**Fiche desc. Source 3 : IAPP – Global AI Law and Policy Tracker**

**Structure :**

La source originale est un rapport de suivi publié par l'IAPP au format PDF. Afin de permettre une analyse programmatique et la fusion avec d'autres jeux de données, les informations tabulées de ce document ont été obtenues dans un format structuré (CSV).

**Métadonnées:**

| Titre | IAPP – Global AI Law and Policy Tracker |
| --- | --- |
| Auteur | International Association of Privacy Professionals (IAPP) |
| Date de mise à jours | Last updated May 2025 |
| Juridictions traitées | 25 (ex. Canada, Chili, EU...) |
| Type | Rapport synthétique / Tableau comparatif |
| Format | PDF, 35 pages |

**Langue :** Anglais

**Source :** IAPP – Global AI Law and Policy Tracker

<https://iapp.org/media/pdf/resource_center/global_ai_law_policy_tracker.pdf>

**Volume approximatif :** 25 “données” (juridictions) : Le document contient 25 données, chacune correspondant à une juridiction analysée.

**Structure du des données:**

| **Sections (variables)** | **Description / Contenu typique** | **Difficultés ou remarques** |
| --- | --- | --- |
| Juridiction (Pays) | Nom du pays ou entité politique (ex : Canada, EU, Chine) | Aucune difficulté, clairement identifiable. |
| Specific AI governance law or policy | C'est l'information principale. Cette section décrit une ou plusieurs initiatives directement liées à l'IA :  • Projets de loi (bills)  • Lois déjà en vigueur  • Stratégies nationales  • Guides ou recommandations officielles. | • **Titre trompeur** : La section est au singulier mais contient souvent plusieurs initiatives.  • **Extraction manuelle** : Il faut lire et "décomposer" le paragraphe pour identifier chaque initiative, son titre et son statut (ex: 'draft', 'released', 'debating'). |
| Relevant authorities | Ministères, agences, comités nommés pour réguler ou superviser l’IA | - **Listes parfois longues et hétérogènes** ; noms à normaliser (ex : traduire ou harmoniser selon le pays)  - **Acronymes :** Les noms peuvent être des acronymes qui demandent une recherche rapide pour être compris. |
| Other relevant laws and policies  FR : Autres lois et politiques pertinentes | Liste les lois relatives à l’IA qui ne sont pas spécifiquement sur l'IA mais qui ont un impact direct sur son déploiement : protection des données, cybersécurité, propriété intellectuelle, etc. | - **Identifier le lien** : La difficulté est de bien comprendre pourquoi cette loi "autre" est pertinente pour l'IA.  - **Statut individuel** : Chaque loi listée ici a également son propre statut, qu'il faut noter : certaines en force, d'autres en révision -> Statut à extraire et classer |
| Wider AI context  FR : Contexte plus large en IA | Informations générales sur le développement, l’approche ou les priorités stratégiques d’un pays concernant l’IA : stratégies nationales, objectifs éthiques, priorités sectorielles, partenariats internationaux, plans d’investissement… | - **Données qualitatives** : L'information est descriptive et non-quantitative.  - **Parfois redondant avec d’autre sections ou d’autre source de données** : OECD apparaît souvent alors que c’est une source de données principal que l’on va utiliser pour nos analyses)  - **Coût en temps** : On peut faire une classification thématique des données via cette section en fonction du temps qu’on aura. Elle contient des liens vers d'autres pages web qui contiennent plus de détails (long à explorer).  - **Utile pour le rapport** : C'est une source précieuse pour la rédaction et pour donner de la profondeur à l'analyse, mais difficile à faire entrer dans des colonnes de données structurées. |

**Variables disponibles**

Depuis les différentes sections de ce tracker (PDF) et en l’analysant, on a pu découvrir et distinguer les variables disponibles pour des analyses statistiques pertinentes. Sachant qu’il y a 25 juridictions dans le PDF et qu’une juridiction peut contenir plusieurs initiatives, on considère une donnée comme étant une initiative et on obtient donc plus que 25 données, initialement perçues.

Pour plus de simplicité, on distingue ici deux types de variables, celles qu’on extrait directement et celle que l’on va créer par classification.

1. **Variables d'Extraction Directe**

| **Variable** | **Description** |
| --- | --- |
| pays | Le nom de la juridiction (ex: "Argentina", "Canada"). |
| titre\_initiative | Le nom de la loi, du plan, de la stratégie ou du guide qu’on extrait au sein du texte de la section “Specific AI governance law or policy” (ex: "National AI Plan", "Resolution 2/2023"). |
| status | L'état d'avancement déduit du texte pour cette initiative précise (ex: "In force", "In development"). |
| categorie | Catégorie de l’initiative politique déduite du texte (Governance, Financial support, Guidance and regulation etc.) |
| details | Un champ texte où vous pouvez copier le résumé de l'initiative. |
| autorites\_relevantes | Un champ dédié pour lister les noms des organismes et autorités mentionnés. |

1. **Variables de Classification**

Ces variables de classification ont été pensées et créées par le rédacteur de ce rapport (le stagiaire Layân ID YASSINE) en fonction des différentes propositions d’analyse suivantes, qu’elles nous permettent de faire.

Ce sont des variables de catégorisation des initiatives politiques à partir desquelles on déduit leurs valeurs via le NLP ou manuellement en analysant chaque politique.

C’est une liste non exhaustive mais aussi une première version des variables à créer à partir de celles présentes. En effet, il se peut qu’au cours de la partie d’extraction durant les semaines 2 et 3, on supprime certaines variables finalement pas intéressantes ou qui ne sont pas facilement exploitables en fonction de la présence.

| **Nom de la variable** | **Question à se poser en lisant le texte** | **Résultats possibles** |
| --- | --- | --- |
| posture\_strategique | Quelle est l'attitude générale du pays envers l'IA (Innovation vs Régulation) ? | Pro-Innovation; Pro-Régulation; Équilibré |
| objectif\_strategique | Quel est le but principal, l'ambition la plus haute, que le pays déclare vouloir atteindre avec sa stratégie IA ? | Devenir un leader mondial; Assurer la souveraineté; Stimuler l'économie/compétitivité; Protéger les citoyens/droits; Améliorer les services publics; Non spécifié |
| collaboration\_internationale | Le texte mentionne-t-il une participation active, une collaboration ou un alignement avec des organisations internationales ou d'autres pays ? | OCDE; UE; UNESCO; G7 / G20; Partenariat bilatéral; Aucune mention |
| modele\_inspire\_par | La politique du pays s'inspire-t-elle d'un modèle international existant ? | EU AI Act; OECD Principles; Approche distincte; Non spécifié |
| mention\_investissement\_public | Le contexte mentionne-t-il un plan d'investissement public majeur ? | Oui; Non |
| focus\_droits\_humains | La stratégie est-elle explicitement ancrée dans la protection des droits fondamentaux ? | Oui; Non |
| mention\_securite\_nat | La stratégie est-elle liée à des enjeux de défense ou de sécurité nationale ? | Oui; Non |
| mention\_impact\_emploi | Le contexte aborde-t-il l'impact de l'IA sur l'emploi ou la formation ? | Oui; Non |
| focus\_soutien\_pme | Le contexte mentionne-t-il un soutien spécifique aux PME ou start-ups ? | Oui; Non |
| mention\_partenariat\_public\_prive | Des partenariats public-privé sont-ils mentionnés comme un élément de la stratégie ? | Oui; Non |
| region | À quelle grande région géopolitique le pays appartient-il ? | Europe; Amérique du Nord; Asie; Amérique du Sud; Afrique; Océanie |
| cible\_principale | La stratégie vise-t-elle surtout à moderniser le secteur public ou à stimuler le secteur privé ? | Secteur Public; Secteur Privé; Mixte/Équilibré |
| approche\_gouvernance | L'approche générale est-elle de prévenir les risques en amont (ex-ante) ou de les corriger après coup (ex-post) ? | Ex-Ante (Préventif); Ex-Post (Réactif); Mixte |
| horizon\_temporel | La stratégie mentionne-t-elle un horizon temporel ou une date cible ? | Texte libre (ex: '2030', '5 ans') |
| themes\_prioritaires | Quels sont les thèmes prioritaires mis en avant dans le texte ? | Liste de thèmes (ex: 'Éthique; Droits humains; Compétitivité; Sécurité') |

| **Variable dérivée** | **Description** | **Méthode de création** | **Résultats possibles** |
| --- | --- | --- | --- |
| Country | Pays d’application de la loi/politique (extrait de la colonne 'Unnamed: 0') | Extraction directe (nettoyage si nécessaire) | France, Germany, Brazil, etc. (noms de pays) |
| Is\_National\_Strategy | Indique si la politique est une stratégie nationale d’IA | Regex sur 'national AI strategy' | Oui / Non |
| Law\_or\_Policy | Indique si le texte est une loi ou une politique | Regex/NLP sur 'law', 'policy', 'act' | Law / Policy / Act / Other |
| Date\_Adopted\_or\_Announced | Date de mise en place si présente | Extraction regex de dates (e.g., YYYY) | 2017, 2020, 2023, etc. |
| Num\_Other\_Related\_Laws | Nombre de lois/politiques connexes mentionnées | Comptage de ';' ou ',' dans 'Other relevant laws and policies' | 0, 1, 2, 3, etc. |
| Num\_Institutions\_Involved | Nombre d’autorités mentionnées | Comptage de ';' ou ',' dans 'Relevant authorities' | 1, 2, 3, etc. |
| Mentions\_Ethics | Présence de termes liés à l’éthique | Regex sur 'ethics', 'trust', 'fundamental rights' | Oui / Non |
| Mentions\_Industry | Présence de termes liés à l'industrie ou innovation | Regex sur 'industry', 'innovation' | Oui / Non |
| Mentions\_Education\_Or\_Skills | Présence de termes liés à l’éducation ou aux compétences | Regex sur 'education', 'skills', 'training' | Oui / Non |
| Mentions\_Human\_Rights | Mention des droits humains | Regex sur 'human rights', 'civil liberties' | Oui / Non |
| Length\_of\_Wider\_Context\_Text | Longueur du texte dans la colonne 'Wider AI context' | len() sur le texte | Nombre entier (longueur en caractères) |
| Wider\_Context\_Topic\_Category | Thème dominant du texte contextuel | NLP : clustering ou classification supervisée | Éthique, Innovation, Éducation, Droits humains, etc. |

**Potentiel Analytique**

Cette source de données est fondamentale pour cartographier l'état et la nature de la gouvernance de l'IA dans le monde. Son potentiel analytique se décline en plusieurs niveaux :

* Analyse de la maturité réglementaire : La source permet de comparer le niveau d'avancement des cadres légaux entre les pays en se basant sur le statut des politiques (ex: "In force", "In development"). On peut ainsi visualiser quels pays ou régions sont des précurseurs en matière de législation.
* Analyse Comparative des Approches de Gouvernance : En catégorisant les initiatives ("Loi principale", "Stratégie", "Guide éthique", "Plan de Financement") et les autorités impliquées, la source permet de comparer les leviers d'action privilégiés par les différents États pour encadrer ou promouvoir l'IA.
* Enrichissement et Validation des Données : Le potentiel le plus élevé de cette source réside dans sa combinaison avec d'autres bases de données. Confrontée aux données de l'OCDE ou aux rapports comme le "AI Index" de Stanford, elle permet non seulement de valider et de compléter la liste des initiatives politiques, mais surtout de croiser des données qualitatives (le type de loi) avec des indicateurs quantitatifs (les investissements, les publications) pour des analyses plus profondes.

**Difficultés éventuelles**

Malgré la richesse de cette source, il existe certaines difficultés à prendre en compte :

* **Extraction des initiatives multiples** : Une difficulté majeure réside dans l'extraction des initiatives politiques individuelles au sein d'une même juridiction. La colonne de description regroupe souvent plusieurs actions distinctes (un plan national, une résolution, un guide, un projet de loi) dans un seul paragraphe. Par conséquent, pour analyser chaque initiative séparément, un travail d'extraction manuelle ou semi-automatisée est nécessaire pour décomposer ces descriptions textuelles et obtenir un niveau de détail plus fin.
* **Interprétation et subjectivité :** La création de variables catégorielles (comme statut, categorie\_loi ou les étiquettes thématiques) nécessite une interprétation du texte. Cette "traduction" d'un texte qualitatif en une catégorie peut comporter une part de subjectivité. Pour garantir la cohérence, il sera nécessaire de définir des règles de classification précises et de s'y tenir tout au long de la collecte.
* **Hétérogénéité des Informations :** Le niveau de détail varie considérablement d'une juridiction à l'autre. Certaines entrées sont très exhaustives (comme celle de l'Argentine), tandis que d'autres sont beaucoup plus succinctes. Cela peut rendre les comparaisons directes entre pays plus complexes.