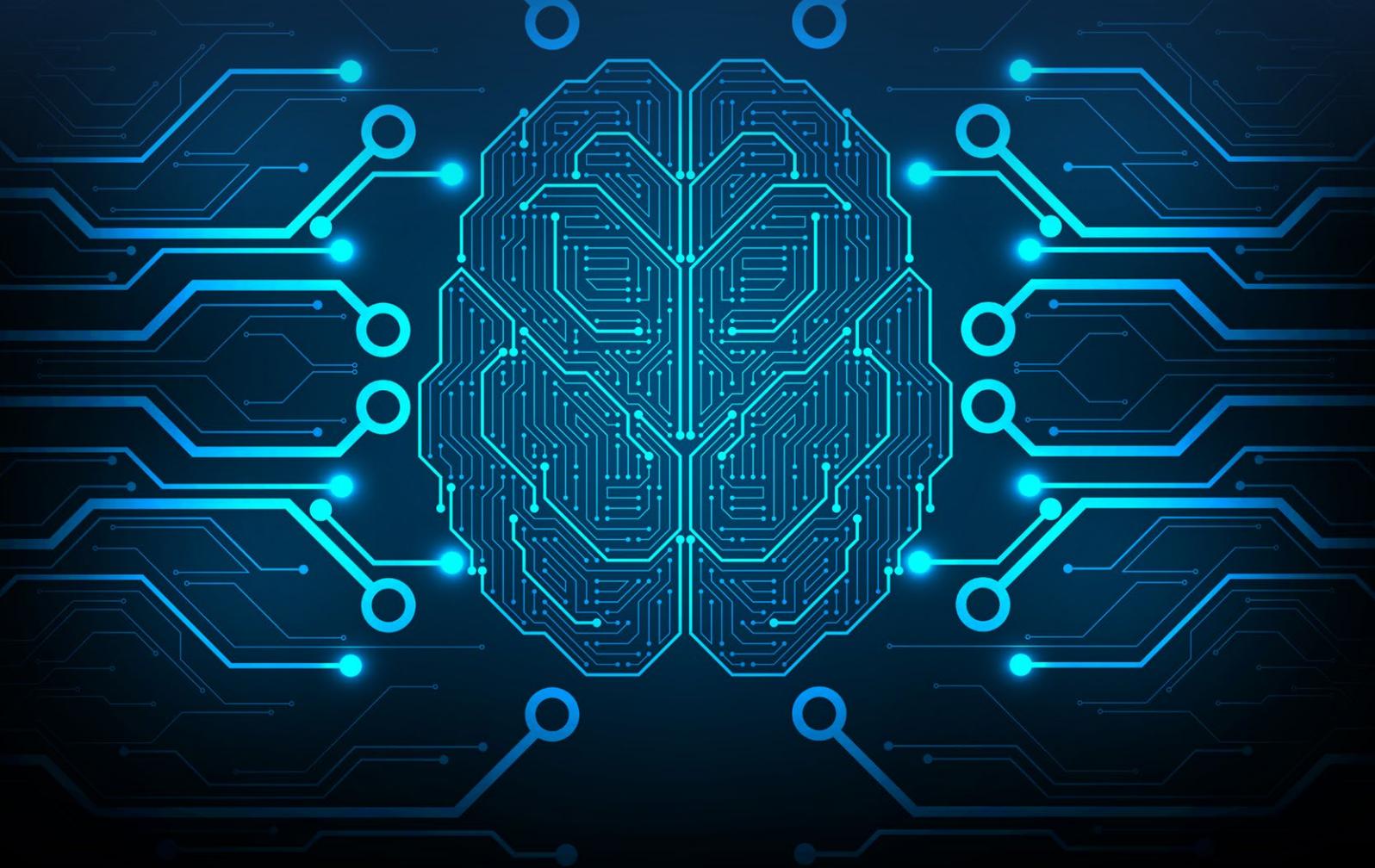


# 4

## Mégadonnées, analyse de données, intelligence artificielle et apprentissage automatique



## **L'analyse de mégadonnées, l'intelligence artificielle et l'apprentissage automatique sont utilisés sous une forme ou une autre par environ la moitié des autorités douanières, et l'autre moitié envisage d'y recourir à l'avenir. La plupart des Membres interrogés voient des avantages évidents à ce type de technologies, au premier rang desquels la gestion des risques et l'établissement d'un profil de risques, la détection des fraudes et un meilleur respect des règles.**

Une stratégie doit être établie pour les données afin d'en garantir une gouvernance et une gestion plus efficaces et de parvenir à la qualité requise. Les obstacles et difficultés liés à l'adoption de ce type de technologies sont notamment le coût et le manque de connaissances spécialisées et de bonnes pratiques.

Il existe une forte demande en ressources humaines très qualifiées sur le plan technique que l'on peine souvent à recruter, s'agissant par exemple des architectes et ingénieurs de données, et des concepteurs de logiciels et de systèmes d'apprentissage automatique. Les lois de protection des données peuvent limiter les possibilités d'exploitation des données.

Des conseils insuffisants sur l'interprétation de cette législation pour l'analyse des données à des fins douanières aboutissent souvent à une prudence excessive dans la conception des projets et limitent l'échange de données entre les organisations et les autorités douanières.

Il peut y avoir cependant des avantages majeurs pour les autorités douanières, parmi lesquels :

- l'extraction de données à des fins de renseignement et de gestion des risques ;
- les audits et les contrôles après dédouanement ;
- l'élaboration de modèles fondés sur l'intelligence artificielle pour l'interprétation des images de radiographie ;
- une meilleure efficacité s'agissant d'utiliser les données financières et fiscales plus largement et stratégiquement pour améliorer le respect des règles et les résultats en matière de perception des recettes fiscales et de lutte contre la contrebande ;
- concevoir des robots conversationnels pour répondre aux questions du public ;
- classer les droits de douane ;
- révéler les anomalies possibles dans les secteurs de fortes recettes (dont celui des marchandises assujetties au droit d'accise) ;
- détecter les anomalies de sous-évaluation et de surévaluation.

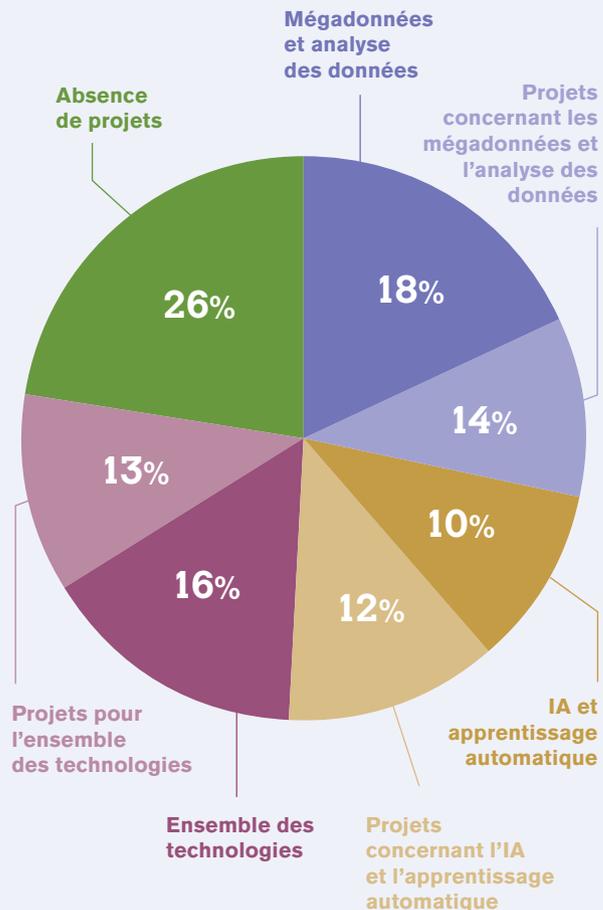
## Des conseils insuffisants sur l'interprétation des lois de protection des données limitent l'échange de données entre les organisations et les autorités douanières.

### Stade d'adoption

Il ressort de l'EAC que 44% des Membres interrogés utilisent l'analyse des mégadonnées, l'intelligence artificielle et l'apprentissage automatique (voir la figure 12), tandis que 33% prévoient de les adopter. Une minorité des Membres interrogés (23%) ne l'envisage pas à l'heure actuelle.

FIGURE 12

Stade d'adoption des mégadonnées, de l'analyse des données, de l'intelligence artificielle et de l'apprentissage automatique



Note : Au total, 94 Membres interrogés ont donné une réponse.

## Avantages

Les principaux avantages signalés concernent l'amélioration de la gestion des risques et de l'établissement d'un profil de risque, de la détection des fraudes et du respect des règles, des audits douaniers et de l'identification des anomalies (voir figure 13).

La prévision des tendances futures, l'amélioration de la facilitation des échanges et l'amélioration de la perception des recettes figurent aussi en bonne place. Les Membres interrogés ont été cependant moins nombreux à mentionner une amélioration de l'imagerie des conteneurs et des recherches visuelles.

Certains Membres interrogés ont signalé une amélioration sensible de la qualité des données lorsque plusieurs sources de données commerciales sont conjuguées aux données douanières. La Belgique, par exemple, a procédé ainsi pour améliorer la qualité des données et, partant, la qualité de l'analyse des données et des renseignements.

En intégrant des modèles d'apprentissage supervisé et non supervisé, il est possible d'améliorer les résultats de l'assertivité de la sélection des transactions, que ce soit pendant le traitement ou après celui-ci. De même, en adoptant des modèles de sélection avancés, il est possible d'interpréter les résultats plus efficacement, et de les réinjecter ensuite dans les mécanismes de sélection, ce qui améliore encore les résultats. Les modèles permettent de réduire l'arbitraire qui découle de décisions reposant seulement et exclusivement sur le jugement d'experts.

FIGURE 13

### Principaux avantages de l'adoption des mégadonnées, de l'analyse des données, de l'intelligence artificielle et de l'apprentissage automatique

Amélioration de la gestion des risques, de l'établissement du profil de risque, de la détection des fraudes et du respect des règles **67**

Facilitation des audits douaniers et de l'identification des anomalies **71**

Prévision des tendances futures **65**

Amélioration de la facilitation **57**

Amélioration de la perception des recettes **57**

Amélioration de l'imagerie (conteneurs) et des recherches **37**

*Note* : Au total, il y a eu 100 réponses. Les chiffres indiquent le nombre de réponses pour chaque catégorie (les Membres pouvaient donner plus d'une réponse).

## Obstacles à l'adoption

Les principaux obstacles à la mise en œuvre dont il est fait état sont le manque de connaissances spécialisées et les coûts (voir la figure 14). Comme dans le cas de l'adoption de l'Internet des objets, le manque de bonnes pratiques, les systèmes déjà en place, le manque d'utilisation par d'autres parties prenantes (18 Membres interrogés) et l'absence de stratégie des pouvoirs publics, ainsi que les problèmes d'ordre juridique, sont perçus comme autant d'obstacles. Les Membres interrogés citent également des problèmes liés à la gouvernance des données, à la qualité des données, à l'organisation de l'information et à la répartition des rôles et des fonctions.

---

**En adoptant des modèles de sélection avancés, il est possible d'interpréter les résultats plus efficacement.**

---

FIGURE 14

### Principaux obstacles à l'adoption des mégadonnées, de l'analyse des données, de l'intelligence artificielle et de l'apprentissage automatique



*Note :* Au total, il y a eu 94 réponses. Les chiffres indiquent le nombre de réponses pour chaque catégorie (les Membres interrogés pouvaient donner plus d'une réponse).

Un Membre interrogé ne demande actuellement qu'une partie des données dont il a effectivement besoin, tout le reste étant communiqué librement par le transporteur ou le prestataire de services. Même dans le cas de données obligatoires comme les renseignements préalables sur les passagers recueillis par les transporteurs aériens, obtenir la qualité nécessaire (données à jour et complètes) demande beaucoup d'efforts. Il peut aussi être difficile de convenir d'un champ d'application suffisamment large pour les accords de partage de données qui doivent être passés avec les fournisseurs de données tiers, y compris au niveau des pouvoirs publics.

Pour une grande organisation, utiliser correctement les mégadonnées et l'apprentissage automatique ne va pas de soi et demande souvent un recalibrage de la méthode de traitement des données selon une perspective élargie consistant entre autres à disposer d'une stratégie de données et à créer un mécanisme de gouvernance et de gestion des données. De tels mécanismes peuvent contribuer à des données de meilleure qualité et optimiser la valeur des renseignements extraits (prise de décision factuelle, automatisation, etc.).

Les Membres interrogés font aussi observer que les autorités douanières risquent de perdre une occasion unique de tirer pleinement parti de ces technologies si elles ne sont pas prêtes à gérer le changement. La formation des individus aux compétences nécessaires a posé des difficultés. La demande d'ingénieurs dûment qualifiés sur le plan technique est forte.

Ces profils sont aussi parfois difficiles à retenir dans les équipes en raison du dynamisme du marché de l'emploi

dans le domaine de l'intelligence artificielle. L'utilisation à grande échelle des données et de la technologie de l'intelligence artificielle passe par un nouvel état d'esprit dans la conception des systèmes informatiques douaniers.

Pour intégrer sur une seule plate-forme des données relatives aux transactions douanières provenant de sources multiples et examiner les mégadonnées selon une perspective à 360 degrés, des changements doivent être apportés aux systèmes informatiques actuels. Dans certains cas, il peut y avoir des obstacles juridiques qui ne permettent pas d'utiliser librement les données. Les autorités douanières sont limitées par les systèmes déjà en place et peuvent avoir encore largement recours à un traitement papier.

De même que les autorités douanières, les lois sur la protection des données sont un facteur qui peut limiter l'utilisation des données à des fins douanières. Pour ce qui est du droit de l'UE, on mentionnera le règlement général sur la protection des données (RGPD) et le code des douanes de l'Union. Selon les définitions du RGPD, presque tous les ensembles de données liés aux douanes comportent sous une forme ou une autre des données à caractère personnel, de façon directe ou indirecte, parmi lesquelles des noms, adresses, numéros de téléphone et données d'identification des entreprises que l'on pourrait faire remonter jusqu'à des personnes ou des entreprises précises. Ces champs de données sont souvent dépourvus de structure et il n'existe pas de moyen facile d'en exclure les données personnelles sans perdre l'essentiel de l'information.

Pour exploiter les données douanières, on doit opter soit pour des données anonymisées qui vont de pair avec

des contraintes minimales de protection des données (mais des données moins utiles), soit pour des données personnelles (auquel cas bon nombre des règles de l'UE et des lois nationales s'appliquent).

La plupart des États membres de l'UE disposaient déjà de lois sur la protection du droit à la vie privée qu'ils ont dû adapter pour se conformer au RGPD, voire aller au-delà de celui-ci. Néanmoins, il n'existe pas d'orientations précises sur la méthode d'interprétation de l'ensemble de ces règles dans le contexte de l'analyse des données à des fins douanières. Ce flou dans les orientations et l'interprétation se solde souvent par un excès de prudence dans la conception des projets. Elle limite aussi fortement les cas où un organisme est habilité à échanger des données avec d'autres autorités douanières. En pratique, on a souvent les situations suivantes :

- (i) Le traitement des données ne peut être effectué que par certaines catégories de personnel (dont les extracteurs de données, les analystes de risques, les gestionnaires de données).
- (ii) Les données peuvent seulement être échangées avec d'autres parties si un accord (ou un protocole spécifique) sur le traitement des données a été conclu. Les négociations peuvent prendre beaucoup de temps ou parfois même échouer. Le contenu doit être modifié et ratifié de nouveau à chaque définition d'un nouvel objectif ou chaque ajout d'une nouvelle source ou catégorie de données.

---

## **Il est possible de travailler à partir de données personnelles, mais de façon limitée et pour certains cas d'utilisation précis, et cela demande beaucoup de volonté, de temps et d'efforts.**

---

- (iii) Des procédures internes doivent être suivies pour s'assurer que tout est étayé comme il se doit par des pièces justificatives et que les données sont protégées (cryptées et pseudonymisées si des données complètes ne sont pas indispensables). La mise en place de l'infrastructure informatique dédiée à l'échange de données (serveurs) n'est possible qu'une fois le processus de confidentialité approuvé par un responsable de la protection des données.
- (iv) Le responsable de la protection des données doit valider l'ensemble de ces procédures et documents avant de pouvoir aller plus loin.
- (v) L'autorisation expresse de la commission de protection des données doit être obtenue pour ajouter des données externes comportant des données à caractère personnel à l'infrastructure informatique.

En dernière analyse, il est possible de travailler à partir de données personnelles, mais de façon limitée et pour certains cas d'utilisation précis, et cela demande beaucoup de volonté, de temps et d'efforts. C'est un processus fastidieux qui doit être répété à chaque ajout d'un nouvel élément. D'un autre côté, travailler sans données à caractère personnel n'est pas beaucoup plus simple, dans la mesure où pour exclure les données à caractère personnel, ces données doivent aussi être traitées au préalable.

Un Membre interrogé a soulevé le problème de la gestion des attentes. Les perspectives sont limitées par les problèmes juridiques et de qualité des données concernant l'analyse des mégadonnées dans le domaine douanier. Pour aller de l'avant, il est essentiel que les gestionnaires et les décideurs soient conscients de ces limites.

### Exemples d'utilisation

Les Membres interrogés ont communiqué de nombreux exemples de la façon dont ils utilisaient les mégadonnées, l'analyse de données, l'intelligence artificielle et l'apprentissage automatique pour tout un ensemble de projets et de cas. Le tableau 2 récapitule ces réponses en fonction du type de technologie.

TABLE 2

## Exemples d'utilisation des mégadonnées, de l'analyse de données, de l'intelligence artificielle et de l'apprentissage automatique

Audits et contrôles après dédouanement

---

Systèmes de ciblage automatisés

---

Erreurs de classification tarifaire et non-respect des indications tarifaires

---

Utilisation abusive des concessions (y compris les décisions et règlements de concession tarifaire, les accords de libre-échange et le masquage de l'origine)

---

Détection d'anomalies dans les secteurs à fortes recettes (y compris les marchandises assujetties au droit d'accise, c'est-à-dire l'alcool, le tabac et le pétrole).

---

Détection des anomalies en matière de dumping et de mesures compensatoires

---

Détection des anomalies de sous-évaluation et de surévaluation

Cas de non-conformité liés aux remboursements et aux ristournes

---

Détection des marchandises prohibées

---

Suivi en temps réel de la bonne exécution des services

---

Donner un éclairage rétrospectif sur l'information statistique douanière pour la planification et la prospective

---

Évaluation du degré de risque de conformité pour les activités des entreprises et le commerce

---

Identification des envois postaux et de messagerie de faible valeur pour améliorer l'évaluation des risques

---

Identifier les personnes à faible risque aux frontières

---

Mettre au point des méthodes d'analyse poussées pour les OEA

### Exemples propres aux mégadonnées et à l'analyse de données

Extraction de données à des fins de renseignement et de gestion des risques

---

Utiliser les données financières et fiscales pour améliorer le respect des règles et l'efficacité de la perception des recettes fiscales et renforcer la lutte contre la contrebande

---

Analyser des données massives provenant d'Internet pour repérer les tendances de la criminalité liées aux droits de propriété intellectuelle

---

Extraction de données sur les importations

---

Analyse graphique du contournement des règles commerciales

---

Analyse des données d'exportation et d'importation

---

Analyse des données relatives aux recommandations de prix

---

Créer un lac de données pour le regroupement de l'ensemble des données douanières en une seule source exacte

### Exemples propres à l'intelligence artificielle et à l'apprentissage automatique

Modèles fondés sur l'intelligence artificielle pour l'interprétation des images de radiologie

---

Utiliser l'intelligence artificielle et l'apprentissage automatique pour améliorer les processus de gestion des risques (boucle de rétroaction des résultats d'inspection, segmentation des clients, évaluation automatique, mise à niveau du profilage des risques)

---

Robots conversationnels pour répondre aux demandes en ligne du public et permettre à des robots physiques de répondre aux demandes de voyageurs aux points de contrôle

---

Cyberpatrouilles interplateformes

---

Utilisation de scanners à intelligence artificielle pour améliorer les capacités de contrôle et l'efficacité du dédouanement

---

Automatiser les procédures manuelles répétitives ayant trait à la protection des recettes (notamment par la recherche en ligne des prix de marché et la mise en place de l'intelligence artificielle pour aider à opérer un filtrage intelligent des résultats de recherche).

---

Utiliser l'intelligence artificielle en relation avec la classification du Système harmonisé