les administrations de sécurité sociale (caisses de sécurité sociale ou les régimes d'assurance privée).

L'appellation des administrations nationales fiscales diffère selon le contexte d'État unitaire ou État fédéral. Par exemple, l'entité gouvernementale en France est appelée la Direction générale des finances publiques (DGFIP) alors qu'elle est nommée Agence du revenu pour le Canada, etc. Pour des raisons pratiques, le terme d'Autorité fiscale (AF) est utilisé tout au long de cette étude pour désigner les Administrations fiscales nationales qui se distinguent par leur mission principale de collecte des impôts directs et indirects.

Les fonctions principales d'une AF sont : l'identification des contribuables, la gestion des déclarations fiscales, l'observation, la vérification, le recouvrement et la gestion des contentieux. En effet, les AF contrôlent la conformité fiscale des obligations fiscales des contribuables (entreprises et individus) en ce qui concerne les impôts directs et indirects (OCDE, 2023).

Pour Atkinson (1977), les impôts directs sont ceux qui peuvent être ajustés aux caractéristiques individuelles du contribuable (exemple : impôt sur le revenu des particuliers et des entreprises). Les impôts indirects sont constitués essentiellement des taxes sur la consommation. Ils sont prélevés sur les transactions, indépendamment de la situation du client ou du fournisseur ; par exemple, la TPS. Selon l'OCDE (2023), « les pays ont centralisé la collecte des impôts directs et (de la plupart) des impôts indirects au sein d'un seul et même service de l'AF ».

Le rapport sur les statistiques des recettes publiques de l'OCDE (2023) précise que le principal impôt direct collecté par les AF est l'impôt sur le revenu des particuliers et des entreprises alors que l'impôt indirect le plus collecté par les AF est la taxe sur la consommation (taxe sur la valeur ajoutée (TVA) ou taxe sur les biens et services (TPS)). Dans les faits, selon les statistiques de l'OCDE (2023), les AF nationales recouvrent près de 100 % des impôts sur le revenu des personnes physiques et des sociétés et environ 93 % des taxes indirectes.

Dans le cadre de cette étude, seules les AF nationales seront considérées. Pour cause, selon le rapport de l'OCDE, « le seul cas d'autorité supranationale dans les pays de l'OCDE est celui des institutions de l'Union européenne (UE) ». Dans ce contexte, les autorités supranationales n'administrent pas l'impôt. Les administrations fiscales nationales versent une part des impôts collectés aux organisations supranationales.

Par ailleurs, les administrations fiscales d'États fédérés/Régions comme l'Agence du revenu du Québec administrent l'impôt, mais leur accès aux données financières transfrontalières est souvent limité. Leur périmètre d'action se limite à la province, car elles ne collectent que les informations financières en lien avec les contribuables de la province. Par exemple, en ce qui concerne la collaboration fiscale, le Québec a signé un seul accord avec la France pour éviter la double imposition (Revenu Québec, s.d.). Même si ces administrations fiscales utilisent les technologies avancées comme les SIA, et administrent des impôts indirects et directs, elles ne disposent pas toujours des ressources informationnelles internationales nécessaires pour alimenter les

algorithmes. Ainsi, « les institutions financières canadiennes sont tenues de déclarer à l'autorité fiscale fédérale (et non aux AF provinciales) certains renseignements sur les titulaires de comptes non-résidents » en vertu de la Norme commune de déclaration (Agence du revenu du Canada, 2016).

Quant aux administrations de sécurité sociale, elles perçoivent des cotisations obligatoires venant des particuliers et des entreprises. Selon l'OCDE, les administrations publiques qui fournissent de telles prestations de sécurité sociale se voient affecter ces montants afin de financer les prestations sociales. En se référant à la classification des impôts de l'OCDE, ces cotisations sont aussi traitées comme des impôts lorsqu'elles sont perçues par une institution publique. Cependant, les administrations de sécurité sociale n'ont pas pour vocation de collecter les impôts directs et indirects comme indiqué dans cette section.

Finalement, les administrations fiscales locales (les municipalités) investissent plus souvent dans les solutions de SIA afin d'avoir une plus grande visibilité sur les propriétés, les services et une amélioration de leurs performances pour une meilleure clarté dans les prévisions budgétaires (Altus Group, s.d.). Par exemple dans la province de l'Ontario au Canada, les municipalités utilisent les SIA pour les processus d'évaluation des propriétés, améliorant ainsi les impôts fonciers (Rapport annuel 2023, Société d'évaluation foncière des municipalités). Au Brésil, l'apprentissage automatique a été implanté dans le but de prédire les infractions fiscales liées à la taxe sur les services dans le système fiscal de la municipalité de São Paulo (Ippolito et Lozano, 2020).

Le tableau 1 ci-dessous fournit une synthèse de la particularité des administrations responsables de l'impôt.

Les raisons évoquées plus haut placent les AF nationales au premier plan dans la lutte contre les comportements fiscaux frauduleux à l'échelle nationale et internationale.

### **Tableau 1 : Caractéristiques générales des AF**

Autorités fiscales	Champ d'application et pouvoirs fiscaux	Exemples
Les administrations supranationales	Pas d'administration de l'impôt	Les pays de l'UE uniquement au sein de l'OCDE
Les administrations centrales ou nationales	<ul style="list-style-type: none"> <li>• Contrôle de la conformité fiscale des entreprises et individus : impôts directs et indirects</li> </ul>	<ul style="list-style-type: none"> <li>• Canada : ARC</li> <li>• États-Unis d'Amérique : IRS</li> <li>• France : Direction générale des finances publiques (DGFIP)</li> </ul>
Les administrations d'États fédérés/Régions	<ul style="list-style-type: none"> <li>• Pas d'accès aux données financières transfrontalières.</li> <li>• Limite du périmètre d'action à la province</li> </ul>	L'Agence du revenu du Québec
Les administrations locales	<p>Contrôle pour une plus grande visibilité sur :</p> <ul style="list-style-type: none"> <li>• les propriétés,</li> <li>• les services et</li> <li>• l'amélioration des performances pour une meilleure clarté des prévisions budgétaires</li> </ul>	Les collectivités régionales ou provinciales et les collectivités locales
Les administrations de sécurité sociale	<ul style="list-style-type: none"> <li>• Cotisations obligatoires des particuliers et des entreprises : financement des prestations sociales</li> </ul>	Caisses de sécurité sociale

## 2.2 L'utilisation des SIA pour la détection des comportements fiscaux frauduleux

Les AF collectent, stockent et utilisent de grands ensembles de données fiscales. Selon l'OCDE (2022, p. 24), « 70 % des AF déclarent qu'elles utilisent déjà des techniques de pointe comme les SIA pour exploiter les données de manière à découvrir des actifs précédemment cachés ou à identifier de nouveaux risques » (source).

La multiplication des sources de données et de renseignements contraint les AF à abandonner progressivement les méthodes traditionnelles de détection des comportements fiscaux frauduleux au profit des nouvelles technologies plus avancées. L'introduction des SIA pour détecter la fraude aide les AF à assurer une surveillance accrue des grandes entreprises multinationales. Zhu et al. (2018, p. 1225) affirment que ces « méthodes de sélection informatisées des cas sont considérées comme l'approche la plus prometteuse et la plus complète, utilisée par les AF pour détecter la fraude fiscale ». Au vu des différents arguments avancés, ces deux concepts sont analysés afin de répondre à l'objectif de recherche.

### **2.2.1 L'intelligence artificielle**

Selon l'*International Association of Privacy Professionals* (2023, p. 2), « la définition de l'IA est en constante évolution, ses contours et ses contraintes changeant pour s'aligner sur les progrès technologiques et les contextes culturels actuels et peut-être futurs ». Depuis lors, la définition de l'IA et les SIA sont fournis en fonction des perspectives gouvernementales, industrielles, technologiques, légales et académiques.

L'IA fait référence à :

« la capacité d'un système informatique à présenter un comportement intelligent semblable à celui de l'homme, caractérisé par certaines compétences essentielles, notamment la perception, la compréhension, l'action et l'apprentissage. Dans cette optique, nous entendons par application de l'IA l'intégration de la technologie de l'IA dans un domaine d'application informatique avec interaction homme-ordinateur et interaction des données » (Wirtz et al., 2019, p. 599, notre traduction).

### **2.2.2 Les systèmes d'intelligence artificielle (SIA)**

Des logiciels et des algorithmes sont requis afin de mettre en œuvre l'IA. Jacob et Lawarée (2022, p. 7) qualifient ces technologies de SIA. Ils affirment qu'« en pratique, les SIA n'incluent pas les activités ou services périphériques concourant au développement de l'IA ». Une définition qui revêt une portée significative est fournie par l'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO, 2021, p. 10).

« Les systèmes d'IA sont des technologies de traitement des informations qui intègrent des modèles et des algorithmes, lesquels génèrent une capacité d'apprentissage et d'exécution de tâches cognitives conduisant à des résultats tels

que l'anticipation et la prise de décisions dans des environnements matériels et virtuels. Les systèmes d'IA sont conçus pour fonctionner avec différents degrés d'autonomie, au moyen de la modélisation et la représentation des connaissances, de l'exploitation des données et du calcul de corrélations. Ils peuvent intégrer plusieurs méthodes, telles que, sans s'y limiter :(i) l'apprentissage automatique, y compris l'apprentissage profond et l'apprentissage par renforcement ; (ii) le raisonnement automatique, y compris la planification, la programmation, la représentation des connaissances et le raisonnement, et la recherche et l'optimisation ».

Selon cette définition, les SIA peuvent intégrer plusieurs branches, dont l'apprentissage automatique. Rahman et al. (2019, p. 536) estiment que l'apprentissage automatique est au cœur de la science des données et qu'il est situé à l'intersection de l'IA et des statistiques. Il est composé de l'apprentissage supervisé, non supervisé, l'apprentissage semi-supervisé et l'apprentissage par renforcement. En tant que sous-domaine de l'IA, l'apprentissage automatique est une discipline dans laquelle les ordinateurs construisent des modèles statistiques probabilistes basés sur des données et les utilisent pour prédire et analyser des données (Li et al., 2024).

Plus précisément, les travaux dans le domaine de l'apprentissage automatique visent à extraire des informations d'un ensemble de données (images, enregistrements de capteurs, texte, etc.) afin de résoudre un problème lié à ces données (classification, reconnaissance, etc.). (Ebers et Tupay, 2023)

Tout d'abord, selon Li et al. (2024), l'apprentissage supervisé a pour objectif principal de comprendre la loi statistique entre les variables et de détecter par lui-même la correspondance entre les données d'entrée et une prédiction de sortie. Quelques techniques d'algorithme d'apprentissage supervisé sont : <sup>2</sup> Arbres de décision, Régression linéaire, Machine à vecteurs de support (*SVM*), Forêt aléatoire, etc.

Ensuite, l'apprentissage non supervisé vise à apprendre les lois statistiques ou les structures potentielles présentes dans les données. Par ailleurs, selon Bal (2022, p. 228), l'apprentissage non supervisé ne génère pas de prédiction. Le type le plus courant d'apprentissage non supervisé utilisé est le regroupement (*Clustering*) c'est-à-dire le regroupement de cas identiques et qui peut être utilisé par les administrations fiscales pour l'identification de la fraude. Par exemple, l'ARC utilise

---

<sup>2</sup> Tous les algorithmes cités dans ce document ont été traduits de l'anglais par ChatGPT.

cette méthode pour effectuer des analyses par groupement. Cela consiste à repérer des sous-groupes semblables par rapport à d'autres groupes dans une population (Agence du revenu du Canada, s.d.). Voici quelques exemples d'algorithmes liés à l'apprentissage non supervisé : Classification par k-moyennes, Modèles de mélange gaussien, Sélection basée sur le modèle, etc.

Quant à l'apprentissage par renforcement, il est défini par Reis et al. (2024) comme « un agent qui doit prendre une mesure dans un environnement donné. L'agent évalue le comportement optimal en fonction de la récompense immédiate. L'objectif est d'augmenter la récompense à long terme et de réduire le risque ». Quelques algorithmes : Techniques de Monte-Carlo, Apprentissage Q, Apprentissage R, etc.

Enfin, l'apprentissage semi-supervisé (hybride ou mixte) combine l'apprentissage supervisé et non supervisé. Voici quelques exemples d'algorithmes : Classificateur naïf bayésien, Régression logistique, Méthode des k plus proches voisins (*k-NN*), *XGBoost*, etc.

L'efficacité des SIA est mentionnée par Huang (2018). Selon lui, le déploiement des SIA dans les AF rend la collecte et la gestion des données fiscales plus systématiques et transparentes. L'intégration des SIA contribue à réduire l'évasion fiscale et la planification fiscale abusive des multinationales en améliorant les opérations des AF. Le tableau 2 ci-dessous fournit un panorama de ces technologies.

**Tableau 2 : Panorama des SIA**

Technologies	Définition	Types d'apprentissages et d'algorithmes	Source
Intelligence artificielle	L'IA fait référence à la capacité d'un système informatique à présenter un comportement intelligent semblable à celui de l'homme, caractérisé par certaines compétences essentielles, notamment la perception, la compréhension, l'action et l'apprentissage. Dans cette optique, nous entendons par application de l'IA l'intégration de la technologie de l'IA dans un domaine d'application informatique avec interaction homme-ordinateur et interaction des données.	<ul style="list-style-type: none"> <li>• Traitement du langage naturel (TLN)</li> <li>• Apprentissage automatique</li> </ul>	(Wirtz et al.,2019)
Système d'intelligence artificielle	Les systèmes d'IA sont des technologies de traitement des informations qui intègrent des modèles et des algorithmes, lesquels génèrent une capacité d'apprentissage et d'exécution de tâches cognitives conduisant à des résultats tels que l'anticipation et la prise de décisions dans des environnements matériels et virtuels. Les systèmes d'IA sont conçus pour fonctionner avec différents degrés d'autonomie, au moyen de la modélisation et la représentation des connaissances, de l'exploitation des données et du calcul de corrélations.	<ul style="list-style-type: none"> <li>• Apprentissage automatique</li> <li>• Assistants virtuels</li> </ul>	(UNESCO, 2021)
Apprentissage automatique	Plus précisément, les travaux dans le domaine de l'apprentissage automatique visent à extraire des informations d'un ensemble de données	<ul style="list-style-type: none"> <li>• Apprentissage supervisé</li> <li>• Apprentissage non supervisé</li> </ul>	(Ebers et Tupay, 2023)

Technologies	Définition	Types d'apprentissages et d'algorithmes	Source
	(images, enregistrements de capteurs, texte, etc.) afin de résoudre un problème lié à ces données (classification, reconnaissance, etc.)	<ul style="list-style-type: none"> <li>• Apprentissage par renforcement</li> <li>• Apprentissage semi-supervisé</li> </ul>	
Apprentissage supervisé	L'objectif principal est de comprendre la loi statistique qui régit la correspondance entre les données d'entrée et de sortie.	<ul style="list-style-type: none"> <li>• Arbres de décision</li> <li>• Régression linéaire</li> <li>• Machine à vecteurs de support (SVM)</li> </ul>	(Li et al., 2024) (Reis et al., 2024)
Apprentissage non supervisé	Il vise à apprendre les lois statistiques ou les structures potentielles présentes dans les données.	<ul style="list-style-type: none"> <li>• Classification par k-moyennes</li> <li>• Modèles de mélange gaussien</li> <li>• Sélection basée sur le modèle, etc.</li> </ul>	
Apprentissage par renforcement	Il apprend la décision séquentielle optimale entre le système intelligent et l'environnement.	<ul style="list-style-type: none"> <li>• Techniques de Monte-Carlo</li> <li>• Apprentissage Q</li> <li>• Apprentissage R</li> </ul>	(Reis et al., 2024)
Apprentissage semi-supervisé	Il combine l'apprentissage supervisé et l'apprentissage non supervisé (par exemple, dans des situations où les données collectées contiennent à la fois des sorties étiquetées et non étiquetées).	<ul style="list-style-type: none"> <li>• Classificateur naïf Bayésien</li> <li>• Régression logistique</li> <li>• Méthode des k plus proches voisins (k-NN)</li> <li>• <i>XGBoost</i>, etc.</li> </ul>	

## **2.3 Les comportements fiscaux qualifiés de frauduleux**

En matière fiscale, les contribuables peuvent utiliser plusieurs stratagèmes en vue de minimiser leurs revenus et l'impôt à payer. Pour y arriver, ils adoptent tantôt des comportements illégaux, qualifiés de fraude fiscale ou d'évasion fiscale, tantôt des comportements légaux, qualifiés d'évitement fiscal (Saxunova et Szarkova, 2018, p. 4). Cependant, lorsque le contribuable détourne la lettre de la loi fiscale à son avantage, l'OCDE considère l'évitement fiscal comme de la planification fiscale abusive illégale. Selon la littérature, les comportements illégaux sont davantage portés sur les impôts sur le revenu et les taxes sur la consommation (De La Feria, 2020 ; Saragih et Ali, 2023 ; Wang et al., 2020).

Plusieurs auteurs ont établi un lien entre les comportements fiscaux illégaux, et le blanchiment d'argent et les fraudes financières en général (Ioana et al., 2021 ; Deen et al., 2022).

### **2.3.1 La fraude fiscale**

Le concept de fraude fiscale se définit comme :

« Toute action du contribuable qui implique une violation de la loi, lorsqu'on peut prouver que l'intéressé a agi dans le dessein délibéré d'échapper à l'impôt. On distingue parfois la faute moins grave, que constitue l'omission - par exemple, le fait de remplir une déclaration de revenus incomplète- des infractions plus graves, comme les fausses déclarations, les factures truquées, etc. (OCDE 1987, p. 17) ».

Selon Vernier (2018), que l'impôt soit dû ou exigible, le fait de ne pas le déclarer ou le payer constitue une fraude fiscale dont la détection constitue un enjeu pour les AF. De La Feria (2020) propose une typologie de la fraude fiscale. L'auteur distingue la fraude fiscale liée à l'évasion et celle liée à la fraude organisée. Selon l'auteur, l'évasion implique principalement les petites et moyennes entreprises nationales qui profitent des limites administratives et de la fiscalité intérieure, tandis que la fraude organisée est le fait de groupes criminels opérant à l'échelle transnationale, exploitant les failles législatives fiscales en matière de commerce transfrontalier.

Par exemple, dans le cas des impôts indirects, l'entreprise étant un intermédiaire qui collecte la taxe pour le compte des AF, les sources de fraudes sont diverses. Les plus importantes sont mises en exergue par De La Feria (2020) : la création de fausses entreprises dans le but de vendre des factures falsifiées et d'effectuer des ventes et des demandes de remboursements de taxes fictives,

le non-reversement de la taxe collectée à l'État à travers une fausse comptabilité, une déclaration de faillite technique et les détournements de numéro de taxes.

En ce qui concerne les impôts directs, plusieurs comportements frauduleux sont révélés par Boita, Paiusan et Ajtay (2021, p. 700) : « la destruction des pièces justificatives financières et comptables, la tenue d'une double comptabilité, l'établissement de bilans comptables avec des données altérées, la non-déclaration partielle ou totale des activités commerciales ou des sources d'imposition et l'élaboration ou l'altération de documents reflétant des opérations d'import-export ».

### **2.3.2 L'évasion fiscale**

Certains citoyens et des entreprises sont parfois réticents aux paiements des impôts et des taxes diverses (Saxunova et Szarkova, 2018). Pour ces auteurs, l'évasion fiscale représente un véritable enjeu pour les économies nationales, car elle a un effet néfaste sur le budget de l'État et en particulier sur les finances publiques. Plusieurs auteurs ont défini les caractéristiques de l'évasion fiscale sous divers angles. Selon Alm (2012), le concept d'évasion fiscale fait référence aux actions illégales et intentionnelles des contribuables pour réduire les impôts à déclarer et à payer. L'évasion fiscale consiste également dans la dissimulation des obligations fiscales légalement dues, quel que soit le secteur d'activités commerciales.

En revanche, Vernier (2018) considère l'évasion fiscale comme une attitude visant la réduction du montant des prélèvements obligatoires, mais sans l'existence de l'élément intentionnel soulevé par Alm (2012). Cela signifie que le contribuable se sert de techniques légales pour payer moins d'impôts. Pour lui, c'est ce critère intentionnel qui diffère la fraude fiscale de l'évasion fiscale. En résumé, pour l'auteur, l'évasion fiscale s'apparente à la fraude lorsque le contribuable « s'appuie sur des techniques illicites ». L'ARC fournit des exemples de techniques illicites : sous-déclarer des revenus imposables ou déduire des dépenses non-déductibles ou surestimées. Les participants à l'évasion fiscale peuvent également tenter d'éluder l'impôt en refusant volontairement de se conformer aux exigences légales en matière de déclaration (Agence du revenu du Canada, 2016). Le ministère des Finances du Québec, dans son rapport sur le phénomène du recours aux paradis fiscaux (ministère des Finances du Québec, 2017), devant l'Assemblée nationale du Québec, indique que les stratagèmes d'évasion fiscale comprennent : la non-déclaration des revenus légaux (travail au noir), la dissimulation de revenus illégaux (le commerce illégal de drogue ou d'armes)

et la désobéissance aux règles fiscales (demandes abusives de remboursement et de crédits d'impôt). Les personnes concernées peuvent faire appel aux paradis fiscaux afin de rendre les stratagèmes très complexes.

L'organisme international *Tax Justice Network* indique que le paradis fiscal « fait généralement référence à un pays ou une juridiction qui permet aux sociétés multinationales et aux particuliers d'échapper à l'État de droit dans les pays où ils opèrent et vivent, et de payer moins d'impôts qu'ils ne le devraient dans ces pays ». À titre d'exemple, les entreprises multinationales utilisent des techniques très avancées, des méthodes comptables complexes et des spécialistes dans divers domaines (finances, technologies...) pour se soustraire de leurs obligations fiscales (Zhu X. et al., 2018, p. 1224). Cette pratique rend le travail de la détection de l'évasion plus difficile par les méthodes classiques. Mais qu'en est-il de l'évitement fiscal ?

### **2.3.3 L'évitement fiscal et la planification fiscale abusive**

La littérature scientifique différencie la notion d'évitement fiscal de l'évasion fiscale. Selon Barilari A. (2018, p. 51), l'évitement fiscal est un « comportement qui n'implique pas en soi de la réprobation ». L'auteur utilise la notion de fuite fiscale pour les personnes physiques et celle de délocalisation pour les activités économiques des multinationales qui s'installent dans des paradis fiscaux afin de profiter des baisses d'impôts. Plusieurs expressions désignent l'évitement fiscal : la fuite fiscale, l'exil fiscal, la délocalisation fiscale, la gestion astucieuse d'une situation fiscale, la soustraction fiscale, l'optimisation fiscale, etc. (Barilari, 2018) L'évitement fiscal inclut divers phénomènes tels que le recours aux conventions fiscales internationales, le transfert de dettes, le report d'impôts, la manipulation des prix de transfert (Baumann et al., 2020), ainsi que l'exploitation des failles législatives, le transfert de revenus entre différentes zones géographiques, la restructuration de l'entreprise et le recours aux paradis fiscaux. (Ftouhi et Ghardallou, 2020)

Pour Wang et al. (2020, p. 793), l'évitement fiscal est composé de la planification fiscale légale et de l'évasion fiscale illégale. Ils fournissent d'abord une définition de la planification fiscale. Pour les auteurs, « elle consiste à réduire la charge fiscale d'une entreprise en investissant et en structurant les activités commerciales dans les limites de la législation fiscale ». Ce comportement n'implique pas la dissimulation et la tromperie. Cockfield (2015) confirme cette assertion en

indiquant qu'aucun crime n'a été commis et le contribuable ne peut être emprisonné. Dans ce cas, le contribuable est condamné à payer des droits plus élevés, des pénalités et des amendes.

Sur un aspect plus pratique, l'Agence du revenu du Canada (s.d.), estime que : « l'évitement fiscal et la planification fiscale ont en commun des arrangements visant à réduire les impôts et qui sont fondés sur le libellé précis de la réglementation pertinente ». Quand les arrangements respectent l'esprit de la loi, l'ARC les qualifie de « planification fiscale efficace ». À l'opposé, l'ARC emploie le concept de planification fiscale abusive pour désigner les stratégies intérieures et internationales qui vont au-delà d'une planification fiscale efficace. Finalement, la partie XVI de la *Loi de l'impôt sur le revenu. L.R.C. (1985), ch. 1 (5e suppl.)* (LIR) et l'article 274 de la *Loi sur la taxe d'accise (L.R.C. (1985), ch. E-15)* (LTA) du Canada, renferment des dispositions visant à contrer la planification fiscale abusive. La section suivante examine de façon comparative les concepts de fraudes discutées.

#### **2.3.4 Analyse comparée des concepts de fraude en matière fiscale**

Dans cette section, il est procédé à une analyse comparée des différents concepts discutés que sont, la fraude fiscale, l'évasion fiscale, l'évitement fiscal et la planification fiscale abusive.

Par définition, l'évitement fiscal est légal, mais il est qualifié de planification fiscale abusive lorsque les arrangements des contribuables vont au-delà de ce qui est prévu par la loi. De nombreuses activités des entreprises peuvent être considérées simultanément comme de l'évasion fiscale et de l'évitement fiscal.

Les prix de transfert en sont un exemple. Ce sont les « prix auxquels une entreprise transfère des actifs corporels ou incorporels ou fournit des services à une société affiliée », Guias et Haineala (2021). De façon pratique, « les entreprises facturent des prix bas pour leurs ventes à des sociétés affiliées à faible taux d'imposition, mais paient des prix élevés pour les achats qu'elles effectuent auprès d'elles » (Gravelle, 2009, p. 727). Le phénomène des prix de transfert peut être considéré comme une planification fiscale abusive, car les multinationales « tirent parti des particularités techniques d'un système fiscal ou des disparités entre deux ou plusieurs systèmes fiscaux dans le but de réduire l'impôt à payer » (Shaw et al., 2019, p. 347).

En ce qui concerne la fraude fiscale et l'évasion fiscale, Shaw et al. (2019) précisent que la dissimulation est le facteur déterminant de distinction. L'évasion fiscale consiste en une dissimulation par omission (omission de déclaration du revenu imposable) et la fraude fiscale consiste en une dissimulation par commission (fournir des documents frauduleux relatifs au revenu imposable). Une synthèse des comportements fiscaux frauduleux est présentée dans le tableau 3 ci-dessous.

**Tableau 3 : Synthèse des principaux comportements fiscaux frauduleux**

<b>Comportements frauduleux</b>	<b>Définition et exemples</b>	<b>Caractéristiques</b>
Fraude fiscale ( <i>OECD</i> , 1987)	Toute action du contribuable qui implique une violation de la loi, lorsqu'on peut prouver que l'intéressé a agi dans le dessein délibéré d'échapper à l'impôt.  Exemples : <ul style="list-style-type: none"> <li>• Fraude de l'impôt sur le revenu</li> <li>• Fraude fiscale des entreprises</li> <li>• Fraude à la TVA</li> </ul>	<ul style="list-style-type: none"> <li>• Illégale avec sanctions administratives.</li> <li>• Infraction criminelle selon le Code criminel.</li> <li>• Une dissimulation par commission.</li> <li>• Exemple : une déclaration de faillite technique.</li> </ul>
Évasion fiscale ( <i>Alm</i> , 2012)	Fait référence aux actions illégales et intentionnelles entreprises par des contribuables pour réduire ou cacher leurs obligations fiscales légalement dues, quel que soit le secteur d'activités commerciales  <ul style="list-style-type: none"> <li>• Exemple : évasion fiscale des entreprises</li> </ul>	<ul style="list-style-type: none"> <li>• Illégale avec sanctions administratives.</li> <li>• Une dissimulation par omission.</li> <li>• Exemple : sous-déclarer des revenus imposables.</li> </ul>
Évitement fiscal ( <i>Barilari</i> , 2018)  ( <i>Wang et al.</i> , 2020)	Un comportement qui n'implique pas en soi de la réprobation.  L'évitement fiscal est composé de la planification fiscale légale et de l'évasion fiscale illégale.	<ul style="list-style-type: none"> <li>• Légale s'il n'y a pas d'abus de la loi.</li> <li>• Respect de la lettre et de l'esprit de la loi.</li> <li>• Exemple : les prix de transfert.</li> </ul>

Comportements frauduleux	Définition et exemples	Caractéristiques
	Exemple : évitement fiscal des entreprises	
Planification fiscale abusive (Wang et al., 2020)	Des arrangements ne respectant pas l'esprit du libellé précis de la loi et qui vise à réduire ou à ne pas payer les impôts.	<ul style="list-style-type: none"> <li>• Illégale et sanctions administratives.</li> <li>• Non-respect de l'esprit de la loi.</li> <li>• Exemple : les prix de transfert.</li> </ul>

La plupart des législations fiscales nationales punissent les comportements fiscaux frauduleux. Au Canada, les sanctions administratives (pénalités) sont régies par les paragraphes 163 (2) et 239 (1) de la LIR et les articles 285 et 326 (1) de la LTA, en ce qui concerne la fraude fiscale. L'article 380 du *Code criminel* (L.R.C. (1985), ch. C-46) du Canada considère la fraude fiscale comme une infraction criminelle. Quant à l'évasion fiscale, elle est sanctionnée par l'article 239 de la LIR. La partie XVI de la LIR et l'article 274 de la LTA, eux, contiennent également des dispositions visant à contrer la planification fiscale abusive (Easter, 2016). En revanche, en Île-Maurice, les contribuables pris sous le coup de la Partie VII de la Loi mauricienne de l'impôt sur le revenu (*The Income Tax Act, 1995*) doivent seulement payer le montant d'impôt dû sans mesures dissuasives visant à empêcher la récurrence des pratiques d'évitement fiscal futures (Beebeejaun, 2021).

Malgré la mise en œuvre de ces mécanismes législatifs nationaux de lutte contre la fraude, plus de 140 pays et juridictions collaborent à la mise en œuvre de quinze mesures visant à lutter contre l'évasion fiscale au sein du Cadre inclusif OCDE/G20 (OCDE, s.d.) sur les prix de transfert. L'objectif est l'amélioration de la cohérence des règles fiscales (OCDE).

Le tableau 3 fournit un aperçu des différents comportements fiscaux frauduleux et non frauduleux en indiquant leurs principales caractéristiques.

Bien que les AF soient organisées au sein d'instances internationales, la détection des comportements frauduleux demeure un enjeu majeur. Selon Tian et al. (2016, p. 2651), il existe trois méthodes couramment utilisées par les AF pour la détection des comportements fiscaux frauduleux. Il s'agit de « la sélection manuelle des cas, la sélection basée sur les dénonciations et

la sélection informatisée des cas ». Selon les auteurs, « la sélection manuelle des cas et les méthodes de sélection basées sur les dénonciations prennent beaucoup de temps et sont fastidieuses ». Cependant, les contrôles informatisés semblent plus efficaces.

Le choix de l'inclusion des notions de fraude fiscale, évitement fiscal et évasion fiscale dans cette étude répond au besoin de prendre en considération les principaux comportements fiscaux frauduleux, car ces concepts sont les plus répandus. Malgré leur distinction sémantique et opérationnelle, ces trois concepts sont interdépendants et représentent des enjeux considérables pour les administrations fiscales. De plus, la prise en compte de ces trois concepts dans la même étude permet d'obtenir une évaluation efficace des SIA dans la détection des comportements frauduleux et leur traitement. Enfin, cette approche intégrative contribue à enrichir la compréhension globale des enjeux fiscaux à l'ère des technologies avancées.

## **2.4 État de la question**

Comme le définissent De Roux et al. (2018), la fraude fiscale est l'acte intentionnel qui consiste à produire une fausse déclaration de revenus dans le but de réduire ses obligations fiscales. De nos jours, les pratiques de fraude fiscale sont de plus en plus sophistiquées (Zheng et al., 2023). Ainsi, la détection de ces pratiques est nécessaire pour garantir la capacité des États à assurer les services essentiels. L'enjeu de la détection de la fraude fiscale est majeur pour les AF qui gèrent la conformité fiscale, détectent et préviennent les comportements criminels (Faúndez-Ugalde et al., 2020) des entreprises, des groupes de contribuables et des individus.

Depuis une dizaine d'années, les SIA sont développés et mis en œuvre par les États pour améliorer la détection de la fraude fiscale. Dans la littérature scientifique, plusieurs recherches ont été publiées sur cette problématique. Un aperçu général de quelques cas d'utilisation des SIA pour détecter les comportements fiscaux frauduleux est présenté.

Tout d'abord, Latulippe et Martel (2021) ont répertorié plusieurs études ayant identifié l'utilisation des SIA par les AF pour détecter les pratiques de fraude fiscale et d'évitement fiscal entre 1999 et 2020. Entre 1999 et 2017, on constate que la détection de la fraude s'appuie sur des méthodes et des applications de SIA. Cependant, en 2018, les avancées technologiques ont permis une

diversification et une sophistication des modèles d'IA ou d'apprentissage automatique pour la détection de la fraude fiscale.

#### **2.4.1 L'intelligence artificielle**

Raikov (2021) propose un système d'aide à la décision pour la prédiction et le diagnostic des cas d'évasion fiscale en s'appuyant sur des technologies d'IA. Cette approche intègre des méthodes de réseaux neuronaux profonds, de modélisation cognitive et d'analyse des mégadonnées. Enrichie par des éléments de sémantique cognitive non formalisables, la solution vise à accroître la transparence des événements fiscaux des entreprises. Cette technique a été utilisée afin de faciliter la détection et l'intervention dans des contextes tels que la fourniture de produits dans la zone arctique de la Russie. La limite de cet article est qu'il ne précise pas la nature de l'AF en Russie ayant appliqué cette technologie.

En Inde, Rathi et al. (2021) ont expérimenté un modèle d'IA incluant des techniques de traitement du langage naturel, de reconnaissance d'images et de reconnaissance vocale dans le but d'identifier l'évasion fiscale et la corruption au sein d'un département fiscal. Selon les auteurs, les modèles ont accru la transparence du processus fiscal, la découverte de contribuables non déclarés, la prévention de la fraude et de l'évasion fiscale. Cependant, les données utilisées par leurs modèles provenaient d'un sondage et de sources secondaires. De plus, l'étude n'a pas suivi un processus méthodologique rigoureux.

#### **2.4.2 Les méthodes d'apprentissage automatique**

La recherche documentaire a fourni peu d'articles dans lesquels les algorithmes d'intelligence artificielle sont mis en œuvre pour détecter la fraude. En revanche, plusieurs études ont démontré l'usage des techniques d'apprentissage automatique pour la détection de la fraude. Cette section donne un aperçu de la littérature sur l'usage de l'apprentissage automatique.

##### **2.4.2.1 L'apprentissage supervisé**

Comme modèle supervisé, Zheng et al. (2019) ont développé, en Chine, un système analytique visuel appelé *ATTENet*. Il détecte les groupes suspects de fraude fiscale basés sur des transactions affiliées. Ce modèle utilise deux algorithmes que sont : le *Structure2Vec* et la Forêt aléatoire. Les

données fiscales réelles ayant alimenté le modèle proviennent des AF coopératives. *ATTENet* détecte efficacement les groupes suspects en fournissant des explications visuelles avec plus de détails. Cependant, le modèle n'a pas été expérimenté dans une administration fiscale afin de mesurer sa performance.

En Amérique du Sud et précisément au Brésil, Xavier et al. (2022) ont modélisé des données publiques avec les algorithmes de Forêt aléatoire, Réseaux de neurones multicouches et Réseaux de neurones pour graphes (RNG) dans le but de détecter les pratiques d'évasion fiscale. En revanche, le modèle de Forêt aléatoire a atteint une précision de plus de 98 % lorsqu'il a été appliqué à l'administration fiscale de Goiás et non à l'administration fiscale fédérale du Brésil.

Shakil et Tasnia (2022) ont mené une étude en Asie et dans le Pacifique qui met en évidence le fait que l'utilisation des modèles de réseaux de neurones et des machines à vecteurs de support ont contribué à réduire l'évasion fiscale. Bien que les modèles d'apprentissage automatique aient été expérimentés dans des AF en Chine, en Malaisie, à Singapour, en Inde et en Nouvelle-Zélande, l'étude n'indique pas un processus méthodologique détaillé.

Par ailleurs, un modèle de prédiction a été élaboré en Espagne par Pérez López et al. (2019). Ces auteurs espagnols ont analysé un modèle d'outils prédictifs avancés d'apprentissage supervisé, réseau neuronal perceptron multicouche, dont le taux d'efficacité est évalué à 84.3 %. Ce modèle détecte la fraude fiscale dans les déclarations de revenus des individus en ce qui concerne les impôts directs (impôt sur le revenu) et les impôts indirects (la taxe sur les produits et services).

En Thaïlande, Visitpanya et Samanchuen (2023), quant à eux, ont mené leur étude sur une méthode basée sur l'apprentissage automatique pour identifier les comportements suspects d'évasion fiscale en utilisant des données financières publiques. L'approche met en œuvre des modèles d'apprentissage supervisé, Forêt aléatoire et *K-nearest neighbors* (KNN), pour détecter des schémas de fraude à partir de critères de risque prédéfinis. Les résultats obtenus montrent une efficacité de plus de 95 % en matière de précision. Les données utilisées pour cette étude proviennent du « *Stock Exchange of Thailand* », la bourse des valeurs de Thaïlande. Aucun test n'a été effectué par une AF afin de valider son efficacité.

#### **2.4.2.2 L'apprentissage non supervisé**

Quelques études ont démontré l'usage des méthodes d'apprentissage non supervisé afin de détecter les comportements fiscaux frauduleux. En Colombie, De Roux et al. (2018) ont proposé une étude avec l'utilisation de la technique de *Clustering spectral* afin d'identifier les anomalies dans les bases fiscales déclarées. La fraude concernait l'impôt sur la délimitation des zones urbaines à Bogota. Le type de fraude identifiée est la sous-déclaration des revenus avec un impact principal tel qu'une meilleure allocation des ressources pour les vérifications. Cependant, bien que les données soient réelles, elles proviennent du Département des Finances de Bogota, un organisme régional responsable de la gestion de la ville de Bogota.

Wei et al. (2019) ont développé une étude qui fournit une description de l'utilisation des techniques de Réseaux adversariaux conditionnels non supervisés (*UCAN*). Cette méthode a été appliquée sur des données fiscales réelles provenant de cinq provinces en Chine sans l'implication d'une AF. La fraude fiscale identifiée est la fraude par facturation fictive liée à la TVA.

En outre, une étude menée par Herrera-Semenets et al. (2024) sur les données fiscales de l'*Internal Revenue Service (IRS)* des États-Unis d'Amérique a utilisé des techniques algorithmiques basées sur l'apprentissage non supervisé. La technique a consisté en la détection des sous-déclarations des revenus dans l'État de l'Alaska. La méthode a démontré son utilité et son efficacité pour la prise de décision et le ciblage des zones suspectes.

#### **2.4.2.3 L'apprentissage semi-supervisé (hybride)**

Dans une étude théorique, un modèle hybride de réseau neuronal graphique (*Eagle*) (Shi et al., 2023) détecte l'évasion fiscale en intégrant les caractéristiques des contribuables et leurs relations pour extraire des fonctionnalités complètes de détection. Le modèle utilise à la fois un apprentissage supervisé et non supervisé. Pour les scénarios de classification avec des échantillons étiquetés disponibles, l'apprentissage supervisé est employé. À contrario, lorsqu'il s'agit de détecter des anomalies, avec des échantillons étiquetés non disponibles, l'apprentissage non supervisé est utilisé.

Un autre modèle d'apprentissage hybride a été présenté par Mahesar et al. (2025) afin de détecter l'évasion fiscale en matière de e-commerce. Les techniques d'apprentissage supervisé sont le

*XGBoost* et le *LightGBM* tandis que le modèle non supervisé inclut l'isolation Forest et l'*Autoencoder*. Ils ont été combinés pour la détection des sous-déclarations des ventes, la fraude à la TVA transfrontalière et l'utilisation de microtransactions pour éviter les seuils de déclaration. Cependant, les données utilisées ne proviennent pas d'une AF en particulier, en plus, ce ne sont pas des données réelles. Selon l'étude, même si les données sont synthétiques, les résultats sont fiables et l'un des impacts est l'optimisation des audits fiscaux.

#### **2.4.2.4 L'apprentissage par renforcement**

Quant à Goumagias et al. (2018), ils ont développé en Grèce un modèle qui fournit des solutions approximatives grâce à une combinaison d'apprentissage Q et d'apprentissage par renforcement profond pour détecter l'évasion fiscale. Cependant, ce modèle qualifié de « in vitro » est un modèle théorique que les administrations fiscales peuvent appliquer à leurs opérations surtout en Grèce. Au fil des ans, les SIA se sont intégrés dans les processus décisionnels des AF. La méthodologie présentée au chapitre suivant permettra de sélectionner les articles pertinents.

## CHAPITRE 3 : DEVIS MÉTHODOLOGIQUE

---

Ce chapitre présente le devis méthodologique adopté dans le contexte de cette RP. Cette méthodologie présente de façon systématique, les différentes étapes de la RP afin d’approfondir la recherche en répondant aux objectifs énoncés dans la question de recherche.

### 3.1 Processus de la revue de portée

Colquhoun (2014) définit la RP comme :

« une forme de synthèse des connaissances qui aborde une question de recherche exploratoire visant à cartographier les concepts clés, les types de preuves et les lacunes de la recherche liée à un domaine ou à un domaine défini en recherchant, en sélectionnant et en synthétisant systématiquement les connaissances existantes ».

Il recommande une RP en cinq principales étapes:

- L’identification de la question de recherche
- L’identification des études pertinentes
- La sélection des études
- La cartographie des données
- Rassembler, résumer et rendre compte des résultats ainsi que la discussion

#### 3.1.1 Le protocole *PRISMA*

Plusieurs auteurs ont développé des protocoles dans la littérature scientifique afin de conduire des RP. Dans une étude sur la RP, Lawarée et al. (2020) ont mené des recherches qui mettent en exergue une méthodologie de RP dans le domaine de l’évaluation des programmes et des politiques publiques entre 1979 et 2018. Cette étude révèle que de nombreuses revues de portée ou de littérature en évaluation omettent de rapporter avec rigueur des éléments méthodologiques fondamentaux, tels que les critères de sélection des études, les procédures d’évaluation de leur qualité, ou encore la stratégie de synthèse retenue (Lawarée et al., 2020, p. 7).

Pour combler cette lacune méthodologique, cette RP est conduite sur la base du protocole *Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)* (Lawarée, Jacob

et Ouimet 2020). Le protocole PRISMA (Moher et al., 2009) consiste en une liste de contrôle de vingt-sept éléments. Cette liste de contrôle porte sur la définition de la question de recherche, la détermination des sources de données et la stratégie de recherche, les critères d'éligibilité et les processus de sélection (exclusion et inclusion), l'extraction des données, leur analyse et, enfin, la discussion.

L'analyse des différents concepts liés aux comportements fiscaux frauduleux a mis en évidence des outils de pointe que sont les SIA capables de prédire, détecter et diagnostiquer la fraude fiscale, l'évasion fiscale et l'évitement fiscal. Plusieurs revues scientifiques ont été élaborées sur le sujet, mais quels sont les critères d'éligibilité de ces articles ?

### **3.2 Les critères d'éligibilité : la méthode *SPIDER***

Le processus de sélection des articles pertinents est basé sur des critères d'exclusion et d'inclusion préalablement définis. Ce processus a été réalisé en deux étapes : la première était basée sur le résumé, le titre et les mots-clés. La deuxième étape a porté sur une lecture complète de l'article. L'outil *SPIDER* (en anglais *Sample - Phenomenon of Interest – Design - Evaluation et Research type*) et en français, Échantillon, Phénomène d'intérêt, Conception, Évaluation, Type de recherche, a été mis en œuvre afin de déterminer les critères d'éligibilité des mots-clés de la recherche documentaire. (Cooke et al., 2012)

D'abord, les AF, indépendamment de la situation géographique, comme définie dans le cadre théorique, ont constitué l'échantillon d'étude. Les AF supranationales, provinciales, régionales et municipales sont exclues de cette étude.

Ensuite, le phénomène d'intérêt a porté sur l'utilisation des SIA par les AF. Les chaînes de recherche ont inclus les termes IA, SIA, y compris apprentissage automatique, y compris apprentissage profond, apprentissage par renforcement, apprentissage supervisé et apprentissage non supervisé. En ce qui concerne les comportements fiscaux frauduleux, les chaînes incluaient les termes suivants : fraude fiscale, évasion fiscale et évitement fiscal. La recherche a été également inclus les types d'impôt les plus à risque, à savoir les impôts directs et indirects. Le dernier critère était lié aux effets et aux impacts de l'application des SIA dans les activités des AF et la perception des utilisateurs internes. Ces inclusions ont limité le champ de recherche, ce qui a

conduit à l'exclusion des termes d'évasions fiscales, évitements fiscaux, fraudes fiscales dans les entreprises privées et dans les autres administrations chargées de collecter les impôts, qui ne sont pas des AF.

La prochaine étape de l'outil *SPIDER* était la conception relative aux techniques utilisées pour la collecte des données. Les données secondaires et primaires étaient prises en considération. Les articles relatifs à des groupes de discussion, des entretiens et des observations ont été pris en compte.

Pour ce qui est de l'évaluation, les articles basés sur une démarche scientifique dotée d'une méthodologie robuste aboutissant à des résultats éprouvés ont été sélectionnés. En outre, les critères de présentation des indicateurs et des mesures de taux de détection et de prédiction des comportements fiscaux frauduleux ont également été retenus. Ainsi, ont été inclus les articles traitant des impacts, des enjeux, des changements apportés ainsi que l'influence et les conséquences de l'utilisation des SIA sur les comportements fiscaux frauduleux. Finalement, le type de recherche incluait les publications de type qualitatif évaluées par les pairs.

Le tableau 4 ci-dessous fournit une synthèse des critères d'éligibilité selon l'outil *SPIDER* :

<b><i>SPIDER</i></b>	<b>Inclusion</b>	<b>Exclusion</b>
Échantillon ( <i>Sample</i> )	Les AF nationales utilisant principalement les outils de SIA pour lutter contre les différents stratagèmes fiscaux frauduleux. Articles se basant sur les données fiscales d'AF nationales.	Les AF supranationales, provinciales, régionales et municipales
Phénomène d'intérêt ( <i>Phenomenon of Interest</i> )	Les comportements fiscaux frauduleux en jeu : fraude fiscale, fraude à la TVA, fraude à l'impôt sur le revenu des entreprises et des particuliers, évasion fiscale, évitement fiscal, planification fiscale abusive.	Évitements fiscaux, évasions fiscales et fraudes fiscales dans les entreprises privées et dans les autres administrations chargées de collecter les impôts n'étant pas des AF.

<b>SPIDER</b>	<b>Inclusion</b>	<b>Exclusion</b>
	Les types d'impôts : les impôts directs et indirects.	Les taxes municipales, les impôts fonciers et les taxes douanières.
Conception (Design)	L'intervention principale mise en œuvre et se basant sur l'utilisation des SIA en vue de la détection et de la prédiction des comportements fiscaux frauduleux.	
	<ul style="list-style-type: none"> <li>• Intelligence artificielle : <ul style="list-style-type: none"> <li>○ Traitement du langage naturel,</li> <li>○ Les grands modèles de langage,</li> <li>○ Agents virtuels intelligents,</li> <li>○ Algorithme,</li> <li>○ Vision par ordinateur et automatisation</li> </ul> </li> </ul>	
	<ul style="list-style-type: none"> <li>• Apprentissage automatique : <ul style="list-style-type: none"> <li>○ Apprentissage profond</li> <li>○ Apprentissage par renforcement</li> <li>○ Apprentissage supervisé</li> <li>○ Apprentissage non supervisé</li> <li>○ Apprentissage semi-supervisé</li> </ul> </li> </ul>	
Évaluation (Evaluation)	<ul style="list-style-type: none"> <li>• Utilisation d'une démarche scientifique avec une méthodologie robuste évaluant le SIA utilisé et qui aboutit à des résultats éprouvés.</li> <li>• Présentation des indicateurs et des mesures de taux de détection et de prédiction des comportements fiscaux frauduleux.</li> <li>• Résultats évaluant la performance des SIA en plus de leurs effets et impacts sur la détection des comportements fiscaux frauduleux par les AF.</li> <li>• Présentation des effets et des impacts des SIA dans la détection des comportements fiscaux frauduleux au sein des AF</li> </ul>	<p>Exclusion :</p> <ul style="list-style-type: none"> <li>- Les articles qui ne présentent pas de méthodologie d'évaluation du SIA</li> <li>- Les articles qui ne comportent pas des effets et des impacts de l'utilisation des SIA sur les AF</li> </ul>

<b>SPIDER</b>	<b>Inclusion</b>	<b>Exclusion</b>
	<ul style="list-style-type: none"> <li>• La perception de l'utilisation de l'IA par les utilisateurs internes.</li> </ul>	
Type de recherche ( <i>Research Type</i> )	<p>Les recherches qualitatives et mixtes. Les données secondaires et les données primaires : articles de revue scientifique, livres, actes de conférence, sections et chapitres de livre et les rapports. Les articles évalués par les pairs.</p>	<p>Exclusion :</p> <ul style="list-style-type: none"> <li>- Article de journal ou de presse</li> <li>- Thèses</li> <li>- Article de magazine</li> <li>- Génériques</li> <li>- Articles scientifiques non évalués par les pairs et les recherches quantitatives.</li> </ul>

### 3.3 Sources d'information

Cette section décrit les différentes étapes de la détermination des bases de données pertinentes ainsi que des méthodes mises en œuvre pour identifier les articles pertinents.

#### 3.3.1 La détermination des sources des références

À partir des critères d'éligibilité définis, la sélection des articles scientifiques a été réalisée à partir de 8 bases de données (BD) couvrant des disciplines variées comme la fiscalité, l'administration publique, l'économie, les sciences de la gestion et la comptabilité, les sciences de l'information, l'informatique et l'ingénierie et les systèmes comme les SIA. Les critères de recherche étaient limités aux articles scientifiques de langue française et anglaise dans les BD dédiées suivantes afin d'obtenir une liste exhaustive des articles publiés et évalués par les pairs : *Web of science*, *Scopus*, *Scimedirect*, *JSTOR*, *IEEE*, *Emerald Journals*, *EBSCOhost* et *ABIIFORM*. L'identification et le choix des BD ont été le fruit d'une collaboration avec la bibliothèque de l'École nationale d'administration publique (ENAP). Une recherche manuelle complémentaire des revues et des références pertinentes a été effectuée dans *Google Scholar*.

### 3.3.2 Stratégie de recherche

Plusieurs séances de travail ont été organisées avec la bibliothèque de l'ENAP pour élaborer les équations de recherche pertinentes. Afin de procéder à la recherche proprement dite des références, une combinaison de mots-clés a été utilisée à travers deux équations en français et en anglais (voir Annexe A). Cependant, les équations de recherche proprement dites pouvaient différer compte tenu des BD. L'application des combinaisons des équations de recherche n'était pas systématique et a souvent été modifiée pour tenir compte des BD.

Pour l'ensemble des BD consultées, les recherches ont débuté le 9 mai 2024 avec des mises à jour régulières. La dernière mise à jour a été effectuée le 28 août 2024 avec des recherches supplémentaires dans *Google Scholar* le 8 septembre 2024. Aucune restriction de filtres concernant les dates et les types de références n'a été appliquée lors de la recherche, car l'objectif était de collecter l'exhaustivité des études sur une période prolongée. Il a été ainsi obtenu un total de 3 896 références (voir Annexe B).

Finalement, toutes les études identifiées dans le logiciel *EndNote* ont été enregistrées et classées par BD afin d'obtenir des références bibliothécaires respectant les critères de l'APA (*American Psychological Association*).

Pour procéder au filtrage des articles, le choix s'est porté sur le logiciel *DistillerSR*.

### 3.4 Processus de sélection des références avec le logiciel *DistillerSR*

Une évaluation formelle de la qualité des recherches primaires a été faite en vue d'identifier les meilleures données disponibles (Baron, 2018) par le biais de logiciels adaptés. Dans *EndNote*, chaque BD a été convertie sous un format spécifique avec une extension «*.ENLX*» pour un transfert automatique dans le logiciel *DistillerSR*.

Selon Fabiano et al. (2024), *DistillerSR* est « un outil d'IA validé qui effectue le tri de la littérature après avoir été entraîné sur une proportion de résultats pertinents ». *DistillerSR* intègre une approche mixte, car le filtrage des références est à la fois manuel et automatique.

*DistillerSR* dispose d'un système de traçabilité, permettant de sauvegarder les actions effectuées. En revanche, il n'est pas adapté aux études comportant un nombre réduit de références.

Le processus s'est déroulé en quatre phases.

#### 1.4.1 Phase préliminaire : Exclusion des doublons

Durant cette phase, les références qui étaient identifiées plusieurs fois dans diverses BD ont été supprimées. Le logiciel *DistillerSR* emploie le terme « mises en quarantaine des références identiques », car il y a toujours la possibilité de les reconsidérer. Elles ne sont pas complètement supprimées.

La détection des doublons s'est effectuée par un filtrage à partir des critères suivants : les titres, les résumés et les noms des auteurs, la date de publication, l'édition, le numéro de ressource électronique (*Electronic Resource Number*), *ISBN (International Standard Book Number)*, Numéro (*Issue*), Revue scientifique (*Journal*), Numéro de la revue (*Journal Issue*), Volume de la revue (*Journal Volume*), page, volume et année de publication. L'application de ces critères a permis de détecter 1 006 références avec un taux de confiance compris entre 81 % et 98 %. La fonction « Quarantaine intelligente » a été appliquée pour mettre en quarantaine 160 références, soit un taux d'exclusion automatique de 4.1 %. Finalement, 846 références dupliquées ont été mises en quarantaine manuellement, soit un taux d'exclusion semi-automatique de 22.6 %. 2 890 références ont donc été soumises au filtrage à la phase une.

#### 3.4.2 Première phase : Élimination des références sur la base des types de référence

Cette section vise à exclure les articles selon les critères du type de référence c'est-à-dire les études qui ne sont pas des documents de recherche comme indiqué dans le tableau 5 ci-dessous. Pour ce faire, l'ensemble des types de référence a été identifié dans *EndNote* avant d'être caractérisés en critères critère d'inclusion ou d'exclusion.

**Tableau 5 : Inclusion et exclusion des types de références**

Types de références à inclure:	Types de références à exclure:
<ul style="list-style-type: none"><li>• Article de revue scientifique</li><li>• Livre</li><li>• Actes de conférence</li></ul>	<ul style="list-style-type: none"><li>• Thèse</li><li>• Article de journal</li><li>• Article de magazine</li></ul>

<ul style="list-style-type: none"> <li>• Section de livre / Chapitre de livre</li> <li>• Rapport</li> <li>• Livre électronique</li> <li>• Article électronique</li> <li>• Communication de conférence</li> <li>• Article</li> <li>• Chapitre</li> <li>• Ouvrage de référence</li> </ul>	<ul style="list-style-type: none"> <li>• Générique (référence générique ou non spécifiée)</li> </ul>
---	--

Puis, une exclusion et une inclusion aléatoire des articles ont été réalisées manuellement pour atteindre une proportion d'inclusion minimale de 2 %. Aussi, dans *DistillerSR*, une fonction de filtrage par l'IA permet une sélection automatique des références en les classifiant comme « inclus » ou « exclus ». Cette technique entraîne l'IA à filtrer les références restantes de façon automatique. Finalement, après avoir appliqué la fonction filtrage par l'IA, toutes les références ont été scannées et filtrées pour un taux d'exclusion de 8.6 % comme indiqué dans le tableau 6.

**Tableau 6 : Nombre de références incluses et exclues à la première phase**

Nombre de références total	Total inclus	Total exclu (Taux)
2 890	2 640	250 (8.6 %)

Cela signifie que 250 références étaient soit des thèses, des articles de journaux, des articles de magazines ou des références génériques ou non spécifiées.

Dans la phase suivante, le filtrage des références s'est poursuivi selon les critères de langue et de mots-clés.

### 3.4.3 Deuxième phase : Les critères d'éligibilité contenus dans le titre et le résumé

À cette étape, la fonction de filtrage par l'IA vise à analyser les titres et les résumés, identifier les mots-clés et inclure les articles pertinents. Les publications ne contenant aucun des mots-clés éligibles dans le titre et le résumé ont été systématiquement exclues, ainsi que tous les documents scientifiques n'étant ni en français ni en anglais.

D'abord, les critères d'éligibilité (*SPIDER*) édictés ont été utilisés pour paramétrer le logiciel *DistillerSR*. Chaque fois qu'un des mots-clés concernant les comportements fiscaux frauduleux, les technologies de SIA et les AF étaient identifiés dans le titre ou le résumé d'un article, celui-ci était inclus. Au cas contraire, l'article est exclu. La norme requise par le logiciel *DistillerSR* est d'exclure ou inclure manuellement au moins 2 % des références afin de pouvoir appliquer le filtrage par l'IA.

Le système a été entraîné avec un filtrage manuel de 73 références. Puis, la fonction de filtrage automatisé par l'IA a été utilisée pour analyser les références restantes. Aussi, cette même fonction a été utile pour inclure les 96 références en français et en anglais. Le processus automatique a été validé manuellement pour préserver sa qualité. Par la suite, il a été procédé à l'importation des 2 640 articles restant dans un fichier Excel afin de les analyser et de s'assurer que les 96 articles respectent les critères d'inclusion (Tableau 7). Cela a permis d'écartier tout biais potentiel.

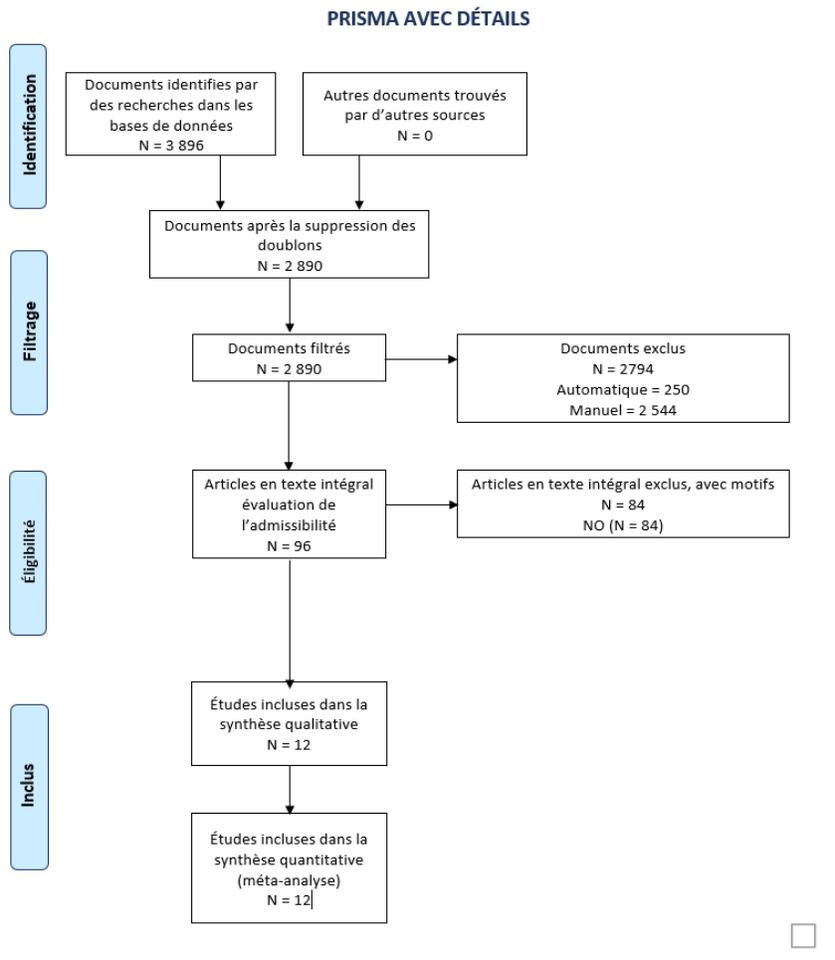
**Tableau 7** : Nombre de références incluses et exclues à la deuxième phase

Nombre de références total	Total inclus	Total exclu (Taux)
2 640	96	2 544 (96 %)

#### 3.4.4 Troisième phase : Critères liés au contenu des articles

Cette phase est liée à l'analyse intégrale du contenu de chaque article afin de décider de son inclusion ou de son exclusion dans la perspective de l'extraction des données. Avec un nombre d'articles à examiner relativement faible, la fonction de filtrage par l'IA de *DistillerSR* ne peut s'appliquer. Une lecture intégrale et une analyse manuelle de l'ensemble des références ont été menées. Chaque article a été analysé en profondeur afin de s'assurer du respect des critères d'éligibilité prédéfinis. Par ailleurs, une évaluation formelle de la qualité des recherches primaires a été réalisée par le directeur de mémoire en vue de produire les meilleures données disponibles pour les chercheurs et les décideurs publics (Baron, 2018). Au terme de cette étape, seuls 12 articles ont été conservés. La qualité de ces 12 articles a été évaluée par le directeur de mémoire comme évaluateur.

Tenant compte des critères d'éligibilité de l'approche *SPIDER*, plusieurs raisons expliquent l'exclusion des 83 articles. Cependant, la raison principale reste l'absence d'un processus méthodologique robuste suivi de résultats et l'application du modèle de SIA impliquant les données d'une AF. Lorsque la technologie utilisée pour la détection de la fraude n'est pas précisée, les auteurs de l'article sont contactés afin d'obtenir plus de précisions. C'est le cas de l'article : *Visual querying and analysis of temporal fiscal networks*. Après une requête envoyée par courriel le 24 décembre 2024, l'auteur a confirmé, le 9 janvier 2025, le fait que la technologie utilisée ne constitue pas un SIA (**voir Annexe C**). Toutes les démarches de sélection des références sont documentées directement dans *DistillerSR* comprenant un « diagramme de flux *PRISMA* 2020 qui suit les références à mesure qu'elles progressent lors des différentes étapes du flux de travail dans cette revue de la portée ». Le diagramme est présenté dans la figure 1 ci-dessous.



Additions to the original PRISMA Flow Diagram, Copyright © 2021, Evidence Partners Inc., All Rights Reserved.  
 Adapted from "Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097"  
 For more information, visit: [www.evidencepartners.com](http://www.evidencepartners.com) , [www.prisma-statement.org](http://www.prisma-statement.org)

**Figure 1 : Diagramme de flux PRISMA 2020**

### 3.5 Analyse de données

Les analyses opérées dans cette recherche portent sur l'échantillon des 12 articles sélectionnés en s'appuyant sur la méthode thématique déductive. L'intérêt de cette approche est d'abord de structurer les renseignements extraits des articles et ensuite de les synthétiser.

### 3.5.1 Analyse thématique déductive des données

Les 12 articles de l'échantillon ont été analysés au moyen d'une technique d'analyse thématique déductive composée de l'analyse thématique et de l'analyse de contenu dirigée (analyse déductive).

L'objectif de l'analyse de contenu dirigée est de valider ou d'explorer davantage un cadre théorique ou des recherches existantes. Cette approche vise à soutenir ou à affiner les théories existantes tout en précisant la question de recherche, ou encore identifier des codes initiaux ou des liens entre les thèmes (Hsieh et Shannon, 2005, p. 1281).

L'analyse thématique vise à « procéder systématiquement au repérage, au regroupement et, subsidiairement, à l'examen discursif des thèmes abordés dans un corpus, qu'il s'agisse d'une transcription d'entretiens ou de divers types de documents (organisationnels, gouvernementaux, littéraires) » (Paillé et Mucchieli, 2021, p.270).

L'étude du cadre théorique a permis d'affiner davantage les principaux concepts en lien avec la question de recherche. L'analyse thématique déductive facilite l'interprétation des thèmes identifiables (Lambert et O'Halloran, 2008) et fournit une analyse plus détaillée des différents aspects des thèmes prédéfinis (Goel et al., 2022) comme l'apprentissage automatique.

L'analyse des données a été réalisée à l'aide de la méthodologie de l'analyse thématique déductive proposée par Crabtree et Miller (1999). Cette méthode se résume en cinq étapes.

Étape 1 : La familiarisation : elle a consisté à pratiquer plusieurs séances de lectures pour obtenir une compréhension spécifique de chaque article, et se familiariser avec leur contenu. Les parties pertinentes des articles ont été surlignées et les notes initiales ont été consignées progressivement dans un fichier Excel. Cette première étape a permis une meilleure compréhension des SIA, y compris les techniques utilisées (les algorithmes), les types de fraudes ainsi que les effets et impacts.

Étape 2 - Le codage déductif : ce processus a consisté en l'attribution de codes aux différents thèmes prédéfinis dans le cadre théorique et la méthode *SPIDER*, à savoir les comportements fiscaux frauduleux liés aux types d'impôts, les SIA utilisés ainsi que les effets et les impacts

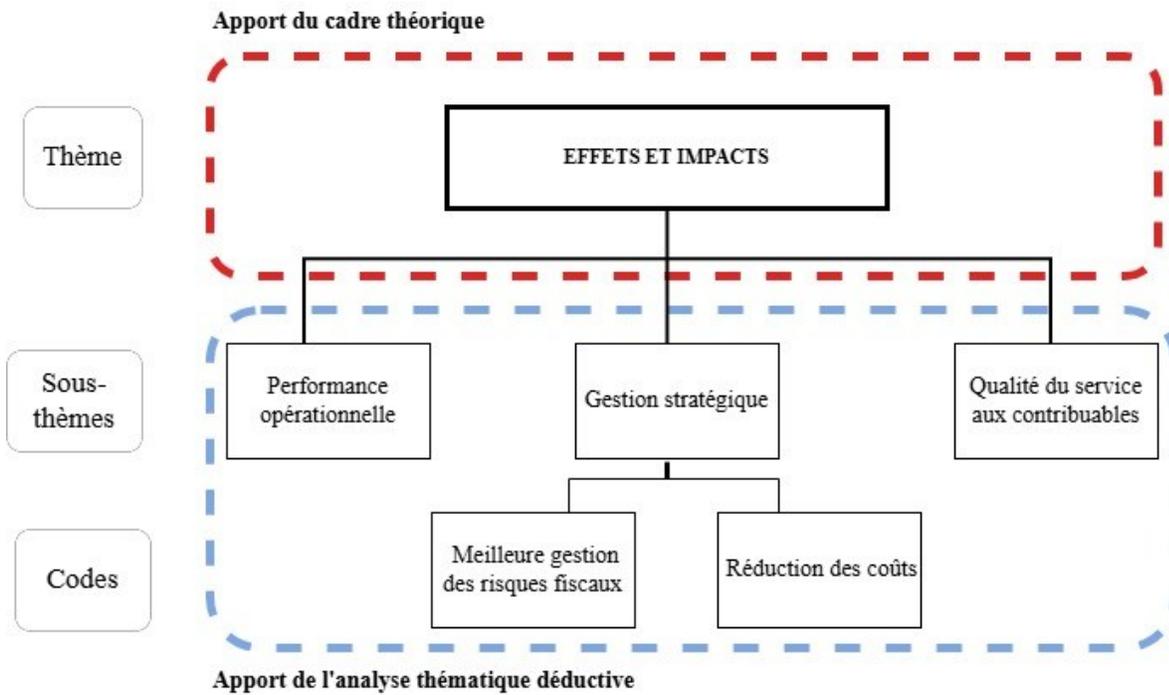
organisationnels. Les logiciels Word et Excel ont été utilisés pour classifier et consigner l'ensemble des codes.

Étape 3 – Le groupement : chaque fois que des codes identiques sont détectés, ils sont regroupés afin de reformuler des sous-thèmes. Une procédure de classification des codes et des sous-thèmes a été réalisée à partir des logiciels Word et Excel. L'utilisation de cette technique a contribué à identifier les sujets pertinents afin de les mettre en corrélation. Par exemple, les données relatives aux enjeux stratégiques ont été regroupées sous le sous-thème « Gestion stratégique ».

Étape 4 – L'immersion: des lectures approfondies ont été effectuées pour mettre en évidence les interprétations répétées ayant permis une émergence d'idées nuancées sur les thèmes préétablis. Par exemple, le sous-thème de la performance organisationnelle a émergé et celui-ci a été abordé dans tous les articles étudiés.

Étape 5 – La génération des thèmes : finalement, tous les thèmes initiaux (SIA, comportements fiscaux frauduleux et effets et impacts) ont été affinés afin de faire émerger de nouveaux sous-thèmes. À titre d'exemple, une fraude de l'impôt sur les acquisitions immobilières a été identifiée dans l'étude menée sur l'AF de la République de Corée. Aussi, trois sous-thèmes relatifs ont été identifiés aux effets et aux impacts, c'est-à-dire la performance opérationnelle, la gestion stratégique et la qualité du service aux contribuables. De plus, deux autres thèmes ont émergé de cette analyse : la méthodologie utilisée pour évaluer la performance des SIA et les caractéristiques des données fiscales. La figure 2 illustre un exemple d'arborescence générée par l'analyse thématique déductive (Lawarée, 2017).

Tous les renseignements consignés dans les tableaux *Word* et *Excel* ont été revus et validés par le directeur de mémoire. Cette révision garantit une cohérence et une validité de l'analyse, facilitant ainsi la présentation claire des résultats.



**Figure 2: Un exemple d'arborescence générée par l'analyse thématique déductive**

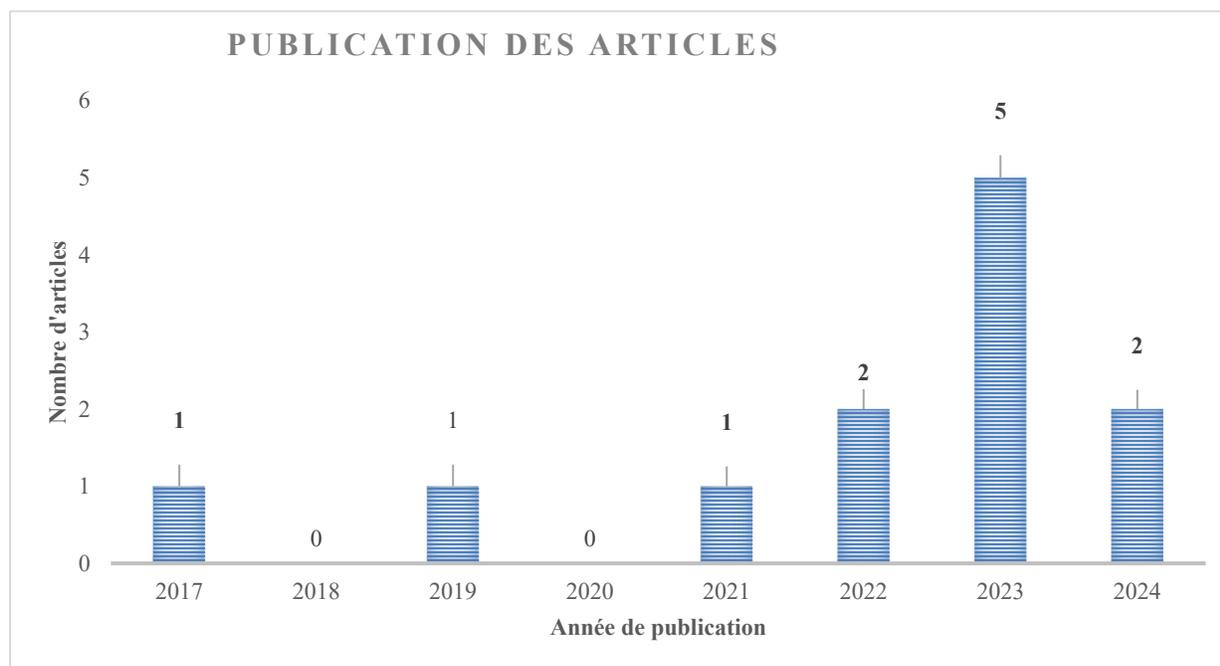
## CHAPITRE 4 : RÉSULTATS

Ce chapitre révèle les résultats de cette RP menée sur la base de l'analyse des 12 études sélectionnées. Cette section présente d'abord les caractéristiques de l'échantillon. Aussi les SIA ayant servi à la détection des comportements fiscaux frauduleux y ont été mis en évidence, ainsi qu'un aperçu des tendances qui ont émergé en lien avec leurs effets et impacts sur les AF.

### 4.1 Caractéristiques générales de l'échantillon

Dans le but d'approfondir l'analyse des études recensées dans cette RP, elles ont été modélisées selon les critères *SPIDER*. Cette démarche permet une lecture structurée et critique des études retenues par rapport à la problématique de l'utilisation des SIA dans la détection des comportements fiscaux frauduleux par les AF.

Parmi les 12 articles sélectionnés, des études ont été menées dans les AF des 9 pays suivants : Espagne, Arabie Saoudite, Yémen, Rwanda, République de Corée, Iran, Mexique, Brésil et Serbie.



**Figure 3 : Répartition des articles par pays et par année de publication**

Comme indiqué sur la figure 3 ci-dessus, aucune recherche en lien avec les critères établis n'a été publiée en 2018 et 2020. Cependant, les trois premiers articles pertinents pour cette étude ont été publiés respectivement en 2017 en Iran, en 2019, en Espagne et au Mexique en 2021. En 2022,

deux articles ont été publiés en République de Corée et au Rwanda tandis qu'en 2024, deux autres articles ont été publiés au Brésil et au Yémen. 2023 est l'année qui a vu le plus grand nombre d'articles publiés sur le sujet dans quatre pays : Espagne, Arabie Saoudite, Serbie et Rwanda. Les détails sont indiqués dans le tableau de l'annexe D.

#### **4.2 Analyse des données et processus d'évaluation des SIA**

D'abord, la plupart des études ont été réalisées à partir de données réelles historiques provenant des AF (voir annexe D). Tous les échantillons de données utilisés dans les 12 articles sélectionnés sont des données collectées sur des années consécutives. Cette méthode fournit une analyse tendancielle et comportementale des contribuables sur un intervalle de temps continu. Il est observé plusieurs cycles de collecte de données historiques en fonction des besoins des études allant d'une année à six ans. En revanche, deux études ne fournissent pas la période correspondant aux données utilisées : l'étude menée avec les données de l'AF du Rwanda (Belle Fille et al., 2022) et celle menée par l'AF du Yémen (Shujaaddeen et al., 2024). Selon les différentes études, les données sont des déclarations fiscales, des données financières, certains attributs fiscaux comme les numéros des taxes, des données socio-économiques pour les individus (le genre, l'état civil, etc.). Les SIA nécessitant une grande quantité de données, celles-ci sont généralement collectées par milliers voire par millions pour mener les études. Par exemple, dans l'étude sur l'AF du Mexique, 6 823 415 757 de données de facturation électronique mensuelles entre contribuables, collectées entre janvier 2015 et décembre 2018, ont été anonymisées et agrégées pour les besoins de la recherche. (Zumaya et al., 2021)

Ensuite, après une analyse de l'ensemble des études, il a été observé que pour chaque étude, un processus de mesure de la performance des modèles de SIA a été mis en œuvre afin de s'assurer de la rigueur de la démarche scientifique. Belavagi et Muniyal (2016) décrivent les six principales étapes de cette mesure de la performance : a) prétraiter l'ensemble de données, b) diviser l'ensemble des données en données d'entraînement et données d'essai, c) construire le modèle de classification sur les données d'entraînement pour les SIA, d) lire les données d'essai, e) tester les modèles de classification de SIA sur les données d'entraînement et f) mesurer la performance par des outils spécifiques comme la précision, l'exactitude, le rappel, etc.

Les deux thèmes émergents, à savoir, la méthodologie utilisée pour évaluer la performance des SIA et les caractéristiques des données fiscales, sont pertinents pour cette revue de portée. En effet, l'évaluation de la performance des technologies de SIA contribue à une validation rigoureuse de la recherche, une comparaison scientifique des algorithmes déployés et enfin fournit une mesure raisonnable des impacts sur les AF. Par exemple, au Rwanda, l'étude menée par Belle Fille et al. (2023) a porté sur quatre techniques de SIA : Arbres de décision, Forêt aléatoire, *Gradient boost* et *XGBoost*. Finalement, l'évaluation de la performance a montré que le modèle de Forêt aléatoire s'est révélé être plus performant pour détecter la fraude à la TVA.

### **4.3 Résultats liés aux comportements fiscaux frauduleux, aux interventions des SIA, aux effets et impacts**

L'analyse des études laisse apparaître une corrélation cohérente entre les comportements fiscaux frauduleux liés aux types d'impôts (évasion ou fraude à la TVA ou sous-déclaration) et les SIA utilisés pour l'intervention. Les caractéristiques et les spécificités des données fiscales et des comportements frauduleux conditionnent ainsi le choix du modèle de SIA approprié pour la détection de la fraude ou de l'évasion comme indiqué sur la figure 3.

Cette section présente donc les résultats issus des interventions des SIA dans la détection de la fraude ainsi que les effets et les impacts constatés.

#### **4.3.1 Intervention des systèmes d'intelligence artificielle sur les comportements fiscaux frauduleux identifiés**

L'échantillon étudié a mis en évidence l'utilisation de 5 méthodes d'apprentissage automatique que sont l'apprentissage supervisé, l'apprentissage non supervisé, l'apprentissage hybride (supervisé et non supervisé), indépendamment du type d'impôt fraudé ou évadé. Sur les 12 études, 58 % (7 études) ont employé l'apprentissage supervisé pour la détection d'impôts directs et indirects frauduleux, 8 % (1 étude) ont utilisé l'apprentissage non supervisé pour détecter la fraude sur les acquisitions immobilières et enfin 33 % des études (4 études) ont déployé un système hybride pour la détection d'impôts directs et indirects frauduleux.

#### 4.3.1.1 Apprentissage supervisé

Dans l'échantillon d'études de cette RP, l'apprentissage supervisé a été déployé au sein des AF en Arabie Saoudite (Mohamed Maher Ben et AlSadhan, 2023), au Rwanda (Belle Fille et al., 2022; Murorunkwere et al., 2023; Belle Fille et al., 2023), au Brésil (Pires, 2024), au Mexique (Zumaya et al., 2021) et en Espagne (Pérez López et al., 2019) pour la détection de la fraude et de l'évasion fiscale liées aux impôts directs et indirects.

L'apprentissage supervisé utilise des données issues de contribuables qui ont déjà fait l'objet de vérification par les AF. Ces données ont déjà été étiquetées c'est-à-dire que ce sont des contribuables classés frauduleux ou non frauduleux. L'apprentissage supervisé s'applique également dans un contexte de comportements fiscaux frauduleux, répétitifs et connus des AF. Plusieurs caractéristiques fiscales et économiques sont associées à ces données comme le type de contribuable (entreprises ou particuliers), le secteur d'activité (service, commerce ou industrie) (Rahimikia et al., 2017), les facteurs démographiques (Murorunkwere et al., 2023), la source de revenus, les dépenses déductibles, les anomalies dans les déclarations, etc. Les différents algorithmes utilisés sont entraînés pour reconnaître les caractéristiques associées aux fraudes et données étiquetées afin de fournir une probabilité de détection de la fraude sur de nouvelles données.

Concernant les impôts directs, différents algorithmes, notamment le réseau neuronal profond (Mohamed Maher Ben et AlSadhan, 2023), le Réseau de neurones artificiels (Murorunkwere et al., 2022), le *Gradient Boosting* (Pires, 2024), le Réseau neuronal perceptron multicouche (Pérez et al., 2019) ont été mis en œuvre pour détecter les fraudes sur les revenus des particuliers et des entreprises.

En revanche, les Réseaux neuronaux profonds, la Forêt aléatoire au Mexique (Zumaya et al., 2021), les Réseaux de neurones artificiels, les Arbres de décision, le *Gradient boost* et *XGBoost* (Belle Fille et al., 2023), les Réseaux de neurones artificiels (Murorunkwere et al., 2022) ont été utiles dans la détection des impôts indirects, notamment la TVA.

Plus précisément, une étude a été menée au sein de l'AF rwandaise en vue de la prédiction des comportements futurs de non-conformité (Murorunkwere et al., 2023) sur des informations émises

par les machines de facturation électronique. Dans ce contexte, 852 239 enregistrements anonymisés et étiquetés, provenant de contribuables déjà vérifiés, ont été utilisés. Les éléments caractéristiques se composent entre autres de la période d'activation de la machine électronique, l'historique des opérations, l'emplacement géographique des entreprises, les informations liées aux reçus émises par les machines, etc. Les trois algorithmes d'apprentissage supervisé adoptés sont les arbres de décision, la forêt aléatoire et le *Gradient boost* et *XGBoost*. Ce dernier modèle a produit le meilleur résultat avec une précision de 92.3 %.

Ainsi, il faut indiquer que l'étude menée au Rwanda par Murorunkwere et al. (2023) a porté simultanément sur la détection de la fraude de l'impôt sur le revenu des personnes physiques et des sociétés et la fraude à la TVA avec l'utilisation du même modèle d'apprentissage supervisé.

#### **4.3.1.2 Apprentissage non supervisé**

Selon Bolton et Hand (2002), les méthodes d'apprentissage non supervisées sont basées sur des données fiscales ne contenant pas d'observations légitimes et frauduleuses. Dans l'échantillon étudié, seule l'étude menée au sein de l'AF de la République de Corée (Lee, 2022) a révélé l'utilisation de l'apprentissage non supervisé afin de détecter la fraude liée à l'impôt sur les acquisitions immobilières. Contrairement au modèle supervisé, l'objectif de l'apprentissage non supervisé est la détection de comportements jugés anormaux, sans supervision et qui auraient été difficilement détectés par un auditeur fiscal. L'apprentissage non supervisé ne se fonde pas sur des données étiquetées c'est-à-dire qu'il fonctionne dans un environnement où il n'existe pas de données de contribuables ayant fait l'objet de vérifications préalables (frauduleux et non frauduleux). Dans cette étude, le jeu de données est composé des déclarations fiscales immobilières de 2 228 contribuables liées aux acquisitions de terrains sur une période de trois années consécutives (2016–2018). Les éléments caractéristiques des données sont les prix d'achat ou de vente, la surface et l'emplacement du terrain : immeuble à but commercial, d'habitation ou autre, etc. Sur l'ensemble des contribuables, trois cas de fraudes ont été détectés.

#### **4.3.1.3 Apprentissage hybride ou mixte**

De façon générale, les systèmes hybrides sont déployés dans la détection de cas complexes pouvant impliquer des réseaux de fraudes relativement bien organisés en combinant plusieurs secteurs

d'activité avec l'emploi de données non structurées. Ces systèmes peuvent aussi s'appliquer à des fraudes émergentes requérant une surveillance continue avec une capacité d'adaptation à un environnement changeant lié à des types de fraudes de plus en plus sophistiqués grâce au numérique.

Dans l'échantillon d'études, il a été développé des modèles de systèmes hybrides pour détecter les comportements fiscaux frauduleux en Arabie Saoudite (Alsadhan, 2023), en Iran (Rahimikia et al., 2017), en Serbie (Savić et al., (2022) et au Yémen (Shujaaddeen et al., 2024) pour combattre la fraude fiscale liée à la TVA. Trois types de systèmes hybrides ont été utilisés. Le premier système combine essentiellement des modèles d'apprentissage supervisés, le second est composé uniquement de modèles d'apprentissage non supervisés et le dernier est un mixte de l'apprentissage supervisé et l'apprentissage non supervisé.

Plus spécifiquement, différents systèmes hybrides supervisés ou non supervisés ont été mis en œuvre pour détecter la fraude de la TVA et la fraude sur les revenus des entreprises et des particuliers. D'abord, en Arabie Saoudite, une combinaison d'apprentissage supervisé (*EXtreme Gradient Boosting (XGBoost)*), réseau de neurones artificiels, *support vector machine*) et d'apprentissage non supervisé (*Autoencoders*) a permis de détecter la fraude à la TVA (Alsadhan, 2023). Par contre, c'est un réseau neuronal hybride supervisé (*Multilayer Perceptron*) et non supervisé (*Self-Organizing Map*) qui a été déployé au Yémen pour la détection de la fraude à l'impôt sur les bénéfices commerciaux et industriels. (Shujaaddeen et al., 2024) Ensuite, un apprentissage hybride supervisé (*Support Vector Machine, Multilayer Perceptron* et Régression logistique avec optimisation de la recherche d'harmonie) a été implémenté en Iran pour combattre l'évasion fiscale de l'impôt sur le revenu des entreprises. (Rahimikia et al., 2017) Aussi, l'étude menée en Serbie a permis l'implémentation d'un apprentissage hybride (le *Clustering* et l'*Autoencoder*) en vue de la détection de l'évasion fiscale de l'impôt sur le revenu des particuliers. (Savić et al., 2022)

Tous ces modèles ont démontré des niveaux de performances élevés grâce à leur évaluation par les outils de mesure de la performance. Cela prouve leur robustesse à lutter contre les comportements fiscaux frauduleux. Plus concrètement, l'AF du Yémen a fourni les déclarations de revenus de 1 083 entreprises commerciales et industrielles pour l'étude. Les variables principales utilisées sont le chiffre d'affaires, les montants des taxes à payer, les amendes et le ratio des taux d'imposition par rapport au chiffre d'affaires. D'une part, le *Multilayer Perceptron*

qui est un apprentissage non supervisé, identifie les variables complexes au sein des données non étiquetées, procède ensuite à un regroupement des contribuables ayant des comportements identiques. D'autre part, le *Self-Organizing Map*, apprentissage non supervisé procède à une analyse plus poussée afin de détecter les niveaux d'évasion fiscale ou de conformité fiscale. Les modèles d'apprentissage automatique capables de détecter les comportements fiscaux frauduleux et leurs caractéristiques sont décrits dans la figure 4 ci-dessous.

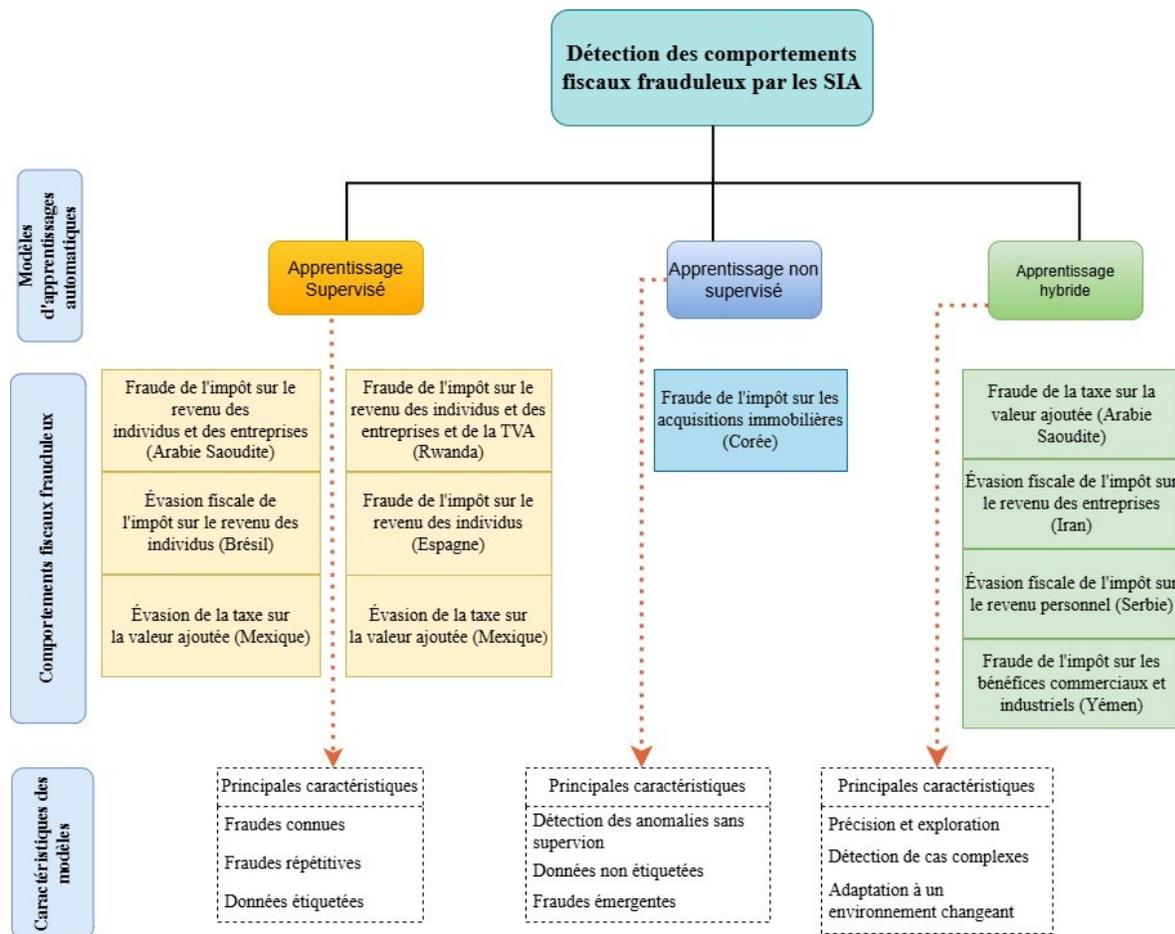


Figure 4: Modèles d'apprentissage automatique, comportements fiscaux frauduleux et leurs caractéristiques

L'examen des SIA pour la détection des comportements fiscaux frauduleux a produit des effets et impacts documentés dans l'annexe E et analysés dans les sections suivantes.

### **4.3.2 Effets et impacts organisationnels**

Un examen transversal de l'échantillon met en lumière les effets et les impacts des interventions de l'apprentissage automatique menées pour la détection des comportements fiscaux frauduleux. Les effets et impacts identifiés dans les articles de cette RP sont indiqués dans le tableau de l'annexe E. Trois axes principaux se dégagent de ces études à savoir le renforcement de la performance des opérations fiscales, des services rendus aux contribuables et l'optimisation de la gestion stratégique et du ciblage fiscal.

#### **4.3.2.1 Renforcement de la performance des opérations fiscales**

Une des principales contributions de l'utilisation de l'apprentissage automatique est l'automatisation et la maximisation des processus de vérification. En plus d'accroître l'efficacité des vérifications par la réduction des efforts manuels des inspecteurs, les études menées au Rwanda soulignent également une augmentation de l'efficacité des vérificateurs (Belle Fille, 2022). Dans cette étude, les réseaux de neurones artificiels ont permis la détection de cas de fraudes liées à l'impôt sur le revenu des entreprises et des particuliers avec une précision de 92 %. Cela a eu pour impact direct, la réduction du temps de vérification et une augmentation de la productivité des vérificateurs.

#### **4.3.2.2 Optimisation de la gestion stratégique et du ciblage**

Même si l'automatisation est considérée comme un pan important de l'intervention des modèles d'apprentissage automatique, il n'en demeure pas moins que ces modèles offrent également des capacités avancées d'analyse stratégique. En effet, les modèles d'apprentissage automatique soutiennent une meilleure planification et une gestion des risques fiscaux. Par exemple, en Arabie Saoudite, le réseau neuronal profond a permis de détecter la fraude de l'impôt sur le revenu des individus et des entreprises avec une exactitude de 99 %, grâce à sa capacité à mieux hiérarchiser les risques. (Mohamed Maher Ben et AlSadhan, 2023)

Cependant, seulement 17 % des études (deux articles) ont fourni une estimation des montants fraudés ou évadés. Les impacts financiers sont estimés à 872 236 297, 69 R\$ pour l'AF du Brésil (Pires, 2024) et un montant estimé entre 135 millions et 1 milliard de pesos pour l'AF mexicaine. (Zumaya et al., 2021)

Les AF peuvent ainsi améliorer la planification des risques et l'allocation des ressources associées à leur gestion.

#### **4.3.2.3 Bonification de la qualité du service aux contribuables**

Un des effets indirects de l'application des SIA pour la détection des comportements fiscaux frauduleux est l'amélioration du service. Les études de l'échantillon mettent en évidence la diminution du harcèlement fiscal injustifié des AF vis-à-vis des contribuables. À titre d'exemple, l'étude menée au Rwanda par Belle Fille et al. (2023) avec l'application de l'apprentissage supervisé pour détecter la fraude à la TVA a engendré une amélioration de la transparence dans l'utilisation des machines de facturation. L'impact sur les contribuables a été la réduction des vérifications inutiles.

Les effets et les impacts soulevés dans l'échantillon d'études n'ont pas fait l'objet de validation. Dans l'ensemble, il manque souvent de preuves pour étayer les affirmations. Ce qui constituera aussi l'objet d'une discussion au chapitre suivant. L'examen des SIA pour la détection des comportements fiscaux frauduleux a produit des effets et impacts documentés dans l'annexe E et analysés dans les sections suivantes.

## **CHAPITRE 5 : DISCUSSIONS**

---

Les résultats obtenus au chapitre précédent révèlent des insuffisances dans l'application des SIA au sein des AF. L'objet de ce chapitre est d'approfondir l'étude à travers un examen critique des tendances observées. Ensuite, il s'agira de discuter les effets observés, les limites et puis proposer des pistes de recherches futures.

### **5.1 Synthèse des principaux résultats et apports de la revue de portée**

Les études ont porté essentiellement sur les principaux impôts directs et indirects. L'évasion fiscale de l'impôt sur le revenu des entreprises et des particuliers, la fraude de l'impôt sur le revenu des sociétés et des particuliers, la fraude de l'impôt sur les bénéfices commerciaux et industriels et finalement la fraude de l'impôt sur les acquisitions immobilières constituent les impôts directs ayant fait l'objet d'études. Quant aux impôts indirects, les comportements fiscaux frauduleux identifiés sont l'évasion et la fraude à la TVA.

Les SIA mis en action pour la détection des comportements fiscaux frauduleux sont diversifiés et ont mobilisé les modèles d'apprentissage automatique que sont l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage hybride.

Cette RP a permis de mettre en lumière l'intérêt grandissant des AF pour l'usage des SIA non seulement comme outil opérationnel de détection de la fraude, mais également comme un outil stratégique de détection proactive et de planification à moyen et long terme. En effet, les différentes études ont montré que la précision des modèles d'apprentissage automatique est d'au moins 90%. Les AF se servent des SIA pour opérer des transformations stratégiques dans le processus d'évaluation des risques et une optimisation des stratégies d'audits fiscaux avec un objectif ultime d'augmenter les recettes budgétaires. Au Rwanda, l'étude menée par Belle Fille et al. (2023) a montré une nette réduction du temps d'évaluation des risques et de facto une meilleure gestion des processus de vérification. En République de Corée, la méthode d'apprentissage supervisée utilisée par Lee (2022) a perfectionné les stratégies de sélection des contribuables à risque.

## 5.2 Examen critique des méthodologies de l'échantillon d'étude

L'analyse des méthodologies déployées pour évaluer la performance des études a permis d'identifier certains enjeux qui méritent d'être discutés.

D'abord, la plupart des AF ont fourni des données réelles en prenant le soin de les rendre anonymes. Un tel procédé accorde une certaine crédibilité aux études. Par exemple, le Service d'administration fiscale au Mexique a fourni des données de facturation électronique anonymisées de 6 823 415 757 émissions mensuelles agrégées entre contribuables (Zumaya et al., 2021). Ces données ont été collectées sur une période de quatre ans, entre janvier 2015 et décembre 2018. Si cette étude n'a pas évoqué les enjeux éthiques en matière de confidentialité et de protection des informations personnelles, il en est de même pour la plupart des études menées.

Un second défi est lié à l'interprétabilité et la transparence des modèles. En raison du nombre important de paramètres qu'ils contiennent et de la difficulté de compréhension du mécanisme interne, les réseaux neuronaux profonds sont parfois assimilés à des boîtes noires Zhang et al. (2021). En utilisant ces algorithmes, les AF peuvent être confrontées à des difficultés lors de litiges avec les contribuables. Elles peuvent avoir de la difficulté à expliquer aux contribuables certaines décisions algorithmiques sources d'enjeux légaux lors des vérifications. Par exemple, dans l'étude menée par Mohamed Maher Ben et AlSadhan (2023), il a été mis en œuvre le réseau neuronal profond pour la détection de la fraude de l'impôt sur le revenu des individus et des entreprises. Cependant, l'article n'indique pas quelles techniques spécifiques ont été utilisées pour expliquer le processus interne du modèle d'où la limite de l'interprétabilité.

En définitive, il a été constaté que plusieurs études n'ont pas mis simultanément en œuvre différentes techniques d'algorithmes. Le recours à plusieurs techniques lors de la même étude permet une comparaison rigoureuse des techniques avec pour avantage le choix de la plus performante. Comme illustration, Belle Fille et al. (2023) ont testé et comparé plusieurs techniques d'apprentissage supervisé comme les Arbres de décision, la Forêt aléatoire, le *Gradient boost* et le *XGBoost* au Rwanda. En se servant des mêmes données fiscales, la technique de Forêt aléatoire a démontré une performance plus élevée pour détecter la fraude à la TVA.

### 5.3 Analyse critique de la relation entre les SIA utilisés et les comportements fiscaux frauduleux

L'ensemble de l'échantillon d'étude indique l'importance de la disponibilité et de la qualité des données afin d'identifier le modèle optimal d'IA. En effet, les choix des modèles d'apprentissage automatique, supervisés, non supervisés ou hybrides déployés sont fonction de la complexité des comportements fiscaux frauduleux, de la quantité des données disponibles, et enfin du niveau de maturité technologique des AF.

D'abord, l'échantillon semble indiquer que les systèmes d'apprentissage supervisé sont plus efficaces avec un taux d'utilisation de 58% par les AF, comparativement aux modèles non supervisés et mixtes. Ils ont contribué aussi bien à la détection des impôts directs qu'indirects. L'étiquetage et la structuration des données, ainsi que les données de vérifications historiques jouent un rôle central dans le fonctionnement du modèle. Ces modèles peuvent être utilisés par des AF avec un niveau de maturité numérique avancée. Cela signifie que les systèmes informatiques existants sont conçus d'une manière à fournir des données prêtes à être utilisées. Les algorithmes tels que le Réseau neuronal profond (Arabie Saoudite), le Réseau de neurones artificiels (Rwanda), le *Gradient Boosting* (Brésil), le Réseau neuronal perceptron multicouche (Espagne), les Réseaux neuronaux profonds, la Forêt aléatoire (Mexique) et les Arbres de décision, la forêt aléatoire, le *Gradient boost et XGBoost* (Rwanda) ont été efficaces pour détecter les fraudes connues. Cette efficacité est due aux données structurées et parfois très volumineuses comme cela a été le cas pour l'étude menée en Espagne avec plus de 2 millions de déclarations fiscales.

Par ailleurs, une comparaison des algorithmes de modèles supervisés de l'échantillon montre que la catégorie des Réseaux de neurones artificiels (Réseau neuronal perceptron multicouche (MLP) et Réseaux neuronaux profonds) est plus utilisée et efficace. La recherche de Delgado et al. (2023) sur la détection de l'évitement fiscal en Europe a mis en évidence la performance des Réseaux de neurones artificiels. En effet, l'étude montre que cette technique s'est révélée efficace en analysant de façon précise de grands ensembles de données non linéaires pour détecter des phénomènes complexes comme l'évitement fiscal. Plusieurs études confirment cette assertion. Au Mexique, deux algorithmes de modèles supervisés, la Forêt aléatoire et les Réseaux neuronaux profonds ont été mis à contribution pour la détection de l'évasion de la TVA. Les Réseaux neuronaux profonds

ont été plus performants avec une capacité à détecter 149 921 contribuables suspects contre 128 227 pour la Forêt aléatoire.

Afin de mesurer efficacement la performance des modèles supervisés, il est primordial d'avoir à disposition des données labélisées. Au cas contraire, le modèle peut être biaisé lorsque de nouvelles catégories de fraudes surviennent. Pour y remédier, Belle Fille et al. (2022) ont fait l'usage de techniques supplémentaire d'apprentissage automatique pour modéliser les données. Ils ont utilisé du *Random Under sampling (RUS)* et du *Synthetic minority over-sampling technique (SMOTE)*. L'usage de ressources supplémentaires pourrait s'expliquer à l'incapacité de l'AF rwandaise de générer des données fiscales améliorées et prêtes pour l'usage.

Ensuite, Vanhoeyveld et al. (2020), ont mis en œuvre une technique d'anomalies non supervisées dans la détection de la fraude à la TVA. Cette étude s'est basée sur des données non labélisées venant de l'AF belge. L'apport de cette technique non supervisée est qu'elle détecte efficacement des contribuables suspects, de divers secteurs d'activités, qui n'avaient pas encore été appréhendés pour cas de fraude, en plus d'analyser des données fiscales et économiques non étiquetées. L'échantillon d'étude indique une seule application de l'apprentissage non supervisé, l'*Autoencoder* (République de Corée) ayant contribué à détecter la fraude de l'impôt sur les acquisitions immobilières. Cependant, la performance de cette technique est à relativiser, car le taux de détection a augmenté seulement de 0.1% (3 fraudes sur 2 228) à 0.9% (3 fraudes sur 352). (Lee, 2022) L'analyse de son efficacité dans le contexte de cette étude risque d'être superficielle, car l'étude manque de comparatif dans l'échantillon. Aussi, selon Alsadhan (2023), l'usage des modèles non supervisés présente des limites lorsqu'ils sont utilisés de façon indépendante sans l'apport simultané d'un autre modèle. Cette étude fait référence aux modèles d'apprentissage hybrides qui constituent une alternative.

En effet, l'usage de l'apprentissage hybride représente 33% de l'échantillon étudié. L'analyse montre que deux études ont utilisé une combinaison de l'apprentissage supervisé et non supervisé (Arabie Saoudite et Yémen). Une étude a employé une combinaison d'algorithmes supervisés (Iran) et une autre un amalgame d'algorithmes non supervisés (Serbie). Chaque système hybride, combinant deux modèles supervisés, a allié au moins des Réseaux de neurones artificiels ou une de ses sous-catégories. Cela confirme la thèse développée par Delgado et al. (2023) sur la

performance et la capacité avérée des Réseaux de neurones artificiels à identifier la fraude et l'évasion fiscales. En Iran, le système hybride a combiné le *Support Vector Machine*, le *Multilayer Perceptron* et la Régression logistique avec optimisation de la recherche d'harmonie (HS) pour détecter l'évasion fiscale de l'impôt sur le revenu des entreprises. La contribution du *Multilayer Perceptron* a été plus significative par rapport aux autres algorithmes dans les deux secteurs du textile et de l'alimentation. Les taux d'exactitude (90.07%), de sensibilité (85.48%) et de spécificité (90.34%) étaient plus élevés que les autres techniques. En Arabie Saoudite, Alsadhan (2023) a utilisé un système hybride composé de quatre modèles : un modèle supervisé (EXtreme Gradient Boosting (XGBoost)), un modèle non supervisé (*autoencoders*), un modèle de prédiction (les Réseaux de neurones artificiels et le *Support Vector Machine*). Selon ses résultats, même si les Réseaux de neurones artificiels ont été efficaces pour détecter les cas de fraude avec une exactitude de 74.78%, le modèle non supervisé a été utile dans le calcul des anomalies pour chaque variable originale. Cette approche démontre la complémentarité des deux systèmes dans un environnement mixte.

Cependant, l'étude de Shujaaddeen et al. (2024) fait remarquer que cette performance n'est pas absolue et dépend des secteurs d'activité. Aussi, ils ont construit un nouveau modèle qui allie l'apprentissage supervisé et non supervisé. L'objectif est de pallier les biais provoqués par le manque de données non labélisées. Pendant que l'algorithme non supervisé, le *Self-Organizing Map* se focalise sur les données non labélisées, l'apprentissage supervisé *Multilayer Perceptron (MLP)* souligne les données étiquetées.

Par ailleurs, Vanhoeyveld et al. (2020) indiquent dans leur étude que l'efficacité des SIA dans la détection de la fraude à la TVA est en fonction du secteur d'activités. Ils précisent que le choix de la méthode de SIA de la détection de la fraude fiscale devrait être déterminé par les caractéristiques du secteur d'activités. En effet, les caractéristiques financières peuvent différer selon le domaine d'activités économiques des entreprises. Dans l'échantillon, seule l'étude menée au sein de l'AF de l'Iran a considéré ce principe. En effet, Rahimikia et al. (2017) ont mené leur étude en considérant les données provenant de deux secteurs, le textile et l'alimentation. L'application du système hybride a produit des résultats différents. Leur étude confirme la thèse de Vanhoeyveld et al. (2020) selon laquelle le modèle de SIA doit s'adapter à chaque secteur d'activités.

L'application des systèmes d'apprentissage automatique par les AF a généré des effets et des impacts qu'il convient d'analyser.

#### **5.4 Analyse des effets et des impacts organisationnels**

L'analyse de l'échantillon d'étude indique que les effets et les impacts de l'utilisation des SIA par les AF se situent à différents niveaux. De façon générale, les mêmes effets et impacts sont observés indépendamment des comportements fiscaux frauduleux détectés et des interventions des SIA comme indiqué à l'annexe F. Trois principaux effets et impacts ont été identifiés à savoir l'impact sur la prise de décisions stratégiques, les services opérationnels et le service aux entreprises et aux particuliers.

##### **5.4.1 Effets et impacts stratégiques**

Dans un premier temps, les SIA contribuent significativement au processus de gestion et d'évaluation des risques fiscaux. La gestion efficace des risques fiscaux est dépendante de la compréhension des différents stratagèmes fiscaux. Les modèles de SIA prédictifs contribuent à une meilleure segmentation des contribuables en fonction des profils de risque.

Pamisetty (2029) démontre dans son étude, comment l'usage de l'apprentissage automatique est efficace pour une détection de la fraude fiscale et une évaluation des risques fiscaux en temps réel dans les AF. L'étude propose une solution permettant aux gouvernements de détecter la fraude au moment de sa réalisation. Plusieurs études de l'échantillon ont indiqué également comment l'usage des SIA contribue à la gestion des risques fiscaux.

Par exemple au Yémen, l'utilisation d'un système hybride a contribué à l'amélioration de la nomenclature des comportements des contribuables (Shujaaddeen et al., 2024). Au Rwanda et en Iran, la gestion des risques fiscaux permet une meilleure hiérarchisation des profils de risques (Murorunkwere et al., 2023 ; Rahimikia et al., 2017). Cependant, ces deux études ne fournissent pas de chiffres détaillés sur la nomenclature ou la hiérarchisation des profils de risques. Une gestion appropriée des risques a plusieurs avantages comme l'amélioration et le perfectionnement des politiques de vérification. À titre d'exemple au Brésil et en Arabie Saoudite (Pires, 2024 ; Mohamed Maher Ben et AlSadhan, 2023)) l'apprentissage supervisé a démontré une excellente capacité de prévision des risques sur l'ensemble des tests. Ce qui indique une forte performance

de la généralisation de la gestion des risques. Une limite de ces études est qu'elles ne mentionnent pas si les AF objet de l'étude ont adopté ces techniques algorithmiques de façon continue et en temps réel, car les recherches ont porté sur des données historiques.

Un autre avantage lié aux enjeux stratégiques est une meilleure planification des activités de vérification et de recouvrement. En Espagne, en Arabie Saoudite et en République Corée, les études indiquent que les modèles supervisés et non supervisés ont augmenté le niveau des prises de décisions relatives aux vérifications et aux recouvrements (García et Caballero, 2023 ; Alsadhan, 2023 ; Lee, 2022). Plus concrètement, avec l'utilisation de l'apprentissage supervisé au Mexique, entre 7 677 et 17 769 contribuables suspects d'évadés fiscaux ont été détectés pour une estimation s'élevant entre 135 millions et 1 milliard de pesos de 2015 à 2018, avec une moyenne de 259 millions de pesos par an (Zumaya et al., 2021). Par ailleurs, l'apprentissage supervisé déployé au Brésil a permis de détecter une évasion potentielle de 872 236 297, 69 R\$ (Pires, 2024). La publication du nombre de contribuables frauduleux et une approximation des potentiels revenus perdus et qui pourraient être récupérés constituent une bonne mesure de planification des audits et du recouvrement. Cependant, les articles ne fournissent pas une approximation des montants ayant été réellement recouvrés par les AF. Cette dernière information fournirait un impact plus concret de l'efficacité des SIA utilisés.

En plus de renforcer leur capacité, les modèles d'apprentissage automatique aident les AF à développer des stratégies de conformité fiscale en fonction des risques et des types de contribuables identifiés, impactant ainsi les opérations fiscales.

#### **5.4.2 Effets et impacts organisationnels**

En ce qui concerne l'amélioration de la performance des activités opérationnelles, une des principales contributions de l'utilisation de l'apprentissage automatique est la maximisation des processus de vérification. La détection de la fraude par les SIA en amont aide les AF à concentrer leurs efforts sur les contribuables présentant un risque plus élevé avec un jugement humain en plus. Pour Baghdasaryan et al. (2022), les informations collectées après l'intervention des SIA peuvent être utilisées par la division de gestion des risques pour améliorer les règles actuellement mises en œuvre dans le système de gestion des risques et pour se concentrer sur la vérification des contribuables les mieux classés, permettant ainsi d'économiser des ressources. En effet, cela

entraîne une réduction du niveau des vérifications manuelles effectuées par les agents vérificateurs du fait de la connaissance par anticipation des cas frauduleux. Cela suppose aussi un gain de temps et de ressources pour les AF et une amélioration de la productivité des vérificateurs.

Par exemple, l'étude menée en République de Corée par Lee, 2022, a porté sur 2 228 contribuables. L'utilisation de l'apprentissage non supervisé l'*Autoanccoder* a permis de détecter 352 échantillons suspects en lieu et place de 2 228 contribuables à vérifier. Par la suite, les vérificateurs ont identifié trois cas frauduleux. Sans l'utilisation d'un SIA, la vérification serait manuelle et aurait porté sur 2 228 contribuables.

En outre, les études menées au Rwanda (Belle Fille et al., 2022) soulignent également une augmentation de l'efficacité des vérificateurs. Ces études révèlent que les Réseaux de neurones artificiels ont occasionné la détection de cas de fraudes liées à l'impôt sur le revenu des entreprises et des particuliers avec une précision de 92%. Cela a eu pour impact direct, la réduction du temps de vérification et une augmentation de la productivité des vérificateurs.

Aussi, la réduction du temps de traitement et d'analyse des opérations par les employés implique une diminution des erreurs causées par le travail manuel ainsi qu'une diminution des budgets investis pour les vérifications. Au Yémen, le système hybride jugé plus performant que les autres modèles (Shujaaddeen et al., 2024)) a permis une maximisation des coûts de détection et des ressources internes en plus d'une amélioration de la précision lors de la détection de l'impôt sur les bénéfices commerciaux et industriels.

En Arabie Saoudite, l'étude de Mohamed Maher Ben et AlSadhan (2023) indique que l'utilisation de l'apprentissage supervisé a démontré une augmentation de l'efficacité opérationnelle par une amélioration de la précision lors de la détection de l'impôt sur les bénéfices commerciaux et industriels avec une baisse des activités manuelles. Par ailleurs, une maximisation des coûts de détection, des coûts des audits fiscaux ainsi que des coûts des ressources internes a été constatée au Yémen, en Iran et en Serbie avec l'implémentation de l'apprentissage hybride (Shujaaddeen et al., 2024 ; Rahimikia et al., 2017 et Savić et al., 2022).

En améliorant l'efficacité des opérations, cela produit des gains opérationnels qui occasionnent une augmentation significative des opérations de vérification avec une main-d'œuvre réduite et un

impact sur la qualité du service rendu aux contribuables. Cependant, ces études ne fournissent pas une estimation du montant des revenus fraudés et de l'économie de ressource effectuée.

#### **5.4.3 Effets et impacts sur la qualité du service rendu aux contribuables**

Les impacts stratégiques et opérationnels produisent un effet collatéral sur le niveau de qualité du service rendu aux contribuables. Selon Idrus (2024), l'usage des SIA entraîne plusieurs gains pour les contribuables tels que la réduction des honoraires payés aux conseils fiscaux lors des vérifications effectuées par les AF, l'amélioration de la transparence, une offre de service personnalisé.

Néanmoins, une diminution du niveau d'identification des contribuables imputables de fraude par erreur est constatée. L'utilisation des SIA permet aux AF de mieux cibler et personnaliser les contribuables auteurs de fraude. Cela peut également produire un effet d'équité parmi les contribuables, car les vérifications plus justes renforcent la croyance des contribuables aux AF. Par exemple, le réseau neuronal hybride utilisé par l'AF du Yémen (Shujaaddeen et al., 2024) a contribué à la maximisation des coûts de détection et des ressources internes avec une amélioration de la nomenclature des comportements des contribuables. Aussi, l'amélioration de la précision lors de la détection de la fraude liée à l'impôt sur les bénéfices commerciaux et industriels a bonifié la qualité du service.

Finalement, en axant les stratégies et les opérations de vérifications ciblées sur la hiérarchisation des risques, les AF démontrent une certaine transparence. Cela renforce leur légitimité et la perception qu'ont les contribuables. En plus, les AF font des économies, car elles concentrent leurs efforts et ressources sur des contribuables dont le risque de fraude est élevé. Au même moment, les contribuables ayant évité une vérification font également des économies en se gardant de payer des conseils fiscaux. L'utilisation par Belle Fille et al. (2023) de l'apprentissage supervisé pour détecter la fraude à la TVA a contribué à une meilleure planification et gestion plus efficace des risques fiscaux. L'AF du Rwanda a constaté une amélioration de la transparence dans l'utilisation des machines de facturation. Pérez López et al. (2019) ont conclu dans leur étude que l'AF de l'Espagne a utilisé l'apprentissage supervisé pour calculer la probabilité de fraude de chaque contribuable de l'échantillon. Ce qui leur a permis une planification réaliste des programmes de

vérifications fiscales. Non seulement les ressources sont optimisées, mais seuls les contribuables susceptibles de fraudes sont contrôlés.

Sur l'ensemble des 12 articles étudiés, il n'a pas été produit une recherche mettant en évidence la satisfaction des contribuables ainsi qu'une estimation des gains financiers que ceux-ci réalisent.

### **5.5 Limites des études analysées et de la revue de portée**

Comme toute étude, les analyses effectuées ne sont pas exemptes de limites. En ce qui concerne la mesure des effets et des impacts de l'usage des SIA pour la détection des comportements fiscaux frauduleux, il apparaît que les conclusions sont souvent d'ordre général et formulées sans données probantes.

D'abord, on constate une absence de preuves ou de documentation concernant les divers changements annoncés en matière de gestion organisationnelle et stratégique. En d'autres termes, les études n'expliquent pas comment les transformations en matière de gestion des ressources humaines, la restructuration des processus de vérification et l'introduction de l'automatisation des ciblage se sont opérées. Par exemple, Mohamed Maher Ben et AlSadhan (2023), García et Caballero (2023) ainsi que Lee (2022), indiquent que l'usage des SIA a contribué à l'optimisation de la priorisation stratégique, à la réduction des coûts opérationnels et à l'amélioration des activités de vérification. Mais ces études ne fournissent pas de détails sur les mécanismes mobilisés pour atteindre de tels résultats.

Ensuite, les études présentent peu d'impacts économiques chiffrés liés à l'utilisation des SIA. En dehors de l'étude menée par Zumaya et al. (2021) qui fournit une estimation du montant de l'évasion fiscale détectée (plus de 346 milliards de pesos mexicains), toutes les autres recherches restent silencieuses sur cette question. Les effets rapportés concernent principalement les mesures métriques tels que l'exactitude ou la précision, etc., sans toutefois établir une corrélation directe avec les gains fiscaux réels des AF.

Par ailleurs, les effets évoqués concernant les contribuables sont des affirmations non fondées sur des sondages ou des enquêtes. Il en résulte une absence d'examen sur les effets de l'utilisation des SIA sur les contribuables. À titre d'exemple, Shujaaddeen, Ba-Alwiet Al-Gaphari (2024) affirme

dans son analyse que la qualité du service est bonifiée sans toutefois apporter de preuves empiriques comme les résultats d'une enquête de satisfaction des contribuables ou des clients.

Plusieurs études concluent que leurs résultats pourraient être généralisés à d'autres AF et sur d'autres types d'impôts (Mohamed Maher Ben et AlSadhan, 2023 ; Belle Fille et al., 2022 ; Lee, 2022). Cependant, ces généralisations apparaissent peu fondées, car les spécificités des autres juridictions semblent ne pas être prises en considération. Il s'agit notamment des contextes juridiques, sociaux et de la maturité numérique qui peuvent fortement influencer l'efficacité des modèles proposés.

Ainsi, une des limites importantes est l'absence d'études portant sur les AF des pays développés comme le Canada, les États-Unis d'Amérique et les pays d'Europe. Une des hypothèses pouvant expliquer cette situation est la sensibilité de l'accès aux renseignements des contribuables. Par exemple au Canada, la *Loi sur la protection des renseignements personnels (L.R.C. (1985), ch. P-21)*, les lois administrées par l'ARC (LIR, LTA, etc.) et les questions éthiques rendent difficiles les collaborations universitaires avec les AF. Les échanges de renseignements sont effectués par des mécanismes complexes et encadrés par diverses lois relatives à la sécurité des informations du Gouvernement du Canada. Par exemple, l'article 241 de la *LIR* au Canada encadre de façon stricte les informations fiscales collectées par l'AF fédérale qu'est l'ARC. Pour terminer les études effectuées au sein des AF ne font pas l'objet de publication scientifique et restent des rapports corporatifs, non accessibles aux chercheurs et exclus de la RP. En définitive, ces limites permettent de porter une réflexion sur les perspectives de recherches futures.

## **5.6 Pistes de recherche futures**

Au vu des différentes limites analysées concernant l'utilisation des SIA dans la détection des comportements fiscaux frauduleux sur les AF, leurs effets et impacts apparaissent peu explorés. Afin de combler ce vide, plusieurs futures études pourraient être envisagées.

Premièrement, des études menées par l'OCDE (2022, p. 24) ont montré que plusieurs pays utilisent les SIA depuis plusieurs années dans la lutte contre les comportements fiscaux frauduleux. Il est préconisé des études à long terme au sein des AF, car la plupart des études analysées sont

ponctuelles. Plus précisément, il est recommandé d'effectuer chaque année des études sur les SIA implémentés dans une AF pour mesurer les effets et les impacts organisationnels à moyen terme.

Deuxièmement, des études devraient être menées sur les impacts économiques de l'utilisation des SIA dans les AF pour lutter contre les comportements fiscaux frauduleux. On peut citer à titre d'exemple l'étude d'impacts économiques au Mexique dont les estimations d'évasion s'élevaient en moyenne à plus de 259 millions de revenus par an. Cependant, l'étude n'indique pas le montant collecté réellement par l'AF. Des études quantitatives pourraient être effectuées dans les AF sur plusieurs années afin de déterminer le niveau réel des détections frauduleuses par les SIA, le nombre de vérifications faites, les revenus détectés et les montants frauduleux réellement recouvrés.

Troisièmement, plusieurs études dans l'échantillon mentionnent une amélioration de la qualité du service aux contribuables, due à l'utilisation des SIA sans nécessairement apporter des preuves. La perception des usagers peut être mesurée par l'intégration d'études qualitatives en menant des enquêtes de perception, de satisfaction, d'acceptabilité et de craintes liées à l'utilisation des SIA.

Quatrièmement, dans leur recherche, García et Caballero (2023) ont inclus un modèle d'IA explicable (*explainable AI*) dans leur processus de détection de la fraude à la TVA et de la fraude à l'impôt sur le revenu des entreprises par une méthode d'apprentissage hybride. De façon concrète, le modèle d'IA explicable permet non seulement aux AF de mieux comprendre le processus de détection de la fraude, mais également de rendre le processus plus transparent et une meilleure interprétabilité au profit des contribuables. Seul Savić et al. (2022), en Serbie, a intégré un modèle d'IA explicable dans leur recherche. L'adoption de ce modèle lors de l'usage des SIA contribuera à améliorer la transparence des décisions prises et l'acceptabilité des contribuables et des vérificateurs.

Une dernière piste à explorer concerne la comparaison des effets organisationnels et stratégiques entre deux AF. Comme indiqué dans le tableau de l'annexe F, plusieurs études ont fait cas de l'optimisation des opérations fiscales et de la planification stratégique des risques et des vérifications. Cependant, des études pourraient mener des explorations plus approfondies, au sein de deux AF ayant des structures organisationnelles similaires pour mesurer leur performance ainsi que les coûts opérationnels.

## CONCLUSION

---

L'amélioration de la détection et la prévention des comportements fiscaux frauduleux tels que la fraude fiscale, l'évasion fiscale et l'évitement fiscal a connu une transformation importante grâce aux processus progressifs de substitution des méthodes traditionnelles digitales par l'intégration des SIA au sein des AF. Malgré les différents niveaux d'efficacité observés, ces SIA doivent suivre des mécanismes rigoureux afin d'évaluer leur performance réelle au sein des AF nationales.

La méthodologie d'étude de cette RP s'est fondée sur le protocole *PRISMA* et les critères d'éligibilité du modèle *SPIDER*. Les résultats indiquent l'usage des méthodes d'apprentissage automatique composé d'apprentissage supervisé, non supervisé et hybride dans la détection des comportements fiscaux frauduleux. Ces technologies ont contribué à la détection de diverses fraudes à l'impôt direct et indirect. Les résultats ont également démontré une amélioration continue de la gestion stratégique des risques, de l'efficacité des opérations fiscales (vérifications et recouvrements) ainsi que de l'amélioration de la satisfaction des contribuables. Cependant, les niveaux de performance, l'efficacité et la fiabilité des SIA varient en fonction de la disponibilité et la qualité des données fiscales et financières provenant des AF.

Il existe une disparité du point de vue de l'implémentation de ces technologies par les États dans la lutte contre les comportements fiscaux frauduleux. L'intégration des SIA dans les AF nécessite des coûts d'investissement souvent onéreux tant en infrastructure qu'en ressources humaines. Comme l'introduction des SIA dans les AF est récente, alors l'évaluation des impacts organisationnels constitue une des principales limites de cette étude. Mais, la mise en place de mécanismes juridiques, éthiques et organisationnels transparents est primordiale afin d'assurer une acceptabilité des contribuables.

En fournissant un inventaire exhaustif des SIA utilisés par les AF pour détecter les comportements fiscaux frauduleux, cette RP a aussi mis en évidence les forces et les limites de la littérature existante en matière d'usage des SIA. Cela s'inscrit dans une démarche novatrice, palliant ainsi un manque de RP antérieure.

Enfin, quelques recherches futures ont été suggérées afin de répondre aux limites identifiées. Il est donc préconisé des études à long terme au sein des AF avec l'utilisation de données probantes pour mesurer les retombées économiques et organisationnelles rigoureuses de l'usage des SIA dans la

détection des comportements fiscaux frauduleux. L'intégration de l'IA explicable contribuera à l'amélioration de la perception des contribuables en leur fournissant une justification transparente et légitime des décisions prises lors des vérifications.

## BIBLIOGRAPHIE ET RÉFÉRENCES

---

### 1. Articles scientifiques

- Alm, J. (2012). Measuring, explaining, and controlling tax evasion: lessons from theory, experiments, and field studies. *International Tax and Public Finance*, 19(1), 54-77. <https://doi.org/10.1007/s10797-011-9171-2>
- Alsadhan, N. (2023). A Multi-Module Machine Learning Approach to Detect Tax Fraud. *Comput. Syst. Sci. Eng.*, 46(1), 241-253.
- Alstadsaeter, A., Godar, S., Nicolaidis, P., & Zucman, G. (2023). *Global Tax Evasion Report 2024*. <https://shs.hal.science/halshs-04563948>
- Atkinson, A. B. (1977). Optimal Taxation and the Direct versus Indirect Tax Controversy. *The Canadian Journal of Economics / Revue canadienne d'Économique*, 10(4), 590-606. <https://doi.org/10.2307/134292>
- Baghdasaryan, V., Davtyan, H., Sarikyan, A., & Navasardyan, Z. (2022). Improving Tax Audit Efficiency Using Machine Learning: The Role of Taxpayer's Network Data in Fraud Detection. *Applied Artificial Intelligence*, 36(1), 2012002. <https://doi.org/10.1080/08839514.2021.2012002>
- Bal, A. (2022). Black-Box Models as a Tool to Fight VAT Fraud. In B. Custers & E. Fosch-Villaronga (Eds.), *Law and Artificial Intelligence: Regulating AI and Applying AI in Legal Practice* (pp. 225-236). T.M.C. Asser Press. [https://doi.org/10.1007/978-94-6265-523-2\\_12](https://doi.org/10.1007/978-94-6265-523-2_12)
- Barilari, A. (2018). La fraude fiscale : les mots et les chiffres [Tax fraud: words and numbers]. *Gestion et Finances Publiques*, (3), 50-57. <https://doi.org/10.3166/gfp.2018.00034>
- Baron, J. (2018). A brief history of evidence-based policy. *The annals of the american academy of political and social science*, 678(1), 40-50. <https://doi.org/10.1177/0002716218763128>
- Bassey, E., Mulligan, E., & Ojo, A. (2022). A conceptual framework for digital tax administration - A systematic review. *Government Information Quarterly*, 39(4), 101754. <https://doi.org/10.1016/j.giq.2022.101754>
- Baumann, M., Boehm, T., Knoll, B. & Riedel, N. (2020). Corporate Taxes, Patent Shifting, and Anti-avoidance Rules: Empirical Evidence. *Public Finance Review*, 48(4), 467–504. <https://doi.org/10.1177/1091142120930684>
- Beebejaun, A. (2021). Mauritius as an offshore financial centre and laws relating to tax avoidance and evasion. *Journal of Financial Crime*, 28(2), 548-565. <https://doi.org/10.1108/JFC-08-2020-0160>
- Belavagi, M. C., & Muniyal, B. (2016). Performance Evaluation of Supervised Machine Learning Algorithms for Intrusion Detection. *Procedia Computer Science*, 89, 117-123. <https://doi.org/https://doi.org/10.1016/j.procs.2016.06.016>
- Belavagi, M. C., & Muniyal, B. (2016). Performance Evaluation of Supervised Machine Learning Algorithms for Intrusion Detection. *Procedia Computer Science*, 89, 117-123. <https://doi.org/https://doi.org/10.1016/j.procs.2016.06.016>
- Belle Fille, M., Jean Felicien, I., Kayijuka, I., Nzabanita, J., & Haughton, D. (2023). Comparison of Tree-Based Machine Learning Algorithms to Predict Reporting Behavior of Electronic Billing Machines. *Information*, 14(3), 140. <https://doi.org/https://doi.org/10.3390/info14030140>

- Belle Fille, M., Tuyishimire, O., Haughton, D., & Nzabanita, J. (2022). Fraud Detection Using Neural Networks: A Case Study of Income Tax. *Future Internet*, 14(6), 168. <https://doi.org/https://doi.org/10.3390/fi14060168>
- Boita, M., Paiusan, L., Ajtay, E. (2021). Underevaluation of the taxable base as a result of policies aggressive accountants source generator of tax fraud. *Ovidius University Annals, Economic Sciences Series*, 21(1), 697-702. <https://stec.univ-ovidius.ro/html/anale/RO/2021/Section%205/3.pdf>
- Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science*, 17(3), 235-255.
- Cockfield, A. J. (2015). Bid data and tax haven secrecy. *Fla. Tax Rev.*, 18, 483. <http://banques.enap.ca:2048/login?url=https://search.proquest.com/?url=https://www.proquest.com/scholarly-journals/big-data-tax-haven-secrecy/docview/2614125354/se-2?accountid=10748>
- Colquhoun, H. L., Levac, D., O'Brien, K. K., Straus, S., Tricco, A. C., Perrier, L., Kastner, M., & Moher, D. (2014). Scoping reviews: time for clarity in definition, methods, and reporting. *Journal of Clinical Epidemiology*, 67(12), 1291-1294. <https://doi.org/https://doi.org/10.1016/j.jclinepi.2014.03.013>
- Cooke, A., Smith, D., & Booth, A. (2012). Beyond PICO: the SPIDER tool for qualitative evidence synthesis. *Qualitative Health Research*, 22(10), 1435-1443. <https://doi.org/10.1177/1049732312452938>
- Crabtree, B., & Miller, W. (1999). Using codes and code manuals: A template organizing style of interpretation. In B. Crabtree & W. Miller (Eds.), *Doing qualitative research* (pp. 163–178). Sage.
- De La Fera, R. (2020). Tax fraud and selective law enforcement. *Journal of law and Society*, 47(2), 240-270. <https://dx.doi.org/10.1111/jols.12221>
- Deen, K., Kemsley, S. A., & Morgan, F. T. (2022). Tax evasion and money laundering: a complete framework. *Journal of Financial Crime*, 29(2), 589-602. <https://doi.org/https://doi.org/10.1108/JFC-09-2020-0175>
- Delgado, F. J., Fernández-Rodríguez, E., García-Fernández, R., Landajo, M., & Martínez-Arias, A. (2023). Tax avoidance and earnings management: a neural network approach for the largest European economies. *Financial Innovation*, 9(1), 19. <https://doi.org/10.1186/s40854-022-00424-8>
- Ebers, M., & Tupay, P. K. o. t. (2023). Artificial intelligence and machine learning powered public service delivery in Estonia: opportunities and legal challenges. *Springer*. <https://doi.org/10.1007/978-3-031-19667-6>
- Fabiano, N., Gupta, A., Bhambra, N., Luu, B., Wong, S., Maaz, M., Fiedorowicz, J. G., Smith, A. L., & Solmi, M. (2024). How to optimize the systematic review process using AI tools. *JCPP Advances*, 4(2), e12234. <https://doi.org/https://doi.org/10.1002/jcv2.12234>
- Faúndez-Ugalde, A., Mellado-Silva, R., & Aldunate-Lizana, E. (2020). Use of artificial intelligence by tax administrations: An analysis regarding taxpayers' rights in Latin American countries. *Computer Law & Security Review*, 38, 105441. <https://doi.org/10.1016/j.clsr.2020.105441>
- Franzoni, L. A. (1998). Tax evasion and tax compliance. *Available at SSRN 137430*. <https://doi.org/https://dx.doi.org/10.2139/ssrn.137430>

- Ftouhi, K. & Ghardallou, W. (2020). International tax planning techniques: A review of the literature. [International tax planning techniques] *Journal of Applied Accounting Research*, 21(2), 329-343. doi: <https://doi.org/10.1108/JAAR-05-2019-0080>
- García, I. G., & Caballero, A. M. (2023). A Comparison Between Bayesian Dialysis and Machine Learning to Detect Tax Fraud and Its Causes: The Case of Vat, Corporate Tax and Customs Duties in Spain. *SN Computer Science*, 4(1), 80. <https://link-springer-com.proxy.bib.uottawa.ca/article/10.1007/s42979-022-01483-5>
- Goel, L., Kale, R. W., Zhang, J. Z., & Arora, D. (2022). Agility and Resilience in Information Systems Research: A Thematic Analysis. *Journal of Organizational and End User Computing (JOEUC)*, 33(6), 1-17. <https://doi.org/10.4018/JOEUC.291510>
- Goumagias, N. D., Hristu-Varsakelis, D., & Assael, Y. M. (2018). Using deep Apprentissage Q to understand the tax evasion behavior of risk-averse firms. *Expert Systems with Applications*, 101, 258-270. <https://doi.org/10.1016/j.eswa.2018.01.039>
- Gravelle, J. G. (2009). Tax havens: International tax avoidance and evasion. *National Tax Journal*, 62(4), 727-753.
- Guias, E., & Haineala, C. (2021). Tax avoidance and tax evasion in eu: trends and effects. *The annals of the university of Oradea. economic sciences*, 30, 229-238. [https://doi.org/10.47535/1991AUOES30\(2\)024](https://doi.org/10.47535/1991AUOES30(2)024)
- Herrera-Semenets, V., Bustio-Martínez, L., González-Ordiano, J. Á., & van den Berg, J. (2024). Tax Underreporting Detection Using an Unsupervised Learning Approach. Mexican International Conference on Artificial Intelligence,
- Hsieh, H.-F., & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, 15(9), 1277-1288. <https://doi.org/10.1177/1049732305276687>
- Huang, Z. (2018) Discussion on the Development of Artificial Intelligence in Taxation. *American Journal of Industrial and Business Management*, 8, 1817-1824. doi: [10.4236/ajibm.2018.88123](https://doi.org/10.4236/ajibm.2018.88123).
- Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O' Cathain, A., Rousseau, M.-C., Vedel, I., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information*, 34(4), 285-291. <https://doi.org/10.3233/efi-180221>
- Huong, V. T. L., Van, N. T. T., & Huyen, V. T. T. (2023). Digital transformation of tax in academic research: A literature review. *World Journal of Advanced Research and Reviews*, 20(01), 610–618. <https://doi.org/https://doi.org/10.30574/wjarr.2023.20.1.2092>
- Idrus, M. (2024). Efficiency of Tax Administration and Its Influence on Taxpayer Compliance. *Economics and Digital Business Review*, 5(2), 889-913.
- Ihnatišínová, D. (2021). Digitalization of tax administration communication under the effect of global megatrends of the digital age. *Les Ulis: EDP Sciences*, (92). <https://doi.org/10.1051/shsconf/20219202022>
- Ioana - Florina, C., Laura - Camelia, F., & Eliza-Angelika, K. (2021). Tax evasion and financial fraud in the current digital context. *Annals of the University of Oradea: Economic Science*, 30(1), 187–194. <https://enap.on.worldcat.org/oclc/9456438303>
- Ippolito, A., & Lozano, A. C. G. (2020). Tax Crime Prediction with Machine Learning: A Case Study in the Municipality of São Paulo. *ICEIS* (1),

- Jacob, S., & Lawarée, J. (2022). Les mesures publiques dans les stratégies gouvernementales en matière d'intelligence artificielle : une perspective internationale. In: Observatoire international sur les impacts sociétaux de l'intelligence artificielle et du numérique, Université Laval.
- Lagodiienko, N., Pozhydaieva, M., & Krylov, D. (2022). Digitalization of tax administration in Ukraine: risks and opportunities. *Management Theory and Studies for Rural Business and Infrastructure Development*, 44(4), 443-450. <https://dx.doi.org/10.15544/mts.2022.44>
- Lambert, S., & O'Halloran, E. (2008). Deductive thematic analysis of a female paedophilia website. *Psychiatry, Psychology and Law*, 15(2), 284-300. <https://doi.org/10.1080/13218710802014469>
- Latulippe, R., Martel, L. (2021). Utilisation de l'intelligence artificielle en fiscalité : état de situation selon la littérature. [https://cfff.recherche.usherbrooke.ca/wp-content/uploads/2021/01/R2021-01\\_Intelligence\\_artificielle\\_fiscalite.pdf](https://cfff.recherche.usherbrooke.ca/wp-content/uploads/2021/01/R2021-01_Intelligence_artificielle_fiscalite.pdf)
- Lawarée, J., Jacob, S., & Ouimet, M. (2020). A scoping review of knowledge syntheses in the field of evaluation across four decades of practice. *Evaluation and Program Planning*, 79, 101761. <https://dx.doi.org/10.1016/j.evalprogplan.2019.101761>
- Lee, C. (2022). Deep learning-based detection of tax frauds: an application to property acquisition tax. *Data Technologies and Applications*, 56(3), 329-341. <https://doi.org/10.1108/DTA-06-2021-0134>
- Mahesar, A. J., Wighio, A. A., Imtiaz, N., Jamali, A., Nawaz, Y., & Urooj, U. (2025). Predicting tax evasion using machine learning: A study of e-commerce transactions. *Spectrum of Engineering Sciences*, 3(4), 840-852.
- Mohamed Maher Ben, I., & AlSadhan, N. (2023). Simultaneous Classification and Regression for Zakat Under-Reporting Detection. *Applied Sciences*, 13(9), 5244. <https://doi.org/https://doi.org/10.3390/app13095244>
- Moher, D., Altman, D., Liberati, A., Tetzlaff, J., Altman, D. G., Antes, G., Atkins, D., Barbour, V., Barrowman, N., Berlin, J. A., Clark, J., Clarke, M., Cook, D., D'Amico, R., Deeks, J., Devereaux, P. J., Dickersin, K., Egger, M., Ernst, E., . Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine*, 6(7), e1000097. <https://doi.org/https://dx.doi.org/10.1371/journal.pmed.1000097>
- Murorunkwere, B. F., Houghton, D., Nzabanita, J., Kipkoge, F., & Kabano, I. (2023). Predicting tax fraud using supervised machine learning approach. *African Journal of Science, Technology, Innovation and Development*, 15(6), 731-742. <https://doi.org/10.1080/20421338.2023.2187930>
- N. Visitpanya and T. Samanchuen (2023). "A Machine Learning Approach to Identifying Suspicious Tax Evasion Behaviour in Public Financial Data," *2023 8th International Conference on Business and Industrial Research (ICBIR)*, Bangkok, Thailand, pp. 1152-1158, doi: 10.1109/ICBIR57571.2023.10147479.
- Pamisetty, V. (2019). Machine Learning Models for Real-Time Tax Fraud Detection and Risk Assessment in Digital Government Systems. *Global Research Development (GRD) ISSN: 2455-5703*, 4(12).
- Pérez López, C., Delgado Rodríguez, M. J., & de Lucas Santos, S. (2019). Tax Fraud Detection through Neural Networks: An Application Using a Sample of Personal Income Taxpayers. *Future Internet*, 11(4), 86. <https://doi.org/10.3390/fi11040086>

- Pires, C. A. (2024). Supervised Machine Learning for Tax Evasion Detection: A Case Study with the Brazilian Tax Administration. *Anais Estendidos do XXXIX Simpósio Brasileiro de Banco de Dados (SBBBD Estendido 2024)*.
- Pollock, D., Evans, C., Menghao Jia, R., Alexander, L., Pieper, D., Brandão de Moraes, É., Peters, M. D. J., Tricco, A. C., Khalil, H., Godfrey, C. M., Saran, A., Campbell, F., & Munn, Z. (2024). “How-to”: scoping review? *Journal of Clinical Epidemiology*, 176, 111572. <https://doi.org/https://doi.org/10.1016/j.jclinepi.2024.111572>
- Rahayu, S.K., (2021). Utilization of artificial intelligence in tax audit in Indonesia. *Manage. Account. Rev.* 20, 135–157.
- Rahimikia, E., Mohammadi, S., Rahmani, T., & Ghazanfari, M. (2017). Detecting corporate tax evasion using a hybrid intelligent system: A case study of Iran. *International Journal of Accounting Information Systems*, 25, 1-17. <https://doi.org/10.1016/j.accinf.2016.12.002>
- Rahman, R. A., Masrom, S., & Omar, N. (2019). Tax avoidance detection based on machine learning of malaysian government-linked companies. *International Journal of Recent Technology and Engineering*, 8(2), 535-541.
- Raikov, A. (2021). Decreasing Tax Evasion by Artificial Intelligence. *IFAC PapersOnLine*, 54(13), 172–177. <https://doi.org/10.1016/j.ifacol.2021.10.440>
- Rathi, A., Sharma, S., Lodha, G., & Srivastava, M. (2021). A study on application of artificial intelligence and machine learning in Indian taxation system. *Psychology and Education Journal*, 58(2), 1226-1233. <http://dx.doi.org/10.17762/pae.v58i2.2265>
- Reis, F. J. J., Alaiti, R. K., Vallio, C. S., & Hespanhol, L. (2024). Artificial intelligence and Machine Learning approaches in sports: Concepts, applications, challenges, and future perspectives. *Brazilian Journal of Physical Therapy*, 28(3), 101083. <https://doi.org/https://doi.org/10.1016/j.bjpt.2024.101083>
- Saragih, A. H., & Ali, S. (2023). Corporate tax risk: a literature review and future research directions. *Management Review Quarterly*, 73(2), 527-577. <https://doi.org/10.1007/s11301-021-00251-8>
- Sarker, T., & Ahmed, M. S. (2022). The role of government reform in improving voluntary tax compliance in the digital economy: the Bangladesh experience. in *Taxation in the Digital Economy*, 240-259,. Routledge, <https://doi.org/10.4324/9781003196020-14>
- Savić, M., Atanasijević, J., Jakovetić, D., & Krejić, N. (2022). Tax evasion risk management using a Hybrid Unsupervised Outlier Detection method. *Expert Systems with Applications*, 193, 116409. <https://doi.org/https://doi.org/10.1016/j.eswa.2021.116409>
- Savić, M., Atanasijević, J., Jakovetić, D., & Krejić, N. (2022). Tax evasion risk management using a Hybrid Unsupervised Outlier Detection method. *Expert Systems with Applications*, 193, 116409. <https://doi.org/https://doi.org/10.1016/j.eswa.2021.116409>
- Saxunova, D., & Szarkova, R. (2018). Global efforts of tax authorities and tax evasion challenge. *Journal of Eastern Europe Research in Business and Economics*, 2018, 1-14. <http://dx.doi.org/10.5171/2018.511388>
- Shakil, M. H., & Tasnia, M. (2022). Artificial intelligence and tax administration in asia and the pacific. In *Taxation in the Digital Economy*, 45-55, Routledge. <https://doi.org/10.4324/9781003196020-4>
- Shi, B., Dong, B., Xu, Y., Wang, J., Wang, Y. & Zheng, Q. (2023). An edge feature aware heterogeneous graph neural network model to support tax evasion detection. *Expert Systems with Applications*, 213. <https://doi.org/10.1016/j.eswa.2022.118903>

- Shujaaddeen, A., Ba-Alwi, F. M., & Al-Gaphari, G. (2024). A New Machine Learning Model for Detecting levels of Tax Evasion Based on Hybrid Neural Network [Article]. *International Journal of Intelligent Systems and Applications in Engineering*, 12(11s), 450-468. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85184869798&partnerID=40&md5=23a6776eb07fdcd4cba45aac24ebc65c>
- St-Amant, G. (2005). E-Gouvernement : cadre d'évolution de l'administration électronique. *Systèmes d'Information et Management*, 10(1), 15-38. <http://banques.enap.ca:2048/login?url=https://search.proquest.com/?url=https://www.proquest.com/scholarly-journals/e-gouvernement-cadre-dévolution-de/docview/222261844/se-2?accountid=10748>
- Tian, F., Lan, T., Chao, K. M., Godwin, N., Zheng, Q., Shah, N., & Zhang, F. (2016). Mining Suspicious Tax Evasion Groups in Big Data. *IEEE Transactions on Knowledge and Data Engineering*, 28(10), 2651-2664. <https://doi.org/10.1109/TKDE.2016.2571686>
- Vanhoeyveld, J., Martens, D., & Peeters, B. (2020). Value-added tax fraud detection with scalable anomaly detection techniques. *Applied Soft Computing*, 86, 105895. <https://doi.org/https://doi.org/10.1016/j.asoc.2019.105895>
- Vernier, É. (2018). Chapitre 1. La fraude fiscale : concepts juridiques et données statistiques. In *Fraude fiscale et Paradis fiscaux* (pp. 13-31). Dunod. <https://www.cairn.info/fraude-fiscale-et-paradis-fiscaux--9782100784684-page-13.htm>
- Wang, F., Xu, S., Sun, J., & Cullinan, C. P. (2020). Corporate tax avoidance: a literature review and research agenda. *Journal of Economic Surveys*, 34(4), 793-811. <https://doi.org/https://dx.doi.org/10.1111/joes.12347>
- Wei, R., Dong, B., Zheng, Q., Zhu, X., Ruan, J., & He, H. (2019). Unsupervised conditional adversarial networks for tax evasion detection. 2019 IEEE international conference on big data (Big Data)
- 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>
- Xavier, O. C., Pires, S. R., Marques, T. C., & Soares, A. d. S. (2022). Tax evasion identification using open data and artificial intelligence. *Revista de Administração Pública*, 56, 426-440. <https://doi.org/10.1590/0034-761220210256x>
- Zhang, Y., Tiño, P., Leonardis, A., & Tang, K. (2021). A survey on neural network interpretability. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 5(5), 726-742.
- Zheng, Q., Xu, Y., Liu, H., Shi, B., Wang, J., & Dong, B. (2023). A survey of tax risk detection using data mining techniques. *Engineering*. <https://doi.org/10.1016/j.eng.2023.07.014>
- Zumaya, M., Guerrero, R., Islas, E., Pineda, O., Gershenson, C., Iñiguez, G., & Pineda, C. (2021). Identifying Tax Evasion in Mexico with Tools from Network Science and Machine Learning. In *Understanding Complex Systems* (pp. 89-113). Springer Science and Business Media Deutschland GmbH. [https://doi.org/10.1007/978-3-030-81484-7\\_6](https://doi.org/10.1007/978-3-030-81484-7_6)

## 2. Communications présentées à des congrès

- De Roux, D., Perez, B., Moreno, A., Villamil, M. d. P., & Figueroa, C. (2018). Tax fraud detection for under-reporting declarations using an unsupervised machine learning

- approach. Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. <https://doi.org/10.1145/3219819.3219878>
- Kumar, R., Malholtra, R. K., & Grover, C. N. (2023). Review on artificial intelligence role in implementation of Goods and Services Tax (GST) and future scope. International conference on Artificial Intelligence and Smart Communication (AISC). Greater Noida, India. <https://doi.org/10.1109/AISC56616.2023.10085030>
- Zheng, Q., Lin, Y., He, H., Ruan, J., & Dong, B. (2019). ATTENet: Detecting and Explaining Suspicious Tax Evasion Groups. IJCAI. <http://dx.doi.org/10.24963/ijcai.2019/964>
- Zhu, X., Yan, Z., Ruan, J., Zheng, Q., & Dong, B. (2018). IRTED-TL: an inter-region tax evasion detection method based on transfer learning. 17th IEEE international conference on trust, security and privacy in computing and communications/ 12th IEEE international conference on big data science and engineering (TrustCom/BigDataSE). New York, NY, USA. <https://doi.org/10.1109/TrustCom/BigDataSE.2018.00169>
- Zhou, L. (2019). Opportunities and challenges of artificial intelligence in the application of taxation system. International conference on economic management and cultural industry (ICEMCI 2019). <http://dx.doi.org/10.2991/aebmr.k.191217.038>

### 3. Livres

- Li, H., Lin, L., & Zeng, H. (2024). *Machine learning methods*. Springer. <https://dx.doi.org/10.1007/978-981-99-3917-6>
- Paillé, P., & Mucchielli, A. (2008). *L'analyse qualitative en sciences humaines et sociales* (2e ed.). Armand Colin. <https://doi.org/10.3917/arco.paill.2016.01>
- Shaw, T. M., Mahrenbach, L. C., Modi, R., & Xu, Y.-C. (2019). The Palgrave handbook of contemporary international political economy. Springer. <https://doi.org/10.1057/978-1-137-45443-0>

### 4. Publications gouvernementales

- Agence du revenu du Canada (2016). *Écart fiscal au Canada : une étude conceptuelle*. Agence du revenu du Canada = Canada Revenue Agency. [http://publications.gc.ca/collections/collection\\_2017/arc-cra/Rv4-84-2016-fra.pdf](http://publications.gc.ca/collections/collection_2017/arc-cra/Rv4-84-2016-fra.pdf)
- Agence du revenu du Canada (2016). Norme commune de déclaration. Gouvernement du Canada. <https://www.canada.ca/fr/agence-revenu/programmes/a-propos-agence-revenu-canada-arc/budgets-gouvernement-federal/budget-2016-assurer-croissance-classe-moyenne/norme-commune-declaration.html>
- Agence du revenu du Canada (s.d.). Comment nous luttons contre l'évasion fiscale et l'évitement fiscal. <https://www.canada.ca/fr/agence-revenu/programmes/a-propos-agence-revenu-canada-arc/observation/comment-luttons-contre-evasion-fiscale-evitement-iscal.html>
- Agence du revenu du Canada (s.d.). Écart fiscal et résultats en matière d'observation pour le système d'impôt fédéral sur le revenu des sociétés. <https://www.canada.ca/fr/agence-revenu/organisation/a-propos-agence-revenu-canada-arc/ecart-fiscal-canada-etude-conceptuelle/ecartfiscal-resultats-conformites.html>
- Assemblée nationale. (2024). Rapport d'information n° 2742 sur la lutte contre la fraude fiscale. [https://www.assemblee-nationale.fr/dyn/16/rapports/cion\\_fin/116b2742\\_rapport-information](https://www.assemblee-nationale.fr/dyn/16/rapports/cion_fin/116b2742_rapport-information)

- Canada, P., Easter, W., Easter, L. h. W., & Bibliothèque numérique, c. (2016). *L'Agence du revenu du Canada, l'évitement fiscal et l'évasion fiscale : Mesures recommandées*. Chambres des Communes. <https://www.deslibris.ca/ID/10063953>  
<https://doi.org/10.1787/bcbdaf1a-fr>.
- G-7 and G-20 (s.d.) U.S. Department of the Treasury. <https://home.treasury.gov/policy-issues/international/g-7-and-g-20>
- Internal Revenue Service. (IRS). (2021). 25.1.1 Overview/Definitions | Internal Revenue Service.
- International Association of Privacy Professionals (IAPP). (2023). International definitions of AI. [https://iapp.org/media/pdf/resource\\_center/international\\_definitions\\_of\\_ai.pdf](https://iapp.org/media/pdf/resource_center/international_definitions_of_ai.pdf)
- Ministère des finances du Québec. (2017). Mémoire – Paradis fiscaux : enjeux et perspectives. [https://www.finances.gouv.qc.ca/documents/autres/fr/autfr\\_memoireparadisfiscaux.pdf](https://www.finances.gouv.qc.ca/documents/autres/fr/autfr_memoireparadisfiscaux.pdf)
- Network, T. J. (2023). *The State of Tax Justice 2023*.
- Organisation de coopération et de développement économiques (OCDE) (2023), Statistiques des recettes publiques 2023 : Dynamisme des recettes fiscales dans les pays de l'OCDE, Éditions OCDE, Paris,
- Organisation de coopération et de développement économiques (OCDE). (s.d.) Inventaire des initiatives de technologie fiscale. OCDE. <https://www.oecd.org/fr/data/datasets/inventory-of-tax-technology-initiatives.html>
- Organisation de coopération et de développement économiques (OCDE). (s.d.). L'échange automatique d'informations. <https://web-archiv.oecd.org/fr/2015-08-25/69329-echangeautomatique.htm>
- Organisation de coopération et de développement économiques (OCDE). (s.d.). Forum mondial sur la transparence et l'échange de renseignements à des fins fiscales. <https://www.oecd.org/fr/networks/global-forum-tax-transparency/who-we-are/history.html>
- Organisation de coopération et de développement économiques (OCDE). (s.d.). Convention concernant l'assistance administrative mutuelle en matière fiscale. <https://www.oecd.org/fr/topics/sub-issues/convention-on-mutual-administrative-assistance-in-tax-matters.html>
- Organisation de coopération et de développement économiques (OCDE) (s.d.). Base Erosion and Profit Shifting (BEPS). OCDE. <https://www.oecd.org/en/topics/policy-issues/base-erosion-and-profit-shifting-beps.html>
- Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO). (2022). Recommandation sur l'éthique de l'intelligence artificielle. [https://unesdoc.unesco.org/ark:/48223/pf0000381137\\_fre](https://unesdoc.unesco.org/ark:/48223/pf0000381137_fre)
- Organisation for Economic Co-operation and Development (OECD). (1987). L'évasion et la fraude fiscales internationales. <https://doi.org/doi:https://doi.org/10.1787/e5d87800-fr>
- Organisation for Economic Co-operation and Development (OECD). (2019). *Tax gap*. <https://doi.org/doi:https://doi.org/10.1787/40126618-en>
- Organisation for Economic Co-operation and Development (OECD). (2022). *Tax Administration 2022 - Moving towards digital transformation*. <https://www.oecd-ilibrary.org/content/publication/1e797131-en>
- Organisation for Economic Co-operation and Development (OECD). (2022). *Administration fiscale 2021*. <https://doi.org/doi:https://doi.org/10.1787/d48f8dc3-fr>

- Revenu Québec. (s.d.). Conventions fiscales internationales. <https://www.revenuquebec.ca/fr/citoyens/votre-situation/nouvel-arrivant/declaration-de-revenus/conventions-fiscales-internationales/>
- Société d'évaluation foncière des municipalités (MPAC). (s.d.). Site officiel de la MPAC. <https://www.mpac.ca/fr>
- Tax Justice Network. (2021). État de la justice fiscale 2021. Tax Justice Network. [The State of Tax Justice 2021 - Tax Justice Network](#)
- Tax Justice Network. (2021). État de la justice fiscale 2024. Tax Justice Network. [The State of Tax Justice 2021 - Tax Justice Network](#)
- U.S. Department of the Treasury. (2023, October 20). Treasury releases 2023 national illicit finance strategy. <https://home.treasury.gov/news/press-releases/jy2650>

## 5. Pages web

- Altus Group. (s.d.). Accélérer la numérisation de la fiscalité foncière. <https://www.altusgroup.com/insights/advancing-property-tax-digitalization/?lang=fr>
- DistillerSR. (s.d.) DistillerSR systematic review software. Evidence Partners. <https://www.distillersr.com/products/distillersr-systematic-review-software>
- HuffPost. (2024, février 15). Fraude fiscale : les contrôles aériens ont permis de trouver 140 000 piscines non déclarées en 2023). [https://www.huffingtonpost.fr/economie/article/fraude-fiscale-les-controles-aeriens-ont-permis-de-trouver-140-000-piscines-non-declarees-en-2023\\_235310.html](https://www.huffingtonpost.fr/economie/article/fraude-fiscale-les-controles-aeriens-ont-permis-de-trouver-140-000-piscines-non-declarees-en-2023_235310.html)

## 6. Mémoires et thèses

- Lawarée, J (2017). *L'intégration d'une innovation politico-administrative au prisme de l'action des acteurs : le cas des pôles de compétitivité en Wallonie de 2005-2014*. [Thèse de doctorat, Université catholique de Louvain. Espace ENAP. <https://espace.enap.ca/id/eprint/379/>

## 7. Lois et documents législatifs

- Loi de l'impôt sur le revenu. L.R.C. (1985), ch. 1 (5e suppl.). <https://laws-lois.justice.gc.ca/fra/lois/i-3.3/>
- Loi sur la taxe d'accise (L.R.C. (1985), ch. E-15). <https://laws-lois.justice.gc.ca/fra/lois/e-15/>
- The Income Tax Act 1995 (Consolidated Version). <https://www.mra.mu/download/ITAConsolidated.pdf>

## ANNEXES

### ANNEXE A

#### Stratégie de recherche

Français				
(« Ministère du Revenu national » OU « Ministère des Finances » OU « Agences fiscales » OU « Administrations fiscales » OU « Autorités fiscales » OU « Services fiscaux » OU « Bureaux fiscaux »)	ET	(« Apprentissage supervisé » OU « Apprentissage par renforcement » OU « Apprentissage automatique » OU « Système d'intelligence artificielle » OU « Intelligence artificielle » OU « Réseaux de neurones » OU « Apprentissage profond » OU « Traitement du langage naturel » OU « Les grands modèles de langage » OU « Agents virtuels intelligents » OU « Apprentissage non supervisé » OU « Algorithme » OU « Vision par ordinateur » OU « Automatisation »)	ET	(« Fraude fiscale » OU « Fraude à la TVA » OU « Fraude d'impôt sur le revenu » OU « Évasion fiscale » OU « Évasion fiscale des entreprises » OU « Évitement fiscal des entreprises » OU « Fraude impôt sur revenus d'entreprise » OU « Évitement fiscal » OU « Planification fiscale abusive »)

Anglais				
("Minister of National Revenue" OR "Department of Finance" OR "Tax agenc*" OR "Tax Administrations" OR "Tax Authorit*" OR "Tax Services" OR "Tax office" )	AND	("Supervised learning" OR OR "Reinforcement learning" OR "Artificial intelligence systems" OR "Artificial intelligence" OR "Neural networks" OR "Deep learning" OR "Natural language modal processing" OR "Large language model" OR "Virtual bots" OR "Unsupervised learning" OR "Algorithm" OR "Computer vision" OR "Automatisation")	AND	("Tax fraud" OR "VAT fraud" OR "Income tax fraud" OR "Tax evasion" OR "Corporate tax evasion" OR "Corporate tax avoidance" OR "Corporate tax fraud" OR "Tax avoidance" OR "Aggressive tax planning")

## ANNEXE B

### Base de données et nombre de références

Base de données	Dates de recherche	Nombre de références	Périodes de publication	Types d'articles
ABIIFORM	09-05-2024 au 28-08-2024	565	1935-2024	<ul style="list-style-type: none"> <li>- Articles de revue scientifique</li> <li>- Livres</li> <li>- Actes de conférence</li> <li>- Sections et chapitres de livre</li> <li>- Rapports</li> <li>- Article de journal</li> <li>- Article de presse</li> <li>- Thèses</li> <li>- Article de magazine</li> <li>- Génériques</li> </ul>
EBSCOhost	09-05-2024 au 28-08-2024	10	2010-2023	
Emerald Journals	09-05-2024 au 28-08-2024	500	1928-2024	
Google Scholar	08-09-2024	81	2003-2024	
IEEE	09-05-2024 au 28-08-2024	13	2011-2024	
JSTOR	09-05-2024 au 28-08-2024	2 409	1959-2024	
Sciencedirect	09-05-2024 au 28-08-2024	281	1963-2024	
SCOPUS	09-05-2024 au 28-08-2024	25	2018-2024	
Web of science	09-05-2024 au 28-08-2024	12	2018-2024	
Total		3 896		

## ANNEXE C

---

### L'article en cause : Visual querying and analysis of temporal fiscal networks

1/11/25, 1:53 AM

Mail - Dao, Adama - Outlook



---

**Re: Question about your article: Visual querying and analysis of temporal fiscal networks**

---

**From** Luca Grilli <luca.grilli@unipg.it>

**Date** Wed 2025-01-08 11:33

**To** Dao, Adama <Adama.Dao@enap.ca>; walter.didimo@unipg.it <walter.didimo@unipg.it>;  
fabrizio.montecchiani@unipg.it <fabrizio.montecchiani@unipg.it>; daniele.pagliuca@agenziaentrate.it  
<daniele.pagliuca@agenziaentrate.it>

Dear Dao,

Thank you for your question.

The answer to your inquiry is no, TeFNet cannot be considered a system that incorporates artificial intelligence. While TeFNet utilizes graph database technologies and dynamic visualization tools, it does not leverage machine learning or other AI techniques.

However, in a subsequent work, we introduced a new system called MALDIVE, which does make use of data mining and machine learning techniques. In this sense, MALDIVE can be viewed as an extension of TeFNet, incorporating AI capabilities.

I hope this clarifies your doubt. Please feel free to reach out if you have any further questions.

Best regards,  
Luca Grilli

On 28/12/24 21:42, Dao, Adama wrote:

Hello dear authors, Walter, Luca, Giuseppe, Fabrizio and Daniele

First of all, I would like to wish you a very happy holiday season.  
Happy new year 2025 with even more intellectual production.

I am a research master's student at the National School of Public Administration of Quebec in Canada. I would like to have a light on the TeFNet (Tax evasion and Fraud Network). You said that its main ingredients are: the visual query language for graph databases and the dynamic graph visualization tools.

My question is whether this association can be considered artificial intelligence because it is not clearly mentioned in the article.

Thank you for your help.

Sent from Outlook for iOS<<https://aka.ms/o0ukef>>

<https://outlook.office.com/mail/inbox/id/AAMkADY5N2Y3OWJhLTg2YjUjNDkNj04ZmI5LWQ5YjY5ODVmNzk5ZjBGAAAAAABB5oF1GoYYTID9%2B...> 1/1

## ANNEXE D

Caractéristiques des publications, données analysées et objectifs de recherche

Articles	Objectif de recherche	Auteurs	Données mobilisées	Pays et AF
Comparison of Tree-Based Machine Learning Algorithms to Predict Reporting Behavior of Electronic Billing Machines	Prédiction de la performance des machines à facturation électronique	(Belle Fille et al., 2023)	852 239 enregistrements, Janvier 2013 à mai 2018	Rwanda (Afrique de l'Est) : Autorité du revenu du Rwanda ( <i>Rwanda Revenue Authority</i> )
Fraud Detection Using Neural Networks: A Case Study of Income Tax	Évaluation de l'efficacité des réseaux neuronaux pour identifier les facteurs de fraude fiscale	(Belle Fille et al., 2022)	7840 contribuables (entreprises et individus) audités (1 655 fraudeurs et 6 185 non fraudeurs), pas de date	
Predicting Tax Fraud Using Supervised Machine Learning Approach	Prédiction de la fraude via apprentissage machine supervisé	(Murorunkwere et al., 2023)	15 732 entreprises et individus, 2014 à 2019,	
Tax Fraud Detection through Neural Networks: an Application Using a Sample of Personal	Segmentation des contributeurs et calcul du risque d'évasion fiscale	(Pérez López et al., 2019)	Données officielles (2M retours fiscaux, particuliers, 2014)	Espagne (Europe de l'Ouest) : Agence espagnole d'administration fiscale de l'État ( <i>La Agencia Estatal de Administración Tributaria</i> )

Articles	Objectif de recherche	Auteurs	Données mobilisées	Pays et AF
A Multi-Module Machine Learning Approach to Detect Tax Fraud	Développement d'un modèle de détection de fraude à partir des déclarations d'impôt	(Alsadhan, 2023)	Environ 2 000 000 de données fiscales (VAT), janvier 2018 et avril 2022	Arabie Saoudite (Moyen-Orient) : L'Autorité de la Zakat, des Impôts et des Douanes de l'Arabie Saoudite ( <i>The Zakat, Tax and Customs Authority</i> )
Simultaneous Classification and Regression for Zakat Under-Reporting Detection	Détection des sous-déclarations de Zakat par l'apprentissage automatique	(Mohamed Maher Ben et AlSadhan, 2023)	51 919 déclarations de Zakat, janvier 2018 à avril 2022	
A New Machine Learning Model for Detecting Levels of Tax Evasion Based on Hybrid Neural Network	Détection de l'évasion fiscale à partir d'un modèle hybride	(Shujaaddeen et al., 2024)	Données fiscales, pas de date, (1083 attributs)	Yémen (Moyen-Orient): Autorité fiscale du Yémen ( <i>Tax Authority of Yemen</i> )
Deep Learning-Based Detection of Tax Frauds: Property Acquisition Tax	Détection des contribuables présentant un risque élevé de fraude fiscale	(Lee, 2022)	2 228 déclarations d'impôt, 2016 à 2018	République de Corée (Asie de l'Est) : Service national des impôts de Corée
Detecting Corporate Tax Evasion Using a Hybrid Intelligent System: A Case Study of Iran	Détection de l'évasion fiscale des entreprises	(Rahimikia et al., 2017)	Données sectorielles (alimentaire (3 097 entreprises), textile (2 356 entreprises)), 2010 à 2011, taille correcte	Iran (Moyen-Orient): Administration fiscale nationale iranienne ( <i>Iranian National Tax Administration (INTA)</i> )
Identifying Tax Evasion in Mexico with Tools from Network Science	Identification des potentiels fraudeurs	(Zumaya et al., 2021)	7 571 093 factures électroniques	Mexique (Amérique latine) : Service de l'Administration fiscale du Mexique

Articles	Objectif de recherche	Auteurs	Données mobilisées	Pays et AF
and Machine Learning				( <i>Servicio de Administración Tributaria</i> )
Supervised Machine Learning for Tax Evasion Detection: A Case Study with the Brazilian Tax Administration	Prédiction du revenu annuel et omissions de déclaration	(Pires, 2024)	390 165 contribuables, 2019 (fonction publique fédérale)	Brésil (Amérique du Sud): Recettes fédérales du Brésil ( <i>Receita Federal</i> )
Tax Evasion Risk Management Using a Hybrid Unsupervised Outlier Detection Method	Développement d'un outil de gestion de risque en matière d'évasion fiscale	(Savić et al., 2022)	Données fiscales, 179 489 entités commerciales, audits réels, mars 2016 à mars 2017	Serbie (Europe de l'Est) : L'Autorité fiscale de la République serbe ( <i>The Republic of Serbia's Tax administration</i> )

## ANNEXE E

### Effets et impacts organisationnels

<b>Comportements fiscaux frauduleux</b>	<b>Performance opérationnelle</b>	<b>Gestion stratégique et du ciblage fiscal</b>	<b>Qualité du service aux contribuables</b>
Fraude fiscale des individus et des entreprises (Arabie Saoudite)	<ul style="list-style-type: none"> <li>- Amélioration de la précision lors de la détection de l'impôt sur les bénéficiaires commerciaux et industriels</li> <li>- Baisse des activités manuelles</li> <li>- Augmentation du recouvrement des impôts</li> </ul>	<ul style="list-style-type: none"> <li>- Performance très élevée du modèle avec une capacité à mieux hiérarchiser les risques</li> <li>- Maximisation de la priorisation stratégique et opérationnelle des activités de vérification</li> <li>- Meilleure gestion des ressources</li> </ul>	- Généralisation du modèle à d'autres types d'impôts
Fraude de l'impôt sur le revenu des individus et des entreprises (Rwanda)	- Réduction du temps d'évaluation des risques, de la vérification et des coûts	<ul style="list-style-type: none"> <li>- Amélioration des prises de décision stratégique pour la planification des activités de vérification</li> <li>- Amélioration des pratiques de vérification</li> </ul>	- Généralisation du modèle à d'autres types d'impôts et à d'autres administrations fiscales
Évasion fiscale de l'impôt sur le revenu des individus (Brésil)	- Capacité de prévision sur l'ensemble des tests	- Indication d'une forte performance de généralisation	- Amélioration et perfectionnement des activités de vérification
Fraude de l'impôt sur le revenu des individus (Espagne)	<ul style="list-style-type: none"> <li>- Amélioration de la détection de la fraude</li> <li>- Augmentation de l'amélioration des prises de décision relatives aux vérifications</li> </ul>	- Réduction des coûts	
Évasion de la TVA (Mexique)	- Évasion fiscale estimée : entre 135 millions et 1 milliard de pesos de 2015 à 2018 avec une	<ul style="list-style-type: none"> <li>- Meilleure planification et gestion des risques fiscaux</li> <li>- Bonne estimation des montants issus de l'évasion</li> </ul>	

<b>Comportements fiscaux frauduleux</b>	<b>Performance opérationnelle</b>	<b>Gestion stratégique et du ciblage fiscal</b>	<b>Qualité du service aux contribuables</b>
	moyenne de 259 millions de pesos par an.		
Fraude à la TVA(Rwanda)	- Amélioration de la précision dans la détection de la fraude - Augmentation du recouvrement des impôts		- Généralisation du modèle à d'autres types d'impôt
Fraude à la TVA (Rwanda)	- Augmentation de l'efficacité des vérificateurs	- Meilleure planification et gestion des risques fiscaux	- Amélioration de la transparence dans l'utilisation des machines de facturation
Fraude sur les acquisitions immobilières (République Corée)	- Amélioration et perfectionnement des activités de vérification	- Perfectionnement des stratégies de sélection des contribuables à risque	- Généralisation du modèle à d'autres types d'impôts et à d'autres administrations fiscales
Fraude à la TVA (Arabie Saoudite)	- Augmentation du recouvrement des impôts - Baisse des activités manuelles - Augmentation de l'efficacité opérationnelle	- Amélioration de la détection des cas de fraude - Amélioration de la gestion des risques fiscaux - Meilleure gestion des stratégies de vérification	
Évasion fiscale de l'impôt sur le revenu des entreprises (Iran)	- Amélioration de la détection de l'évasion fiscale	- Accroissement de l'efficacité des vérifications par la réduction des efforts manuels - Planification stratégique des vérifications	
Évasion fiscale de l'impôt sur le revenu personnel (Serbie)	- Baisse du coût des audits fiscaux	- Bonne allocation des ressources, mettant l'accent sur la gestion des risques - Optimisation des stratégies d'audits fiscaux	

<b>Comportements fiscaux frauduleux</b>	<b>Performance opérationnelle</b>	<b>Gestion stratégique et du ciblage fiscal</b>	<b>Qualité du service aux contribuables</b>
Fraude de l'impôt sur les bénéfices commerciaux et industriels (Yémen)	<ul style="list-style-type: none"> <li>- Maximisation des coûts de détection et des ressources internes</li> <li>- Efficacité dans l'analyse des données</li> </ul>	<ul style="list-style-type: none"> <li>- Amélioration de la précision lors de la détection de la fraude</li> <li>- Amélioration de la nomenclature des comportements des contribuables</li> </ul>	- Bonification de la qualité du service bonifiée

## ANNEXE F

### Comportements fiscaux frauduleux, performances et types d'apprentissage

Comportements fiscaux frauduleux	Algorithmes	Types d'apprentissage	Résultats des évaluations de performance	Effets et impacts
Fraude fiscale des individus et des entreprises (Arabie Saoudite)  (Mohamed Maher Ben et AlSadhan, 2023)	Réseau neuronal profond	Supervisé	<ul style="list-style-type: none"><li>- Exactitude : 99%</li><li>- Rappel : 99%</li><li>- Précision : 99%</li><li>- Mesure F1 : 99%</li><li>- Erreur quadratique moyenne (<i>MSE</i>) : 0.402</li><li>- Racine de l'erreur quadratique moyenne (<i>RMSE</i>) : 0.583</li><li>- Erreur absolue moyenne (<i>MAE</i>) : 26%</li></ul>	<p>Performance très élevée du modèle avec une capacité à mieux hiérarchiser les risques.</p> <ul style="list-style-type: none"><li>- Amélioration de la précision lors de la détection de l'impôt sur les bénéfices commerciaux et industriels</li><li>- Maximisation de la priorisation stratégique et opérationnelle des activités de vérification</li><li>- Baisse des activités manuelles</li><li>- Augmentation du recouvrement des impôts</li><li>- Meilleure gestion des ressources</li><li>- Généralisation du modèle à d'autres types d'impôts</li></ul>

Comportements fiscaux frauduleux	Algorithmes	Types d'apprentissage	Résultats des évaluations de performance	Effets et impacts
Fraude de l'impôt sur le revenu des individus et des entreprises (Rwanda)  (Belle Fille et al., 2022).	Réseau de neurones artificiels (RNA)	Supervisé	<ul style="list-style-type: none"> <li>- Exactitude : 92%</li> <li>- Précision : 85%</li> <li>- Rappel : 99%</li> <li>- Mesure F1 : 92%</li> <li>- <i>AUC-ROC</i> : 95%.</li> </ul>	<p>Le modèle est optimisé lorsqu'il est utilisé avec des échantillons de données hors échantillon.</p> <ul style="list-style-type: none"> <li>- Réduction du temps d'évaluation des risques, de la vérification et des coûts</li> <li>- Amélioration des prises de décision stratégique pour la planification des activités de vérification</li> <li>- Amélioration des pratiques de vérification</li> <li>- Généralisation du modèle à d'autres types d'impôts et à d'autres administrations fiscales</li> </ul>
Évasion fiscale de l'impôt sur le revenu des individus (Brésil)  (Pires, 2024))	<i>Gradient Boosting</i>	Supervisé	<ul style="list-style-type: none"> <li>- Score <math>R^2</math> : 0.84</li> <li>- <i>EAM</i> : R\$ 26,839.25</li> </ul>	<p>Le modèle a démontré une excellente capacité de prévision sur l'ensemble des tests, ce qui indique une forte performance de généralisation.</p> <ul style="list-style-type: none"> <li>- Amélioration et perfectionnement des activités de vérification</li> </ul>
Fraude de l'impôt sur le revenu des individus (Espagne)  (Pérez López et al., 2019))	Réseau neuronal perceptron multicouche ( <i>MLP</i> )	Supervisé	<ul style="list-style-type: none"> <li>- Confusion matrix : 84%</li> <li>- Aire sous la courbe ROC (<i>AUROC</i>) : 0.918</li> </ul>	<ul style="list-style-type: none"> <li>- Amélioration de la détection de la fraude</li> <li>- Augmentation de l'amélioration des prises de décision relatives aux vérifications</li> <li>- Réduction des coûts</li> </ul>

Comportements fiscaux frauduleux	Algorithmes	Types d'apprentissage	Résultats des évaluations de performance	Effets et impacts
Évasion de la TVA (Mexique)  (Zumaya et al., 2021)	Réseaux neuronaux profonds Forêt aléatoire	Supervisé	Réseaux neuronaux profonds : - Mesure F1 : 0.87 pour les entreprises émettant des opérations simulées ( <i>EFOS</i> ) définitifs et 0.94 <i>EFOS</i> présumé - 149 921 contribuables suspects Forêt aléatoire : - Aire sous la courbe ROC ( <i>AUROC</i> ) : 0.893 - Erreur : 0.157 - 128 227 contribuables suspects	Les modèles de SIA permettent une amélioration de la détection de l'évasion fiscale : - Évasion fiscale détectée d'environ 40 milliards et 346 milliards de pesos entre 2015 et 2018 - Entre 7 677 et 17 769 contribuables suspects détectés - Meilleure planification et gestion des risques fiscaux - Bonne estimation des montants issus de l'évasion
Fraude à la TVA (Rwanda)  (Murorunkwere et al., 2023)	Réseau de neurones artificiels	Supervisé	- Précision : 0.848 - Rappel : 0.867 - Mesure F1 : 0.854 - Exactitude : 0.866 - Perte logarithmique : 0.408 - Aire sous la courbe ROC ( <i>AUROC</i> ) : 0.899	Le réseau des neurones artificiels est un modèle robuste, car sa performance lui permet de détecter les comportements fiscaux frauduleux et non frauduleux. - Amélioration de la précision dans la détection de la fraude - Augmentation du recouvrement des impôts - Généralisation du modèle à d'autres types d'impôts

Comportements fiscaux frauduleux	Algorithmes	Types d'apprentissage	Résultats des évaluations de performance	Effets et impacts
Fraude à la TVA (Rwanda)  (Belle Fille, et al., 2023)	Arbres de décision, forêt aléatoire, <i>Gradient boost</i> et <i>XGBoost</i>	Supervisé	<ul style="list-style-type: none"> <li>- Forêt aléatoire :</li> <li>- Précision : 92,4 %</li> <li>- Rappel : 92.1%</li> <li>- Mesure F1 : 92.2%</li> <li>- Exactitude : 92.3%</li> <li>- Perte logarithmique :40.9%</li> <li>- Aire sous la courbe ROC (<i>AUROC</i>) : 84.79%</li> </ul>	La technique forêt aléatoire ressort comme la plus performante pour détecter la fraude à la TVA. <ul style="list-style-type: none"> <li>- Meilleure planification et gestion des risques fiscaux</li> <li>- Augmentation de l'efficacité des vérificateurs</li> <li>- Augmentation des revenus fiscaux</li> <li>- Amélioration de la transparence dans l'utilisation des machines de facturation</li> </ul>
Fraude sur les acquisitions immobilières (République de Corée)  (Lee, 2022)	<i>Autoencoder</i>	Non supervisé	<ul style="list-style-type: none"> <li>- Moyenne de l'erreur de reconstruction : 0.88</li> <li>- Précision : augmenté de 0.1% (3 fraudes sur 2 228) à 0.9% (3 fraudes sur 352)</li> </ul>	L'apprentissage non supervisé d'autoencodeur a détecté les déclarations fiscales avec des attributs inhabituels, constituant des cas avérés de fraude. <ul style="list-style-type: none"> <li>- Amélioration et perfectionnement des activités de vérification</li> <li>- Perfectionnement des stratégies de sélection des contribuables à risque</li> <li>- Généralisation du modèle à d'autres types d'impôts et à d'autres administrations fiscales</li> </ul>

Comportements fiscaux frauduleux	Algorithmes	Types d'apprentissage	Résultats des évaluations de performance	Effets et impacts
Fraude à la TVA (Arabie Saoudite) (Alsadhan, 2023)	Supervisé : <i>EXtreme Gradient Boosting (XGBoost)</i> , Réseau de neurones artificiels, <i>Support Vector Machine</i> Non supervisé : <i>Autoencoders</i>	Système hybride supervisé et non supervisé	<ul style="list-style-type: none"> <li>- Précision : 0.77</li> <li>- Rappel : 0.90</li> <li>- Exactitude : 74.78%</li> <li>- Mesure F1 : 0.83</li> </ul>	Le système utilisant les 4 modèles est efficace. <ul style="list-style-type: none"> <li>- Amélioration de la détection des cas de fraude</li> <li>- Amélioration de la gestion des risques fiscaux</li> <li>- Meilleure gestion des stratégies de vérification</li> <li>- Augmentation du recouvrement des impôts</li> <li>- Baisse des activités manuelles et augmentation de l'efficacité opérationnelle</li> </ul>
Évasion fiscale de l'impôt sur le revenu des entreprises (Iran) (Rahimikia et al., 2017)	<i>SVM (Support Vector Machine)</i> , <i>MLP (Multilayer Perceptron)</i> et Régression logistique avec optimisation de la recherche d'harmonie (HS)	Système hybride supervisé	Secteur alimentaire : <ul style="list-style-type: none"> <li>- Précision <i>MLP</i> (90.07%)</li> <li>- Précision <i>SVM</i> (87.47%)</li> </ul> Secteur textile : <ul style="list-style-type: none"> <li>- Précision <i>MLP</i> (82.45%)</li> <li>- Précision <i>SVM</i> (84.65%)</li> </ul>	MLP a fourni une meilleure performance avec l'utilisation des données hors échantillon pendant que SVM a performé avec les données de test. <ul style="list-style-type: none"> <li>- Détection avec précision par le système hybride</li> <li>- Amélioration de la détection de l'évasion fiscale</li> <li>- Accroissement de l'efficacité des vérifications par la réduction des efforts manuels</li> <li>- Planification stratégique des vérifications</li> </ul>

Comportements fiscaux frauduleux	Algorithmes	Types d'apprentissage	Résultats des évaluations de performance	Effets et impacts
<p>Évasion fiscale de l'impôt sur le revenu personnel (Serbie)</p> <p>(Atanasijević et al., 2022)</p>	<p><i>HUNOD</i> : détection hybride non supervisée des valeurs aberrantes</p>	<p>Système hybride non supervisé</p>	<p>VT : fraction des instances d'entraînement de l'autoencodeur,  - anomalies détectées par <i>K-means</i>: inférieur à 5,5% pour les entreprises avec moins de 10 employés (L10) et inférieur à 2.6% pour les entreprises avec plus de 10 employés (V10)s.</p> <p>V0 : fraction des anomalies détectées par l'autoencodeur  - anomalies détectées par <i>K-means</i>: 98.7% pour L10 et 90% pour A10</p> <p>Précision : 100% (L10 et V10)  Précision relative : 98.7% (L10) et supérieur à 90% (A10)</p>	<p>Le modèle <i>HUNOD Autoencoder</i> fournit des résultats plus précis par rapport aux autres méthodes. Il est aussi est une bonne alternative aux autres modèles en termes de précision.</p> <ul style="list-style-type: none"> <li>- Impact positif sur l'administration fiscale</li> <li>- Adresser les enjeux stratégiques digitaux pour la lutte contre l'évasion fiscale</li> <li>- Optimisation des stratégies d'audits fiscaux</li> <li>- Baisse du coût des audits fiscaux</li> <li>- Bonne allocation des ressources, mettant l'accent sur la gestion des risques</li> </ul>

<b>Comportements fiscaux frauduleux</b>	<b>Algorithmes</b>	<b>Types d'apprentissage</b>	<b>Résultats des évaluations de performance</b>	<b>Effets et impacts</b>
Fraude de l'impôt sur les bénéfices commerciaux et industriels (Yémen)  (Shujaaddeen et al. 2024)	Réseau neuronal hybride : <i>Multilayer Perceptron (MLP) et Self-Organizing Map (SOM)</i>	Système hybride supervisé et non supervisé	Modèle hybride : - Mesure F1 (93.97%) - Précision (99.07%)	Le système hybride est plus performant que les autres modèles. - Maximisation des coûts de détection et des ressources internes - Amélioration de la précision lors de la détection de l'impôt sur les bénéfices commerciaux et industriels - Amélioration de la nomenclature des comportements des contribuables - Efficacité dans l'analyse des données - Amélioration de la qualité du service bonifiée