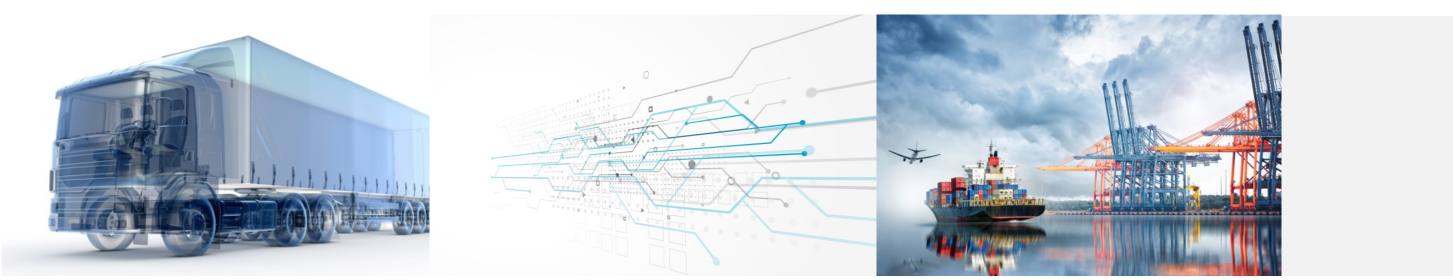
PEN-CP
Pan-European Network of Customs Practitioners

AI-Ready Customs Enforcement

Building Capabilities for the Future

Written by Toni Männistö

PEN-CP Annual Study | January 2025



Author

Toni Männistö, Cross-border Research Association (CBRA)

Credits

The author would like to express sincere thanks to the many customs experts with whom he has had the pleasure of collaborating on customs technology projects over the past five years. Their insights and contributions have been invaluable in shaping the content of this study. Special thanks go to the PEN-CP Expert Team on Detection Technologies for their exceptional input.

The author also appreciates Vladlen Tsikolenko for his work in gathering and analyzing AI literature, Professor Christiano Morini for sharing his perspectives on AI implementation in Brazil, and the late Dr. Juha Hintsa for his encouragement and inspiration in motivating the creation of this study. Finally, a big thank you goes to Mohammed Aqueel (HM Revenue & Customs), Trevor Francis (UK Border Force), and Kevin Humphreys (Revenue Commissioners – Irish Tax and Customs) for their thoughtful feedback on the final draft of the study.

Copyright

Cross-border Research Association (CBRA), 2025, www.cross-border.org

ISBN

978-2-9701478-5-5

Disclaimer

Any reference in this study to any company, product or service is for reader information only and does not constitute endorsement or recommendation by the PEN-CP consortium or the author.

This study has received funding from the European Union's Horizon 2020 research and innovation programme under the Grant Agreement No 786773.

Table of Contents

1 Foundations of AI-driven customs

Recent advancements in artificial intelligence (AI)—fueled by rapid progress in computing power, the proliferation of data, and innovative algorithms—have positioned AI as one of the most transformative technologies of the century. Amid rising global trade volumes and growing demands for trade facilitation, AI can help customs authorities maintain effective control over international goods flows, even as enforcement duties expand and inspection resources remain constrained. To fully harness its potential in customs enforcement, however, customs administrations must first build a solid foundation of AI capabilities.

1.1 Artificial intelligence as a transformative force

Artificial intelligence refers to the simulation of human intelligence in machines designed to think and learn like humans (Russell and Norvig 2021). It serves as an umbrella term encompassing a wide range of technologies, including machine learning, computer vision, and natural language processing, which—individually or in combination—bring intelligence to various applications (WCO/WTO 2022). Rather than representing a single innovation, AI acts as a transformative force, much like electricity or the internet, driving numerous applications across industries. While there is no universally accepted taxonomy for AI technologies, one intuitive way to classify them is by the human capabilities they aim to replicate and enhance.

Machine learning focuses on enabling computers to learn and adapt from data. This field is typically divided into two main approaches: supervised and unsupervised learning[1]. Supervised learning uses labeled datasets—consisting of problem-answer pairs—to train models that can predict outcomes for similar problems in the future (Chen and Chau 2005). In contrast, unsupervised learning identifies patterns and anomalies within data without relying on predefined labels, uncovering insights through statistical and mathematical reasoning. Machine learning models are trained on large datasets, and their performance is heavily influenced by both the quality and quantity of the data used for training and evaluation.

Machine vision enables machines to interpret and understand visual information from their surroundings. This branch of AI employs techniques such as image recognition, object detection, and facial recognition. Machine vision systems rely on either feature-based methods, which classify visual data by analyzing specific characteristics like edges and colors, or deep learning-based methods, which use neural networks to automatically extract and learn features from large datasets.

[1] In addition to its two core learning mechanisms, machine learning encompasses various techniques, each with its own set of algorithms. These techniques include decision trees, which use a tree-like model of decisions and their possible consequences, and neural networks, which mimic the structure and function of the human brain to recognize patterns and make predictions. Each technique is suited to different types of tasks and data, providing flexibility in how machine learning can be applied to real-world challenges.

In the realm of communication, natural language processing empowers machines to comprehend, interpret, and generate human language. The technology encompasses a diverse set of techniques for tasks such as sentiment analysis, language translation, text summarization, and speech recognition. By processing and analyzing vast amounts of textual data, natural language processing facilitates more seamless and intuitive interactions between humans and machines, enhancing the efficiency and usability of AI-driven systems.

Intelligent automation integrates technologies such as robotic process automation and machine learning to mimic and augment human capabilities, particularly in cognitive tasks, decision-making, and repetitive workflows (Williams 2021). It combines process redesign with digital agents to automate tasks traditionally performed by humans, while continuously learning from data to improve over time. By automating repetitive tasks, generating valuable insights, and supporting informed decision-making, intelligent automation serves as a key driver of digital transformation across industries and government functions.

1.2 AI opportunities in customs enforcement

Customs enforcement encompasses activities aimed at ensuring the regulatory compliance of international trade in goods. Key tasks include risk assessment, non-intrusive inspection, and the physical examination of goods and documents. Customs administrations widely recognize that AI presents a transformative opportunity to enhance productivity of enforcement activities, strengthen regulatory control at borders, and facilitate trade through more informed decision-making (WCO and WTO 2022)[2]. There are also clear use cases demonstrating how AI can support customs officers in carrying out their enforcement duties.

AI-powered risk assessment

Risk assessment is a central pillar of customs' risk-based approach to border controls. By employing risk-based, selective controls, customs can focus their limited inspection resources on high-risk traffic, facilitate low-risk trade flows, and oversee cross-border trade with minimal disruption to the movement of goods (EC 2018).[3]

Selective controls rely heavily on data and analytics, enabling customs to identify anomalies and suspicious patterns in trade documents, customs declarations, and other available information. AI technologies enhance this process by providing broader data access and more advanced analytics, leading to more effective risk assessment. For instance, Indian customs use a sophisticated machine learning approach to detect and prevent smuggling, leveraging both import documents and open-source intelligence (Ramesh and Vijayakumar 2023). This AI capability

---

[2] Customs administrations rank artificial intelligence and machine learning among the top three most promising technologies for customs operations, alongside big data analytics and blockchain (WCO 2024a).

[3] The Union Customs Code (UCC Art. 5[25]) defines risk management as "the systematic identification of risk, including through random checks, and the implementation of all measures necessary for limiting exposure to risk."

supports both the targeting of high-risk shipments and the de-risking of low-risk ones, which are critical for maintaining a balance between regulatory control and trade facilitation.

Machine learning can be harnessed to analyze large datasets and detect deviations from expected trade patterns. In maritime traffic, for instance, AI-driven anomaly detection is particularly valuable for monitoring cargo ship routes and identifying suspicious activities, such as unexplained detours or rendezvous with other vessels. These situational awareness capabilities enhance risk assessment and provide early warning systems for potential threats (EMSA 2024). Machine learning also has the potential to uncover anomalies in customs declarations, including tariff misclassification, undervaluation, and overvaluation, thereby helping customs authorities combat fiscal fraud (WTO and WCO 2022).

Natural language processing and machine learning applications have the potential to significantly streamline the process of verifying commodity codes, which is the basis for calculating duties and taxes at the border. In Singapore, an auto-categorization system has been developed that accurately classifies products into commodity codes based on the free-text descriptions provided in goods declarations (Ding et al. 2015). Similar approach could be used to extract unstructured data from diverse sources, such as e-commerce platforms, company websites, and trade documents for customs risk assessment purposes.

AI applications can also address the challenge of data linking by connecting millions of commercial and logistics data points, offering greater visibility into the international supply chain (Morgan, 2025). Machine learning and intelligent automation, for instance, are emerging as technologies that allow customs authorities automate to repetitive, rule-based activities typically carried out by humans. These technologies can streamline the integration of data records across datasets such as container bookings, bills of lading, and import declarations, facilitating the identification and targeting of high-risk shipments. As a driver of productivity, intelligent automation allows repetitive tasks to be handled automatically, enabling customs officers to focus on higher-value activities (WCO and WTO 2022).

AI-powered non-intrusive inspection

Non-intrusive inspection (NII) technologies, such as X-ray scanners, are indispensable tools in modern customs enforcement, allowing for the detection of concealed contraband with minimal disruption to cross-border trade. Machine learning models, combined with computer vision, enhance detection algorithms capable of recognizing threat objects and flagging anomalies in images generated by non-intrusive inspection equipment. AI-powered algorithms are tailored to target specific threats using specialized models; examples of off-the-shelf detection products include applications for high-density cargo, vehicles, empty containers, non-homogeneous cargo, bottles, cigarettes, and firearms (Rapiscan 2024).

Many customs agencies have developed and implemented AI-based detection systems to enhance their enforcement operations. For instance, French customs has created an AI-powered system to detect fraud and identify potential smuggling cases (Ghiran et al. 2020), while Japanese customs uses AI-driven X-ray inspection systems to detect illegal goods in postal and express parcels (Science Japan 2022).

AI-driven predictive analytics also help customs optimize the maintenance of non-intrusive inspection equipment, reducing downtime and preventing costly breakdowns (Thompson 2020). By analyzing data from scanners and inspection tools, AI can predict when maintenance is needed, drawing insights from usage patterns, error logs, and environmental conditions.

Predictive scheduling is another potential AI application for customs, which optimizes the timing and prioritization of non-intrusive inspections (WTO and WCO 2022). By leveraging predictive analytics, customs can analyze factors such as shipment volumes, peak traffic times, historical inspection data, and known risk factors to forecast periods of high demand for inspection resources. This enables better alignment of resources during peak times, helping customs streamline inspection schedules, reduce bottlenecks, and improve efficiency.

AI-powered physical examination

Physical examinations aim to verify the true nature of shipments by assessing the goods and reviewing related documentation. These examinations typically follow non-intrusive inspections, where initial scans flag potential issues, prompting customs officers to conduct manual inspections of goods and documents. The examination process includes verifying the tariff classification, value, country of origin, and quantity of the goods, as well as ensuring that no prohibited or restricted items are present in the cargo. The scope of a physical examination can range from a simple check, such as visually confirming the integrity of container seals, to a more detailed assessment involving the opening of cargo units and a thorough inspection of the contents (WCO 2010).

AI technologies are improving the examination process by enhancing both efficiency and accuracy. For instance, computer vision systems can automate the counting of pallets, packages, and other cargo items during visual inspections, significantly reducing manual effort and minimizing the risk of human error (Kafondo 2020). Additionally, machine vision, optical character recognition, and natural language processing technologies can be used effectively to identify counterfeit goods (Daoud et al. 2020).

Artificial intelligence also enhances document verification, assisting customs officers in detecting forged documents, such as invoices or financial statements, and identifying misleading information within them (Dixit 2024). Trade documents sometimes fail to accurately reflect the true contents of a shipment, but AI technologies, such as machine learning and natural language processing, can improve the accuracy and efficiency of cross-checking and fraud detection. A notable example comes from Brazil, where a machine learning model has been implemented to reduce customs officers' reliance on manual verification of customs declarations and other documents (Filho 2022).

1.3 Roadblocks to AI-ready customs enforcement

The applications of AI in customs enforcement are vast and rapidly growing, yet their implementation remains complex due to various challenges. Major barriers to AI adoption include limited political and managerial commitment, a lack of AI-savvy personnel, and difficulties in integrating AI technologies with legacy IT systems (Kafondo 2020). Moreover, successful deployment of AI requires comprehensive data strategies, strong governance frameworks, and high-quality data (WTO and WCO 2022)—prerequisites that are often lacking withing customs organizations. Effective cross-border AI operations also rely on efficient data sharing among stakeholders, which remains a critical requirement (Ghiran et al. 2020).

While these important challenges hinder the adoption of AI in customs, they are only part of the picture. Through my recent work on customs innovation projects, I have identified three additional AI capability gaps that severely limit AI implementation in enforcement operations: 1) barriers to AI analytics, 2) a reference data gap in detection technologies, and 3) limited control result data. Addressing these gaps is essential to better prepare customs for the adoption of AI technologies across various areas of enforcement.

- **Customs risk assessment**: AI has the potential to enhance data access and analytics through machine learning, intelligent automation, and natural language processing, leading to more accurate risk profiling of cross-border movements. However, the adoption of AI analytics faces several barriers, including a lack of explainability, the risk of algorithmic bias, potential counterproductive outcomes, and challenges in linking unconnected datasets.
- **Non-intrusive inspection**: Advancements in machine vision and machine learning are driving AI-powered automatic threat and anomaly detection in X-ray and other scanning technologies. Despite these advancements, progress is limited by a reference data gap—the lack of curated, high-quality datasets necessary to develop reliable detection models through machine learning.
- **Physical examination**: Manual control activities by frontline customs officers present promising opportunities for AI integration. The control result data collected during physical examinations can serve as valuable training inputs for supervised machine learning, helping to develop more accurate risk profiles for risk assessment and automatic threat and anomaly detection models for non-intrusive inspection. However, several challenges must be addressed, including the imbalanced nature of control result data, its lack of granularity, non-representativeness, and delays in its collection and integration into machine learning models.

The table below highlights the key opportunities AI offers for customs enforcement, along with the critical capability challenges that hinder its full potential.

Table 1 AI opportunities and capability challenges across customs enforcement areas

ENFORCEMENT AREA	AI OPPORTUNITY	AI CAPABILITY CHALLENGE
Risk assessment to identify suspicious shipments	**More accurate risk profiling** • Machine learning • Natural language processing • Intelligent automation	**Barriers to AI analytics** • Lack of explainability • Algorithmic biases • Counterproductive AI • Difficulty of data linking
Non-intrusive inspection to detect contraband	**Automatic threat and anomaly detection** • Machine vision • Machine learning	**Reference data gap in detection technologies** • Unavailability of quality reference data • Resource intensive curation of reference data
Physical examination to determine the identify of goods	**Feedback to supervised machine learning models** • Machine learning	**Limited control result data** • Extreme imbalance of data • Lack of data granularity • Non-representative data • Delays in data processing

1.4 About this study

This study examines three critical capability challenges that hinder the adoption of AI in customs enforcement: barriers to AI analytics, a reference data gap in detection technologies, and limited control result data. It offers a comprehensive set of technical, operational, and organizational solutions to address these gaps. Additionally, the study proposes strategies for developing in-house AI capabilities within customs organizations and emphasizes the importance of sector-wide collaboration to accelerate AI adoption.

The next three chapters explore the critical capability challenges in detail, offering practical solutions for addressing them. Chapter 2 focuses on risk assessment, discussing barriers to implementing AI-driven data analytics, such as model explainability issues, algorithmic biases, counterproductive outcomes, and difficulties in integrating diverse datasets. Chapter 3 addresses the critical shortage of high-quality reference data, which hampers the development of AI-powered threat and anomaly detection models for non-intrusive inspection technologies. Chapter 4 examines challenges related to physical examination, particularly the need for high-quality control result data to train and validate supervised machine learning models.

The final chapter provides strategic recommendations for strengthening in-house AI capabilities within customs organizations and fostering collaboration across the customs sector to ensure the successful adoption of AI technologies.

2 Risk assessment and barriers to AI analytics

Data analytics is the application of computer systems to analyze large datasets to support decision-making (WCO and WTO 2022). As a critical link between data collection and actionable insights, it enables customs authorities to interpret the vast volumes of information generated by modern international trade. Serving as a cornerstone of risk assessment systems, data analytics helps customs officers identify high-risk cross-border movements. AI technologies present a transformative opportunity for customs data analytics: AI-powered tools can process and interpret data with unmatched accuracy, identifying and predicting high-risk patterns often beyond human capability. Nonetheless, significant challenges persist, preventing customs from fully leveraging the potential of AI technologies in data analytics.

2.1 Lack of explainability

Explainability in AI refers to the extent to which a machine's reasoning is understandable to people. Many AI models function as 'black boxes', characterized by decision-making processes too complex and opaque to human users to fully comprehend. There is a trade-off between predictive accuracy and explainability: the most accurate methods, such as deep learning, are typically the least explainable, while more intuitive methods, like decision trees, tend to be less accurate (DARPA 2016).

To build trust in AI as a partner in risk assessment, customs officers need clear explanations of why the AI model makes specific decisions, such as flagging a particular shipment for inspection. Transparent reasoning allows officers to validate AI recommendations and confidently integrate them into broader enforcement strategies. Without transparency, officers may feel excluded by technology they do not fully understand, leading to reduced trust in AI assessments and resistance to its adoption. Explainability is also an ethical imperative, as it prevents unchecked reliance on AI systems that might perpetuate biases or errors.

Here are two general strategies for enhancing the explainability of AI models in customs data analytics. These approaches provide valuable, though partial, insights into machine learning models, helping to improve their transparency and trustworthiness.

Simplified explanations to AI models

Simplified explanation techniques can help clarify an AI model's predictions and bridge the gap between complex AI algorithms and practical, human-centered decision-making. Feature importance methods like SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) break down how each input parameter contributes to the model's output, making the decision-making process more transparent (Lindberg and Lee 2017). This is particularly valuable for customs targeting officers, who need to understand why a specific shipment was flagged for inspection. For example, SHAP values can reveal how factors such as the country of origin, routing, shipper, consignee, and goods description influence the overall risk score. Another valuable approach is counterfactual explanations, which illustrate how changing specific inputs could produce a different outcome. For instance, they can reveal how altering the country of origin impacts risk scores while keeping other factors of a shipment constant.

For interpreting customs X-rays, Grad-CAM (Gradient-weighted Class Activation Mapping) is a powerful technique for uncovering the most critical areas of an input X-ray that influence predictions of threat and anomaly detection models. It generates a heatmap that highlights the regions with the greatest influence on the prediction (Li et al. 2020), such as identifying areas where the model pinpoints a hidden stash of cocaine in a shipping container. This visual representation is especially valuable in the customs domain, as it helps frontline officers understand why a X-ray was flagged as suspicious.

Data visualizations

Using visualization techniques can significantly improve the explainability of AI models in a customs context. Smart data visualizations—such as heatmaps, bar charts, and graph-based analytics—allow customs officers to cluster, sort, and link data points, offering deeper insights into complex datasets. These visual tools not only help in analyzing patterns but also in identifying correlations that may not be immediately obvious. User-friendly dashboards further enhance the experience by enabling officers to explore raw data alongside AI-driven analytics, making it easier to spot emerging risks, new threat patterns, and potential indicators of smuggling or other illicit activities. When combined with simplified explanation techniques like SHAP, LIME, and Grad-CAM, these visualizations can provide even greater clarity on model predictions.

2.2 Algorithmic biases

Algorithmic biases—unjustified prejudices that favor or disadvantage specific individuals or groups (Aslett et al. 2024)— present a serious threat to the reliability of AI-based risk assessment. Just as customs officers may bring subjective biases to their decisions, AI models trained on historical data can inadvertently inherit or even amplify those biases. For instance, if shipments from certain ethnic neighborhoods have historically been flagged more frequently due to officer bias, the AI may learn to replicate this skew. This could result in unfair targeting while potentially overlooking genuine risks in areas that receive less attention. Fortunately, customs have several strategies available for mitigating the risk of algorithmic bias in AI analytics.

Human oversight

AI systems do not inherently adhere to ethical standards, and without proper human oversight, they can produce biased or undesirable outcomes. To mitigate these risks, it is essential for human operators to ensure that AI systems operate ethically and align with clearly defined objectives that support fair and proportionate customs enforcement. Key measures include systematic oversight of AI models, independent audits, randomized inspections during training phases, the establishment of robust ethical guidelines, and clearly defined roles and responsibilities within the customs organization. Addressing bias must also incorporate considerations of diversity and inclusion, allowing customs authorities to move beyond merely avoiding bias to actively fostering diversity, equity, and fairness. This approach helps ensure equitable treatment of trade flows from emerging markets and developing countries, for example. Finally, enhancing the explainability of AI models is critical for identifying and addressing potential biases, enabling more transparent and accountable decision-making (Shin 2021).

Careful curation of training data

The saying ‘a model is only as good as the data it relies on’ is especially relevant for AI analytics, which depend on vast amounts of high-quality data to perform effectively. In fact, poor-quality data can be even more harmful than no data, as it may result in biased decisions, wasted resources, and substandard enforcement performance.

One source of algorithmic bias is discriminatory training data, which teaches AI algorithms how to analyze information and draw conclusions (Ojanen 2011). In the customs context, such data often reflects historical biases stemming from past targeting priorities, the subjective views of frontline officers, or entrenched operational practices. Additionally, training datasets can suffer from sampling biases, producing non-representative data that fails to capture the full scope of trade flows. For example, these biases could lead to disproportionately high inspection rates for shipments from specific regions without valid justification. This risk underscores the importance of carefully curating training data and continuously monitoring AI systems for biased outcomes.

2.3 Counterproductive AI

The use of AI technologies in customs risk assessment brings significant benefits but can also result in unintended consequences. One key example is the application of AI tools designed to generate commodity codes from free-text descriptions of goods[4]. These tools assist customs brokers in ensuring that the declaration data they submit to customs is compliant. However, in doing so, the AI may sanitize the data provided by the original sender. This sanitization can strip away critical risk-relevant details, such as inconsistencies in the descriptions of goods, variations in declared values, or other anomalies that might have signaled potential non-compliance or fraud. As a result, these tools—despite their efficiency—can inadvertently reduce the value of information available for customs risk assessment.

Another significant concern arises from over-reliance on AI in customs enforcement. When customs officers lean too heavily on AI-generated recommendations, it risks undermining their professional judgment and intrinsic motivation. AI systems often lack the ability to provide sufficient context or clear explanations for their outputs, which can leave officers in the dark about the rationale behind certain decisions. In the worst-case scenario, this may lead to a ‘rubber-stamping’ approach, where officers blindly follow AI-generated instructions. Such an approach can

[4] A commodity code is a sequence of digits, with the first six—known as HS codes—being globally standardized, while the remaining digits are defined by national or regional classification systems. In contrast, a goods description is a free-text account intended to clearly and precisely identify the goods. However, these descriptions are often vague, inaccurate, or missing. Both commodity codes and goods descriptions are vital for customs risk assessments, aiding in identifying potential risks and ensuring compliance.

result in critical errors, such as underestimating the need for inspections or overlooking contextual risk indicators that the AI may not be programmed to consider. Furthermore, when AI systems are lack transparency, officers may find it challenging to question or refine AI's recommendations, further compounding the risks of misjudgment. In these ways, the uncritical adoption of AI in customs risk assessment may accidentally weaken enforcement processes rather than strengthen them. Below are two effective strategies to address the challenge counterproductive AI.

Dual-visibility to both original and AI-improved declaration data

Granting customs officers access to both the original sender-provided data and the AI-generated sanitized data can significantly enhance their ability to assess risk-relevant information. This dual-visibility approach enables officers to cross-check the unaltered data against the AI-processed results, ensuring that critical details are not sanitized or lost during the data conversion process. By doing so, officers can identify discrepancies, anomalies, or omissions that the AI may have overlooked or inadvertently removed, thereby improving the reliability of risk assessment.

In the postal domain, for example, many senders lack familiarity with the precise commodity codes required for customs declarations. To address this, some postal operators rely on AI tools to convert sender-provided goods descriptions into accurate commodity codes that comply with customs requirements. While this process is efficient, it risks obscuring subtle discrepancies or errors in the original sender data. In such scenarios, providing customs officers with both the unprocessed sender data and the AI-generated declarations allows them to better analyze potential risks, such as mismatches in product descriptions or values.

Human-in-the-loop

Implementing a 'human-in-the-loop' approach is a powerful safeguard to ensure that customs officers actively review and validate AI-generated recommendations. This approach prevents over-reliance on AI systems and preserves the critical role of human judgment, especially in complex or high-stakes situations. For instance, in scenarios where customs officers must decide whether to intercept a cargo ship suspected of violating international sanctions, relying solely on AI could lead to serious errors. AI systems, while efficient, may lack the ability to consider intricate contextual factors, such as geopolitical tensions, the credibility of intelligence sources, or unusual but legitimate trade patterns. By involving customs officers in the decision-making process, the human-in-the-loop model ensures that these contextual subtleties are carefully assessed. This collaborative AI-officer approach enables officers to challenge or refine AI outputs, ensuring that critical decisions are not made in a vacuum. Moreover, it provides a feedback loop where officers can identify potential flaws or biases in the AI system, which can be addressed to improve future performance.

2.4 Difficulty of data linking

Better access to risk-relevant information can enhance the quality and quantity of data used in machine learning for AI-powered cargo targeting models. A larger, more diverse data pool enables customs data scientists to develop accurate, predictive risk assessment models that better adapt to evolving trade patterns and trafficking strategies. Access to information beyond declaration data allows customs analysts to build a clearer picture of trade activities and cross-check crucial details such as company addresses, contact information, and business financials. However, broader data access requires linking diverse datasets, a complex task due to the fragmentation of information across various stakeholders, IT systems, and documents. This fragmentation creates a landscape of data sources, each with unique accessibility and integration challenges.

Data integration involves connecting data records from different datasets to create a unified view of shipments, containers, traders, or other points of interest. Unique identifiers, such as tracking numbers, container seal numbers, transaction IDs, and voyage numbers, can sometimes provide clear links between corresponding data records. However, many datasets lack these identifiers, making it difficult to establish direct connections. Probabilistic linking techniques, as well as the promotion of standardization and the adoption of unique identifiers, are essential strategies for improving data linking and enhancing the effectiveness of customs risk assessments.

Probabilistic linking of data records across datasets

When two datasets lack common unique identifiers, customs data analysts can employ probabilistic linking to identify corresponding data records. For instance, if two records share the same date, list the same shipper and consignee, and report identical goods values, they are likely associated with the same shipment. By combining multiple data points, these linking keys serve as proxies for unique identifiers. This probabilistic approach offers a high level of confidence that records from separate datasets correspond to the same entity, such as a shipment, container, or trader.

Probabilistic linking of data records tends to be a time-consuming and error-prone process. However, robust statistical models can significantly reduce linking errors by improving the accuracy of record matching. Additionally, intelligent automation offers a powerful AI-based solution by automating the data search and integration process across datasets. In principle, a digital workforce of software robots can efficiently handle repetitive tasks, accelerate data linking, and ensure consistent, error-free results, ultimately enhancing the speed and reliability of data integration efforts (Aguirre and Rodriguez 2017).

Standard unique identifiers for data linking

The most effective approach to addressing the challenge of data linking is to establish common identifiers that appear consistently across different datasets. Using unique identifiers streamlines the linking and integration of diverse data sources, which is crucial for enhancing data accessibility and enabling the full potential of AI analytics in customs risk assessment.

There are standard unique identifiers tailored to specific sub-domains of international trade. For instance, the Universal Postal Union's (UPU) S10 barcode standard facilitates the linking of information related to international postal items. In the maritime domain, container numbers maintained by the Bureau International des Containers (BIC) under the ISO 6346 standard provide a consistent way to link documents such as bills of lading and customs declarations. Nevertheless, at the global supply chain level, there is a lack of linking keys to connect datasets across various sub-domains. In the future, the use of standard identifiers should extend beyond isolated functions like customs, logistics, and commerce to facilitate cross-domain data integration. This broader standardization is essential for seamless data integration and enriching datasets for machine learning applications.

3 Non-intrusive inspection and reference data gap in detection technologies

Customs detection technologies encompass a broad range of solutions[5] that enable cargo, container, and vehicle inspections without the need for manual examination. Machine vision and machine learning AI technologies have the potential to significantly enhance these systems, boosting their performance. However, a key challenge in implementing such technologies is the limited availability of high-quality reference data, such as images and spectral data. This reference data gap arises from two primary factors: the general scarcity of available reference data and the time-intensive process of curating reference datasets for machine learning purposes.

3.1 Unavailability of quality reference data

A key rule for training AI-powered detection algorithms is that larger reference datasets generally lead to better detection performance. In machine learning, quantity often translates to quality, as training and validating high-performing AI models typically require tens of thousands of reference data points (Vukadinovic and Anderson 2022).

The size of the dataset is not the only factor that determines its quality; diversity is equally important. A good reference dataset must represent the variety of cargo types, packaging methods, techniques, and potential threats that could occur in the real world. The dataset must also be balanced, ensuring there are enough examples for machine learning to learn how true positives—targets such as firearms or contraband cigarettes—appear in the reference data.

Other key qualities of reference data include high resolution, meaning the data contains sufficient detail. For instance, a high-resolution spectral library for illicit drugs would provide precise information about the chemical composition, including active ingredients, additives, common impurities, molecular structures, and atomic proportions. Similarly, a high-resolution X-ray image database would include detailed images captured from multiple angles, utilizing material discrimination, high-intensity, and other functional modes. Additionally, this reference data should have high real-world fidelity, accurately reflecting the conditions under which the data is captured, including factors like weather conditions, signal distortions, and occlusions in X-ray images caused by dense objects. The table below summarizes the key attributes of quality reference data, followed by sections outlining solutions for improving reference data quality.

[5] X-ray and gamma-ray systems use electromagnetic radiation to create images from the inside that reveal shapes, densities, and material properties. Dual-view X-rays capture images from two angles, while dual-energy X-rays emit radiation bursts at different intensities to differentiate between organic and inorganic materials. Backscatter X-rays detect both transmitted and reflected rays, offering additional insights. Computed tomography combines cross-sectional scans to generate 3D images. Some technologies also use spectral analysis to identify substances, such as X-ray diffraction, which reveals a material's atomic arrangement.

Table 2 Key attributes of reference data from the data quality viewpoint

ATTRIBUTE OF REFERENCE DATA	
Size	Sufficient data to train and evaluate algorithms for various real-world conditions and rare or emerging threats
Diversity	Diversity of background cargo, cargo units, packaging methods, concealment techniques, and threat manifestations (e.g., angles, depths, and forms such as powders, liquids, or mixed solutions)
Balance	Sufficient representation of threat objects in the data, including rare examples
High resolution	High-resolution images and detailed spectral data
Real-world fidelity	Realistic conditions such as noise, occlusions, and distortions to reflect real-world scenarios

Sharing of reference data among customs

Sharing is an effective strategy for expanding the size and diversity of reference data. X-ray image sharing, for example, has long been a goal for EU customs due to its value in machine learning applications and operator training. However, practical implementation has been challenging. The Baltic X-ray Images Exchange (BAXE) has made progress, but it is currently limited to Estonian, Latvian, and Lithuanian customs. At the EU level, the Customs Risk Management System (CRMS) has been facilitating the exchange of risk-related information across EU customs since 2005. Its second iteration, launched in 2022, further enhances cooperation among EU customs administrations and introduces a centralized database for risk and control information (EC 2024). It also includes a ‘detection corner’ section, which enables customs officers to share X-ray images (ibid.). Despite these initiatives, there remains a strong need for a broader and more systematic exchange of reference data among customs administrations and their strategic industry partners who are developing AI-driven detection models.

Complementing real-world X-rays with synthetic or staged images

A typical customs administration generates thousands of real-world X-ray images daily as a by-product of routine border control operations. These images offer a realistic representation of cargo types, packaging, threats, concealment methods, and other contextual details encountered by customs officers. However, creating a high-quality dataset from these images is challenging due to factors such as the non-random targeting of controls, imbalanced data with limited instances of specific threats, and the fact that some illegal items pass through controls undetected.

Complementing real-world X-ray images with synthetic or staged images can improve the quality of X-ray reference libraries. Synthetic X-ray images are created by digitally overlaying threat objects onto actual images of cargo, containers, and vehicles, or by using entirely synthetic 3D models. This process simulates threats in a realistic environment, improving training and testing scenarios. Synthetic images also avoid privacy concerns, as they contain no authentic data, and enable the simulation of rare threats, such as explosives or high-caliber weapons, in realistic smuggling scenarios.

Staged images provide another method for creating X-ray datasets for training and validating machine learning models. This approach involves crafting test kits of parcels, containers, or cargo units that simulate various cargoes, targets, and concealment techniques. These kits are X-rayed from multiple angles to generate staged images. The downside of the technique is that staged images lack the unpredictability of real traffic and do not fully capture evolving smuggling techniques. Additionally, creating realistic test kits can be difficult due to restrictions on access to threat materials like explosives or drugs. The table below summarizes the strengths and weaknesses of real-world, synthetic, and staged images.

Table 3 Techniques for generating reference X-ray images for machine learning applications

IMAGE TYPE	PROS	CONS
Real-world images from actual customs operations	• Reflect real traffic and scenarios • Capture actual diversity of cargo types, threats and concealment methods • Generated as a byproduct of routine inspections	• Risk-based targeting introduces representation bias • Highly imbalanced dataset with few actual threats • Undetected threats (false negatives) compromise the data
Synthetic images from computer-generated images that simulate threats embedded in realistic smuggling scenarios	• Can generate large and diverse datasets • No handling of restricted or dangerous materials needed • Customizable, enabling models to train on rare scenarios • No privacy concerns	• Lack of realism, as synthetic images may not fully capture nuances of real-world images • Models trained solely on synthetic images may perform poorly in real-world applications
Staged images using physical test kits that emulate real-world smuggling scenarios	• Can generate large and diverse datasets • Customizable, enabling models to train on rare scenarios • Useful for testing models in controlled settings • No privacy concerns	• Lack of realism, as staged images may not fully capture nuances of real-world images • Handling of restricted or dangerous materials needed (e.g., drugs or explosives)

Unlocking proprietary reference data for broader use

One way to accelerate the development of AI-powered detection algorithms is by making existing proprietary reference libraries accessible to a broader group of AI developers. Currently, most reference data is owned and controlled by private or government entities. Private companies protect their libraries primarily for competitive reasons, while government agencies may restrict access due to privacy and law enforcement concerns.

Opening up these proprietary libraries would allow for the pooling of reference data into larger, more representative datasets, helping to address the chronic shortage of real-world datasets that includes actual targets—the true positives. Broader sharing would also benefit small and medium-sized research and commercial organizations by giving them access to valuable data, fostering innovation, and leveling the playing field with larger corporations. There is no need to make these libraries fully public; instead, access could be granted through data processing agreements or reasonable licensing fees.

Federated learning—bringing model to data to alleviate privacy concerns

A practical way to address privacy concerns is to deploy machine learning models directly to where the training data resides, rather than transferring data to centralized systems. One approach is federated learning, a decentralized machine learning technique where models are trained locally on servers that house the data. Instead of sharing sensitive data, only the model updates—such as learned patterns or weights—are transmitted to a central server. These updates are aggregated to improve the global model, ensuring that individual data remains securely on its original system. Federated learning is particularly useful for organizations like customs, where sensitive data, such as trade transactions, shipping records, or personal details, cannot be shared freely due to privacy regulations or data protection concerns.

Standard file formats for reference data

Standardized file formats are crucial for advancing AI-driven detection algorithms, which depend on large, representative reference data libraries. Interoperable formats facilitate the creation of high-quality datasets essential for training and validating machine learning-based detection models. While several standard file formats are available, their adoption and suitability for machine learning applications can be enhanced, particularly in areas such as data fields for annotations and labeling. A notable example of AI-focused innovation is the Unified File Format (UFF), a standardized format for non-intrusive inspection equipment, widely adopted by major X-ray manufacturers. The latest 2024 release includes features designed with future machine learning applications in mind. Other customs-relevant standards that could benefit from AI-centric updates include DICOS (Digital Imaging and Communications in Security), gaining traction in aviation security, and JCAMP-DX (Joint Committee on Atomic and Molecular Physical Data), a widely used format for storing Raman spectroscopic data along with relevant metadata.

Standardized file formats are essential for advancing AI-driven detection algorithms, which rely on large and representative reference data libraries. Interoperable formats enable the development of high-quality datasets that are critical for training and validating machine learning-based detection models. While several standard file formats exist, their adoption and suitability for machine learning solutions could be improved, particularly in areas like data fields for annotations and labeling. A notable example of AI-focused development is the Unified File Format (UFF), a standardized format for non-intrusive inspection equipment, widely adopted by major X-ray manufacturers. The latest 2024 release includes features designed with future machine learning applications in mind[6]. Other customs-relevant standards that could benefit from AI-focused updates include DICOS (Digital Imaging and Communications in Security)[7], which is gaining traction in aviation security, and JCAMP-DX (Joint Committee on Atomic and Molecular Physical Data), a widely used standard for storing Raman spectroscopic data along with relevant metadata.

3.2 Resource-intensive curation of reference data

Data curation is the process of preparing reference data for machine learning applications by adding annotations to highlight key features, incorporating metadata for context, and potentially anonymizing the data. This process can be time-consuming and resource-intensive.

Data annotation involves labeling X-ray images to make them suitable for machine learning algorithms, a process that demands significant time and expertise from specialized analysts. Interpreting customs X-rays is particularly complex, especially in the case of international parcel traffic. Customs officers must inspect parcels for a broad array of threats, including drugs, weapons, and counterfeit goods. These illicit items vary widely in size, shape, and configuration, often hidden within a range of background cargo. Their different orientations, depths, and levels of complexity further complicate the annotation process, requiring careful attention to detail.

Adding context to reference data through metadata is another resource-intensive task. For X-ray images, metadata can include details such as scan ID, date and time, scanner model, and resolution. It may also encompass information like container or package ID, declared contents, weight, volume, origin, destination, and carrier details. Further context could involve the scanning location, purpose of the scan, and any associated risk flags. Additionally, inspection outcomes, including detected anomalies or subsequent actions, are often recorded within the metadata, providing a more comprehensive picture of the scanned shipment.

Customs agencies generate thousands of X-ray images daily, but sharing these images with other customs. or research partners is often restricted due to privacy concerns. X-ray images and their metadata may contain personal data—such as names, phone numbers, and addresses—that must be removed before sharing. The process of blurring sensitive areas in the images or removing

---

[6] The UFF 3.0.0, developed by the Technical Experts Group on Non-Intrusive Inspection under the World Customs Organization with support from leading technology companies and customs experts, states: "The assessment type Label has been included for future expansion to support the labeling of data for machine learning purposes. This feature will be addressed in a future minor update to the standard"(WCO 2024b, p. 34).

[7] DICOS is a flexible and interoperable data standard designed to integrate various security screening imaging technologies, their images, and related metadata. It enables seamless transmission, storage, and viewing of images and associated data from any X-ray system across DICOS-compliant devices and networks (Battelle 2017).

personal data from the metadata can be time-consuming, which reduces the incentive to share X-ray data. The table below summarizes the key attributes of reference data from a data curation perspective, followed by strategies to enhance data curation within the customs context.

Table 4 Key attributes of reference data from a data curation perspective

ATTRIBUTE OF REFERENCE DATA	
Annotations	Precise labeling of threats, benign materials, and other objects of interest
Metadata	Detailed metadata including item descriptions, packaging methods, classifications, and shipping context
Privacy	Anonymization of personal data

Clear annotation guidelines for customs X-ray images

Establishing clear annotation guidelines is essential for creating consistent datasets of customs X-ray images that are suitable for machine learning. These guidelines should define standard categories for threats (e.g., weapons or drugs) and benign goods (e.g., electronics or clothing) to ensure uniform labeling across all images. They should include clear instructions and visual examples to help annotators maintain consistency. Additionally, the guidelines should address quality assurance in the annotation process. For instance, they might recommend using multiple annotators to independently label the same image, followed by a comparison of results to reach a consensus on the correct interpretation.

Funding for X-ray image curation work

Annotating X-ray datasets is a highly labor-intensive process. For example, labeling the 328,000 images in the Microsoft Common Objects in Context dataset requires around 30,000 person-hours, averaging just eleven images per hour (Liang 2020). This estimate covers basic object identification and does not include more intricate tasks, such as drawing bounding boxes around objects to highlight areas of interest. Adding detailed annotations and metadata could further reduce the annotation rate to about five images per hour (ibid.). Assuming an experienced customs X-ray analyst charges €30 per hour and two annotators are required per image for accuracy, the cost of annotating each X-ray image would be approximately €11.

Customs administrations often hold extensive archives of unannotated X-ray images, which represent an untapped resource for training AI-driven threat detection models. However, realizing the full value of these datasets requires significant investment in large-scale image annotation. By investing in annotation now, customs administrations can establish a foundation for more accurate and efficient threat detection algorithms, offering substantial long-term benefits.

Software-operator hybrid models for X-ray image annotation

Software-operator hybrid models combine human expertise with AI-assisted tools to enhance the efficiency and accuracy of annotating X-ray images. These models use software algorithms to perform initial tasks, such as object detection or segmentation, reducing the workload for human annotators. They then refine and validate the AI-assisted annotations, ensuring high-quality outputs while addressing the limitations of automated systems. This collaborative approach significantly speeds up the annotation process and reduces costs, making it particularly valuable for large-scale datasets. Hybrid models also enable iterative learning, where operator corrections improve the software's performance over time.

4 Physical examination and limited control result data

Control result data refers to the outcomes of physical inspections aimed at determining whether international goods comply with relevant laws and regulations. Comprehensive control result data includes key details such as the types of illegal goods discovered, concealment methods, violations (e.g., underreporting value, quantity discrepancies, or incorrect commodity codes), inspection techniques used, and shipment details like declaration data. It also captures the final verdicts of customs inspections, such as the release of compliant goods or the confiscation of illegal items.

Control result data is important for customs enforcement as it provides vital feedback to risk assessment systems on emerging smuggling patterns, new concealment techniques, and other trends. True positives—the successful detection of smuggling activities—offer valuable insights into the factors most closely linked to non-compliant shipments. This information allows customs data analysts to refine risk indicators and profiles, improving the accuracy of targeting over time. Importantly, control result data plays a key role in AI adoption, serving as the essential training and validation input for machine learning models. However, limitations in the quality and availability of control result data hinder its full potential in AI-driven customs enforcement solutions.

4.1 Extreme imbalance of data

Control results—the outcomes of customs inspections—often feature class-imbalanced datasets, where legal goods typically far outnumber illegal ones. This imbalance becomes even more pronounced when focusing on rare targets, such as explosives or high-caliber firearms.

Let's assume that a customs administration receives one million Entry Summary Declarations per month, and that 1%—or 10,000—of these declarations are selected for control upon arrival[8]. Since pre-arrival risk assessment primarily targets security and safety risks[9], which are fortunately rare, we can further assume that only 0.1% of these controls uncover something dangerous. This results in 10 true positives in total, representing one in 100,000 of the original declarations. The rarity of true positives (dangerous shipments) in the example makes it a textbook case of a heavily imbalanced dataset: the distribution of classes is extremely skewed, meaning that the minority class (dangerous shipments) is much less frequent than the majority class (non-dangerous shipments)[10]. This imbalance makes it difficult for machine learning models to learn patterns effectively because they are dominated by the majority class. Fortunately, customs can implement various strategies to alleviate the issue of class imbalance.

---

[8] The entry summary declaration enables pre-loading risk assessments aimed at detecting immediate security and safety threats before goods are loaded onto an EU-bound mode of transport. In the aviation sector, for example, security threats include incendiary and explosive devices—'a bomb in a box'—that have the potential to bring down an airplane. In addition, security and safety concerns at the pre-loading stage may also include risks such as contagious diseases (EC 2012) as well as radiological and nuclear hazards (Bakshi et al. 2011).

[9] As per Union Customs Code Art. 128.

[10] According to Leevy et al. (2018), severely imbalanced data, or high class imbalance, is typically characterized by a majority-to-minority class ratio ranging from 100:1 to 10,000:1.

Resampling and class weighting techniques

To address the effects of class imbalance, customs data analysts can create one or more modified datasets from the original, each featuring an adjusted class distribution. Two common strategies for this are oversampling and undersampling.

Oversampling increases the representation of the minority class (e.g., dangerous shipments), and when additional instances are added randomly, the method is referred to as random oversampling. In contrast, undersampling reduces the size of the majority class (e.g., non-dangerous shipments) by discarding a portion of its instances (Batista et al. 2004). An example of a more advanced oversampling technique is the Synthetic Minority Over-sampling Technique (SMOTE) that generates synthetic instances for the minority class by interpolating between closely located minority samples. This approach provides a more nuanced and effective way to balance the dataset (Chawla et al. 2002).

In addition to resampling methods, data scientists can adjust class weights to address imbalance. This involves modifying the learning process to impose a greater penalty for misclassifying instances of the minority class. By assigning higher weights to the minority class, algorithms such as decision trees, random forests, and neural networks can better account for the imbalance and improve predictive performance of cargo targeting models.

Data pooling to gain deeper insights into minority class instances

Collecting additional control result data may not solve the class imbalance issue, but it can enhance the understanding of minority instances. By incorporating a wider variety of detailed information—such as data on real-world security and safety threats— customs data analysts can gain deeper insights into the patterns and characteristics of these rare events. This enriched dataset not only aids in resampling efforts but also ensures that targeting models are based on more realistic data, improving their ability to detect subtle indicators of minority classes.

There are several strategies for expanding the pool of control result data, each with its specific strengths and disadvantages. One approach is to look back in time. If proper archives are available, customs authorities can extend the historical dataset used to train and evaluate risk assessment models. However, this retrospective approach requires caution, as data from different periods may have inconsistent formats or may no longer align with current traffic patterns and risk priorities.

Another approach is to expand the range of threat categories used in model training. For example, data analysts might group alcohol and cigarette smuggling under the broader category of fiscal fraud, as these activities often share similar patterns. By combining control result data from both types of seizures, a more robust and versatile model could be developed to detect such fraudulent activities. It should be noted, however, that overly broad aggregation of target classes can undermine the model's specificity and performance.

The third approach to pooling control result data is geographic expansion. Expanding the data across different regions can improve the model's ability to detect broader patterns and enhance risk assessment. However, pooling data from diverse areas may introduce inconsistencies or data quality issues due to variations in local practices, regulations, or reporting standards. Additionally, broadening the data pool across regions could result in the loss of important regional specifics, potentially affecting the accuracy and effectiveness of the models.

Choosing the most suitable machine learning approach

Machine learning models can be enhanced using techniques like resampling, class weighting, and data pooling, but these methods have their limitations. Ultimately, AI algorithms depend on real-world examples of true positives, which are scarce in the control result datasets that customs collect. For this reason, customs data analysts should address the class imbalance problem from the outset when designing and implementing machine learning models, including selecting the appropriate learning approaches and techniques to mitigate the problem.

Supervised machine learning models can be effective for tasks like detecting fiscal contraband or counterfeit goods, which occur relatively frequently in international shipments. In such cases, the problem of class imbalance is more manageable, as there can be sufficient labeled data to train and evaluate risk assessment models effectively. However, for rare security threats, supervised learning is often impractical. Customs data analysts can instead frame the problem as anomaly detection, focusing on identifying outliers in the data. Models like random forests or autoencoders, which are better suited for imbalanced datasets, may offer more effective solutions.

4.2 Lack of data granularity

A key issue with control result data is that, even when available, it often lacks the necessary details for developing effective AI-powered targeting models. In some cases, control feedback is provided in a simple binary format—indicating only whether non-compliant goods were found during physical inspections—without offering deeper insights into the nature or context of the findings. This is problematic as rich control result data is essential for AI-driven solutions because machine learning can uncover insights from large datasets that may not be immediately apparent to human analysts. For this reason, the more detailed the control data, the more effective the AI-driven risk assessment solutions.

Expanding data fields in control result forms

Control result forms should be designed to collect diverse and detailed information. They should capture specific details about intercepted goods, such as the type of contraband (e.g., drugs, weapons, or counterfeit goods) and the quantity involved. For drug seizures, the feedback forms should include detailed information like the drug type (e.g., cocaine or heroin), its form (e.g., powder or pills), and attributes such as purity or packaging method. Moreover, capturing contextual data about shipments—such as concealment techniques, inspection methods used, and time and place of control—adds significant value for targeting models based on machine learning. Additionally, a dedicated field should be included in the feedback form for customs officers to record their observations and remarks in a free-text format, enabling them to provide detailed contextual information about the circumstances surrounding a control with greater flexibility.

Convenient reporting of control results by frontline officers

Control result forms should not only capture as much risk-relevant data as possible but also minimize the time and cognitive effort required from customs officers who complete them. Achieving this balance requires forms that meet both machine learning needs and practical usability standards. A smart, digital solution would enable officers to access feedback forms on smartphones, tablets, or other handheld devices they use in their daily tasks. These devices could streamline data entry by enabling officers to scan barcodes, automatically populating shipment identifiers and declaration details, thus reducing the need for manual input.

Moreover, the interface for reporting control feedback should prioritize ease of use, incorporating features like intuitive drop-down menus for quick option selection, speech recognition for dictating notes, and auto-correction to enhance the clarity of free-text entries. Context-aware prompts could further simplify the process by guiding customs officers with follow-up questions based on their initial responses. For instance, if a cocaine seizure is reported, the system might prompt additional details about the drug's form or packaging.

4.3 Non-representative data

A non-representative dataset is one that fails to accurately reflect the characteristics, diversity, or distribution of the phenomenon it is meant to represent. Using such data for analysis can lead to biased or inaccurate conclusions because it does not align with reality. In machine learning, training models on non-representative data can cause the model to learn patterns that are too specific to the training set, reducing its ability generalize to new or unseen data. This often results in overfitting, where the model performs well on the training data but struggles with real-world scenarios.

In the case of customs control result data, common causes of non-representativeness include non-random sampling, excessively small samples, and data collection errors[11]. Non-random sampling is a fundamental issue because most control result data is not derived from a random selection process. Instead, it is shaped by automated targeting systems and field officers' judgments, which determines whether a shipment is deemed risky enough to be selected for inspection. Although this risk-based approach is effective for customs enforcement in general, it skews the control result data by overrepresenting high-risk cross-border movements.

[11] There are numerous sources of biases and systematic errors in control result data, each of which warrants a separate research paper to fully explore and analyze in detail.

Small sample sizes are another cause of non-representative control result data[12]. The European Court of Auditors (2021) reports significant differences in customs control rates across the EU Member States, ranging from under 1% of import declarations in some countries to over 60% in others. However, it appears that the typical control rate is generally closer to the lower end.

A third major cause of non-representative control data is incomplete or inaccurate reporting by customs officers. This can happen for various reasons, such as forgetting to log results, becoming disengaged, or lacking proper training. Additionally, officers might inspect a shipment flagged as high-risk by targeting models but fail to detect irregularities that are actually present, leading to false negatives.

Randomized customs controls

In randomized customs controls, every shipment in the entire traffic has an equal chance of being selected for inspection[13]. Random checks can serve various purposes: they may be used to identify risks missed by the risk analysis process, evaluate the effectiveness of existing risk profiles, or assess the performance of an economic operator (Heijmann and Peters 2022). The unpredictability of random controls can also deter smugglers from attempting to exploit perceived gaps in the customs control system.

Randomized controls also help address the problem of non-representative customs control data. By performing random inspections, customs authorities can gather representative data points that provides accurate estimates of the proportion of non-compliant goods in overall trade or its specific segments, such as postal traffic. Although random inspections may temporarily disrupt legitimate trade and place additional strain on customs' limited control resources, the data they yield is invaluable for improving future AI-driven targeting models, offering significant long-term benefits. The rate of randomized controls can be flexible, allowing adjustments based on resource availability and traffic volume.

100% control campaigns

One approach to addressing the issue of non-representative control data is to conduct periodic campaigns targeting 100% of a specific segment of cross-border traffic. Joint customs-police operations sometimes incorporate 100% inspection actions as part of their efforts to combat specific types of illicit trade. Notable examples include Operation Pangaea, targeting counterfeit medicines (INTERPOL 2023), Operation Thunder, focused on wildlife and timber trafficking (WCO 2023), and Operation Athena, aimed at illicit trade in cultural goods (WCO 2020).

These campaigns provide representative data points on cross-border traffic, which can be valuable for AI-driven data analytics. However, it is difficult to consistently generate statistically significant

---

[12] In statistics, larger sample sizes can help reduce issues related to non-representativeness, although a large sample size does not necessarily guarantee that the sample will be representative.

[13] This can be done through a statistical random selection run by a computer program or through a manual system based on an agreed technique that eliminates individual subjectivity.

sample sizes through such campaigns. This is because the benefits of law enforcement campaigns must be weighed against their disruptive impact on legitimate trade and the strain they place on limited customs control resources. It should be noted as well that periodic campaigns may introduce time-based bias, offering only a snapshot of illicit trade during the campaign period.

4.4 Delays in data processing

Crime and law enforcement are often compared to a game of cat and mouse. Criminals constantly develop new methods, only for law enforcement to adapt and counter them. In turn, criminals rapidly adjust their operational approaches to evade detection again. For customs enforcement to respond effectively, timely access to indications of shifting trafficking patterns is essential: the older the data, the less useful it becomes for updating AI-powered risk assessment models and identifying emerging trends.

Academics use the term ‘crime displacement’ to describe the shifting strategies of criminals in response to law enforcement efforts. In the context of customs and illegal trade, the primary types of displacement are geographical and tactical[14]. Geographical displacement occurs, for example, when criminals switch to a different seaport to smuggle goods due to increased customs controls at their original port of activity[15]. Tactical displacement, on the other hand, involves altering concealment methods, such as shifting cocaine smuggling cover loads from fresh flowers to bananas in response to heightened customs scrutiny of flower shipments.

Rapid capture and analysis of control result data

The rapid capture and analysis of control result data are essential for effective customs enforcement and risk management. By equipping customs officers with the tools to quickly collect and transmit data from the field—such as shipment details, irregularities, or compliance checks—the process greatly enhances both the speed and accuracy of risk assessments. Modern digital solutions, including handheld devices, offer frontline officers a convenient and efficient way to report control result data directly from the field. These devices facilitate real-time, structured data entry, ensuring that critical information is captured accurately and transmitted promptly to targeting officers for integration into risk assessment models. With secure cloud synchronization, the data could be instantly uploaded, making it readily available for immediate analysis and decision-making by customs risk management departments.

[14] Literature identifies five general types of crime displacement (Rengert and Hakim 1981): offense displacement, where criminals shift from one illegal activity to another (e.g., from illicit cigarette production to cigarette smuggling); target displacement, involving a change in focus to a different type of contraband (e.g., switching from cigarette smuggling to drug trafficking); spatial displacement, which entails relocating operations to a different area (e.g., moving smuggling activities from one seaport to another); tactical displacement, characterized by the adoption of new methods, such as changing concealment techniques; and temporal displacement, where criminals adjust the timing of their activities to avoid detection (e.g., altering the time of smuggling operations).

[15] For example, some experts predict that stricter customs controls in Belgium and the Netherlands will shift cocaine smuggling routes to German ports (NZZ 2024). This shift is commonly referred to in customs language as 'port-hopping' or 'port-shopping.'

5 Adopting strategic AI orientation at customs

The final chapter provides strategic recommendations for building in-house AI capabilities within customs organizations and promoting sector-wide collaboration to ensure the successful implementation of AI technologies in enforcement operations.

5.1 Building in-house capability within customs organizations

Building an AI-ready customs organization means equipping every level of the workforce—from frontline officers to engineers and top management—with the necessary capabilities to implement AI-powered solutions. Achieving this requires a holistic approach, integrating efforts across all organizational layers to foster AI competency.

Frontline officers as a catalyst for AI implementation

Frontline customs officers, who manage goods and document inspections at border posts, play a pivotal role in operationalizing AI solutions and ensuring their effectiveness in daily enforcement operations. Their involvement is critical to integrating AI tools into routine workflows, making their contributions essential for successful technology implementation.

One priority area is helping officers accept AI-powered tools for risk assessment and data analytics. Field officers must understand how AI supports their work, such as identifying high-risk shipments for inspection, and they should be trained to recognize the value AI brings to their duties. Additionally, frontline officers must be encouraged to critically assess AI outputs, reporting any potential biases or suboptimal decisions to their superiors. Awareness-building initiatives and targeted training programs are essential to fostering this understanding and confidence.

Another significant contribution of frontline officers lies in bridging the reference data gap in detection technologies. Field officers skilled in interpreting customs X-ray images play a crucial role in creating annotated datasets for machine learning applications. High-quality image annotation is essential for developing accurate and reliable AI-powered detection models. To ensure success, customs organizations need to invest in specialized training and provide ongoing motivation for expert image analysts involved in the highly specialized annotation efforts.

Furthermore, frontline officers are essential in addressing the limitations of control result data. Control result forms are a critical source of information for refining AI-powered risk assessment models, but these forms are sometimes overlooked or completed inconsistently by frontline officers. To maximize the value of control result data, customs organizations must encourage frontline officers to approach form completion diligently and efficiently. Training programs should emphasize the importance of accurate data entry and equip officers with tools and techniques to streamline the data collection process.

Engineers as AI innovators and drivers

Customs organizations rely on a diverse group of subject matter experts who are responsible for tasks such as data analysis, process design, and technology integration. These professionals—who can be collectively referred to as engineers—should form the core of an AI-competent customs workforce. This cadre of experts can be built both by hiring new talent and by upskilling current staff with technical AI skills[16].

To overcome barriers to AI-powered data analytics, data scientists, data analysts, and other specialized engineers should build deep in-house technical capacity to harness AI technologies within customs. Their skillset should include selecting appropriate AI techniques, identifying and mitigating potential biases and inefficiencies, preparing data for AI-driven solutions, and designing processes to efficiently gather the necessary data. Benefits of in-house data analytics teams include their agility in responding evolving risks, dynamic trade patterns, and shifting control priorities. Moreover, internal data analytics teams enhance data confidentiality by minimizing the need to share sensitive information with third-party collaborators.

When it comes to building reference datasets for non-intrusive detection technologies, customs organizations should recognize the limitations of their in-house resources. Industry players often possess the engineering expertise, scale, and technical capabilities needed to develop effective threat detection models that are powered by machine vision and machine learning applications. Customs should actively pursue strategic partnerships with these industry players, providing support and collaboration to ensure the development of best-in-class detection models that benefit customs enforcement operations.

Addressing the limitations of control result data is another area where engineers play a key role. By leveraging their expertise, they can design optimal control result forms that include the right data fields and facilitate efficient, officer-friendly reporting of control data. These forms must also be tailored to the needs of AI systems, ensuring that the collected data is structured in a way that supports machine learning applications.

Managers as enablers of AI transformation

Customs organizations must adopt a strategic approach to AI innovation, with managers acting as key enablers of its successful implementation. This entails addressing critical capability gaps and fostering a supportive environment for AI adoption. Customs management should also maintain realistic expectations for AI solutions, recognizing that initial setbacks and modest early outcomes are part of the learning curve and the disruption of traditional customs workflows.

To enhance AI-driven data analytics, customs managers should allocate resources strategically, prioritizing essential investments in hardware and software. Given the challenges of developing proprietary software tools and operating resource-intensive servers for AI applications, the management should secure access to these critical resources through trusted partnerships with reliable external providers. Managers also play a central role in overseeing AI development and deployment. This includes ensuring compliance with legal and ethical standards, mitigating the risks of unintended consequences, and steering AI initiatives toward productive outcomes.

---

[16] For example, Korea Customs has invested heavily in building in-house data analytics capabilities, training three hundred officers as Big Data experts by 2023. While prioritizing the upskilling of existing staff, Korea Customs is also hiring data analytics specialists from outside the organization, drawing on expertise from the private sector and academia to strengthen future operations (Kang 2018).

Attracting, developing, and retaining AI talent is another critical managerial responsibility. As highlighted by the WTO and WCO (2022), the scarcity of data scientists with expertise in customs processes poses a significant barrier to the adoption of AI in customs enforcement domain.

Bridging the reference data gap in detection technologies is also a priority for customs managers. They should actively facilitate collaboration by pooling reference data with other customs administrations and encouraging data-sharing partnerships with trusted industry stakeholders. Additionally, managers should prioritize the systematic curation of existing reference data, such as X-ray images. This involves dedicating resources to data annotation, contextualization, and anonymization work, ensuring the data is prepared and optimized for machine learning purposes.

Managers can also tackle the challenge of limited control result data. They can strategically organize enforcement activities to generate more representative control datasets by implementing periodic 100% control campaigns or increasing the proportion of randomized customs controls. Additionally, managers have the authority to incentivize frontline officers to invest the necessary time and effort in accurately completing control result forms, ensuring the feedback data collected is reliable, comprehensive, and fit for machine learning processes.

In summary, the table below outlines the strategies for frontline officers, engineers, and managers in tackling critical AI capability challenges in customs enforcement.

Table 5 Strategies for overcoming AI capability challenges at three levels of a customs organization

ENFORCEMENT AREA	AI CAPABILITY CHALLENGE	FRONTLINE OFFICERS	ENGINEERS	MANAGERS
Risk assessment	Barriers to AI analytics	Build awareness of AI's value and limitations	Develop in-house AI competence	• Invest in AI hardware and software • Mitigate algorithmic biases and counterproductive outcomes • Prioritize strategic hiring and training
Non-intrusive inspection	Reference data gap in detection technologies	Use expertise to improve data annotations	Support industry partners	• Advocate for reference data sharing • Invest in data curation
Physical examination	Limited control result data	Train for effective reporting of control results	Design officer- and AI-friendly control result forms	• Organize enforcement activities to gather quality control result data

5.2 Collaborative AI capability building in the customs sector

Implementing AI is a challenging endeavor for any customs administration to undertake alone. Therefore, developing AI-ready customs enforcement requires partnerships and strategic cooperation across the broader customs community. This includes collaboration not only among customs administrations worldwide but also with security innovators specializing in data analytics, non-intrusive inspection, and AI technologies.

Developing centralized AI services and data standards for AI analytics

To overcome critical barriers to AI-powered data analytics, customs organizations can join hands to establish centralized AI competencies and analytics services. Centralized services distribute the substantial costs of infrastructure, operational integration, and R&D among multiple contributors, who ultimately share the benefits. In the European Union, initiatives like the redesigned Common Risk Management Framework and Import Control System 2 already provide centralized data analytics for the 27 EU customs administrations. Also the envisioned Customs Data Hub seeks to compile data from end-to-end trade lanes with the support of machine learning, artificial intelligence and other advanced technologies, to provide customs a more detailed view on cargo movements. These EU-wide services offer an ideal platform to advance combined expertise in AI technologies and customs enforcement operations in the future.

International standardization forums play a key role in promoting unique identifiers that enable customs targeting officers to link data across datasets for risk assessment. This integration provides a unified view of global trade flows, enhancing the power of machine learning models by leveraging larger datasets. Widespread adoption of unique identifiers would address data-linking challenges, a critical barrier to effective AI analytics. Platforms such as UN/CEFACT, Universal Postal Union (UPU), International Maritime Organization (IMO), and World Customs Organization (WCO) standardization committees are well-suited to drive the design and adoption of these common identifiers.

Expanding reference data to enhance non-intrusive inspection

Nearly all AI applications rely on some form of machine learning, where AI systems learn by being trained on large volumes of data. As a rule of thumb: the more training data provided, the more confident the model becomes in its predictions. For this reason, it would be greatly contribute to AI implementation if customs had a broader access to large and high-quality reference data libraries.

Improving the availability of quality reference data for machine-learning-based non-intrusive inspection solutions requires collaborative efforts within the customs community. Standard file formats, such as UFF and DICOS, play a vital role in compiling comparable reference data into large libraries for training machine learning models. Designing these formats to be AI-friendly is essential, ensuring they support easy integration into machine learning workflows. For X-ray images, for example, this includes clear guidelines for metadata inclusion and image annotation. Standards are often developed on established platforms that bring together customs practitioners and leading technology companies. These platforms can further facilitate the sharing and pooling of reference data among machine learning developers.

The next challenge after pooling reference data like X-ray images is curating it by adding annotations, metadata, and data privacy safeguards. While archived X-ray images are technically free for customs—as they are a byproduct of routine control activities—cleaning, organizing, annotating these images requires significant time, effort, and expertise. Customs administrations could collaborate to share financial resources and expertise, enabling the efficient curation of large reference datasets for the common good of developing more accurate threat detection models.

The limitations of real-world X-rays and the challenges of sharing X-ray images are likely to result in a persistent shortage of quality reference data for machine learning applications in the customs context. One solution is to generate additional reference data through staged or synthetic images. However, this process requires significant investment, specialized expertise, and tools—resources often beyond the capacity of a single customs administration. Therefore, customs could collaborate to develop the capability to produce synthetic or staged images, or jointly fund industry players to do so on their behalf.

The customs community could also offer incentives to industry players to encourage the sharing of their proprietary reference libraries for broader use. These incentives might include financial compensation or collaborative programs. Alternatively, customs could incorporate clauses into purchasing contracts that require suppliers to align with the vision of greater sharing of proprietary libraries. Alternatively, customs could build their own reference databases and make a parts of them available for academics and industry players, this way fueling competition in the AI industry. As an example, there are public labelled X-ray datasets for baggage inspection for aviation security and safety purposes, such as GDXray and SIXray libraries (Mery et al. 2020).

Improving the availability and quality of control result data

Control result data is crucial for managing customs enforcement operations and developing AI-powered risk assessment models. Customs should seek to improve and standardize the forms used to collect this data, ensuring they are easy for frontline officers to complete and readily understandable for machine learning applications. The forms should be consistently structured with clearly defined fields for key data points, such as threat type, seizure details, and shipment characteristics, along with standardized categorization of irregularities. This standardization would ensure consistent data collection, streamline benchmarking of enforcement performance, minimize the time-intensive data preparation required by analysts, and promote the seamless sharing of trends and insights.

Additionally, there is a need to facilitate the systematic sharing of control result data among customs authorities, enabling the pooling of feedback from customs controls to enhance machine learning-based risk management systems. A step in the right direction is the EU Strategy and Action Plan for Customs Risk Management (EC 2018), which highlights the importance of sharing risk-relevant information and control results among EU customs. It emphasizes that control feedback data should be fully accessible and utilized to enhance effective risk management.

In conclusion, the table below highlights collaborative strategies that the customs sector can adopt collectively to bridge critical AI capability gaps in customs enforcement.

Table 6 Collaborative strategies for overcoming AI capability challenges in the customs sector

ENFORCEMENT AREA	AI CAPABILITY CHALLENGE	ACTIONS IN THE CUSTOMS SECTOR
Risk assessment	Barriers to AI analytics	• Establish central AI competency and analytics services • Promote standards for unique identifiers to support data linking and integration
Non-intrusive inspection	Reference data gap in detection technologies	• Include AI elements in standard reference data formats to support machine learning applications • Leverage existing standardization platforms for pooling reference data • Combine resources for data curation • Generate additional reference data through staged and synthetic X-ray images • Provide incentives for unlocking proprietary reference data libraries
Physical examination	Limited control result data	• Improve and standardize control result forms • Facilitate systematic sharing of control result data

References

Aguirre, S., and Rodriguez, A. (2017). Automation of a business process using robotic process automation (RPA): A case study. In Applied Computer Sciences in Engineering: 4th Workshop on Engineering Applications, WEA 2017, Cartagena, Colombia, September 27-29, 2017, Proceedings 4, pp. 65-71. Springer International Publishing.

Aslett, J., González, I., Hadwick, D., Hamilton, S., Hardy, M., and Pérez., A. (2024). Technical Notes & Manuals: Understanding Artificial Intelligence in Tax and Customs Administration. International Monetary Fund (IMF), TNM/2024/06. Available at https://www.imf.org/-/media/Files/Publications/TNM/2024/English/TNMEA2024006.ashx.

Bakshi, N., Flynn, S., E., Gans, N. (2011) Estimating the Operational Impact of Container Inspections at International Ports. Management Science 57(1), pp. 1-20.

Batista, G. E., Prati, R. C. and Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. ACM SIGKDD explorations newsletter, 6(1), 20-29.

Battelle (2017). Digital Imaging and Communications in Security (DICOS). Available at www.battelle.org/docs/default-source/infrastructure/brochures/dicos_612.pdf.

Chawla, N. V., Bowyer, K. W., Hall, L. O. and Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research, 16, 321-357.

Chen, H., and Chau, M. (2005). Web mining: Machine learning for web applications. Annual Review of Information Science and Technology, 38(1), pp. 289-329.

Daoud, E., Vu, D., Nguyen, H., and Gaedke, M. (2020). Improving fake product detection using ai-based technology. In 18th International Conference e-Society (Vol. 12, pp. 1-9).

DARPA, Defence Advanced Research Projects Agency (2016). Broad Agency Announcement Explainable Artificial Intelligence (XAI). DARPA-BAA-16-53.

Ding, L., Fan, Z.-Z., Chen, D.-L. (2015). Auto-Categorization of HS Code Using Background Net Approach. Procedia Computer Science, 60.

Dixit, S. (2024). Generative AI-Powered Document Processing at Scale with Fraud Detection for Large Financial Organizations. Authorea Preprints.

EC, European Commission (2012). Communication from the Commission to the European Parliament, the Council and the European Economic and Social Committee on Customs Risk Management and Security of the Supply Chain. COM(2012) 793 final. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A52012DC0793.

EC, European Commission (2018). Report from the Commission to the Council and the European Parliament. Second Progress Report on the implementation of the EU Strategy and Action Plan for customs risk management. COM(2018) 549 final. Available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0549.

EC, European Commission (2024). Customs Risk Management Framework (CRMF). Available at https://taxation-customs.ec.europa.eu/customs-4/customs-risk-management/customs-risk-management-framework-crmf_en.

EMSA, European Maritime Safety Agency (EMSA) (2024). EMSA Outlook 2024. Available at www.emsa.europa.eu/publications/corporate-publications/download/7698/5101/23.html.

European Court of Auditors (2021). Customs controls: insufficient harmonisation hampers EU financial interests. Available at https://www.eca.europa.eu/Lists/ECADocuments/SR21_04/SR_Customs_controls_EN.pdf.

Filho, J. (2022). Análise de Risco com uso de Inteligência Artificial na Aduana Brasileira. Presentation.

Ghiran, A., Hakami, A., Bontoux, L. and Scapolo, F. (2020). The Future of Customs in the EU 2040. JRC121859. Available at https://publications.jrc.ec.europa.eu/repository/bitstream/JRC121859/the_future_of_customs_in_the_eu_2040_report_final2_online_1.pdf.

Heijmann, F. and Peters, J. (2022). Customs: Inside Anywhere, Insights Everywhere. Trichis Publishing BV.

INTERPOL (2023). Global illicit medicines targeted by INTERPOL operation. Available at www.interpol.int/News-and-Events/News/2023/Global-illicit-medicines-targeted-by-INTERPOL-operation.

Kafondo, I. (2020). How Can Customs Better Leverage Emerging AI Technologies for More Sustainable and Smarter Operations? World Customs Journal, 14(2), pp. 143–156. Available at www.worldcustomsjournal.org/article/116424-how-can-customs-better-leverage-emerging-ai-technologies-for-more-sustainable-and-smarter-operations.

Kang, T. (2018). Clearance of express cargo and postal items: Korea tests new analytical tools to root out fraud. WCO News 86 — June issue / 2018. Available at https://mag.wcoomd.org/magazine/wco-news-86/clearance-of-express-cargo-and-postal-items-korea-tests-new-analytical-tools-to-root-out-fraud/.

Leevy, J., L., Khoshgoftaar, T., M., Bauder, R., A., Seliya, N. (2018) A survey on addressing high-class imbalance in big data. Journal of Big Data 5(42).

Li, Y., Yang, H., Li, J. Chen, D., Du. M. (2020). EEG-based intention recognition with deep recurrent-convolution neural network: Performance and channel selection by Grad-CAM. Neurocomputing, 415:225-233.

Liang, K., J. (2020). Deep Automatic Threat Recognition: Considerations for Airport X-Ray Baggage Screening. Dissertation, Duke University. Available at https://dukespace.lib.duke.edu/bitstreams/b90ca4cd-973a-4da5-a925-fe8e2c9dca80/download.

Lundberg, S. M., and Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. Adv. Neural Inf. Process. Syst. 30, 4765–4774. Available at: https://arxiv.org/abs/1705.07874.

Mery, D., Saavedra, D., and Prasad, M. (2020). X-ray baggage inspection with computer vision: A survey. Ieee Access, 8, 145620-145633.

Morgan, A. (2024). Leveraging AI for Proactive Customs Compliance: Giving Shipments a Voice. WCO News 104 – Issue 2 / 2024. Available at https://mag.wcoomd.org/magazine/wco-news-104-issue-2-2024/leveraging-ai-for-proactive-customs-compliance-giving-shipments-a-voice/.

NZZ (2024). “Like a game of cat and mouse”: Drug gangs are smuggling record amounts of cocaine through the port of Hamburg. Available at www.nzz.ch/english/cocaine-smuggling-routes-shifting-to-port-of-hamburg-ld.1829335.

Ojanen, A. (2021). AI governance to avoid bias and discrimination. Available at https://demoshelsinki.fi/ai-governance-to-avoid-bias-and-discrimination/.

Ramesh, M. and Vijayakumar, S. (2023). How Indian Customs leverages data analytics for efficient risk management. WCO News 102 – Issue 3 / 2023. Available at https://mag.wcoomd.org/magazine/wco-news-102-issue-3-2023/how-indian-customs-leverages-data-analytics-for-efficient-risk-management.

Rapiscan (2024). Intelligent Image Analytics. Available at www.rapiscan-ase.com/products/software/insight-operator-assist-tools.

Rengert, G., F. and Hakim, S. (1981), Crime spillover. Sage Publications Beverly Hills, Calif.

Russell, S. and Norvig, P. (2021). Artificial Intelligence: A Modern Approach. 4th Edition.

Science Japan (2022). Automatic real-time identification of international mail: system utilizing AI for X-ray image recognition and processing. Available at https://sj.jst.go.jp/news/202203/n0324-01k.html.

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI, International Journal of Human-Computer Studies, 146:102551.

Thompson, R. (2020). Harnessing the Power of AI. Airline Cargo Management 16 - Issue 4 / 2020. Available at www.aviationbusinessnews.com/digital-issue/airline-cargo-management-winter-2020-digital-issue.

Vukadinovic, D. and Anderson, D. (2022). X-ray baggage screening and artificial intelligence (AI): A technical review of machine learning techniques for X-ray baggage screening. RC129088. Available at https://publications.jrc.ec.europa.eu/repository/bitstream/JRC129088/JRC129088_01.pdf.

WCO, World Customs Organization (2010). Kyoto Convention. General Annex Guidelines Chapter 6: Customs Control. Available at www.wcoomd.org/-/media/wco/public/global/pdf/topics/wto-atf/dev/rkc-guidelines-ch-6.pdf.

WCO, World Customs Organization (2020). 101 people arrested and 19,000 artefacts recovered in international crackdown on cultural goods trafficking. Available at www.wcoomd.org/en/media/newsroom/2020/may/101-arrested-and-19000-artefacts-recovered-in-international-crackdown.aspx.

WCO, World Customs Organization (2023). 2,114 seizures of endangered animals and timber as part of a major international law enforcement operation. Available www.wcoomd.org/en/media/newsroom/2023/december/2114-seizures-of-endangered-animals-and-timber-as-part-of-a-major-international-operation.aspx.

WCO, World Customs Organization (2024a). Results of the WCO Smart Customs Survey. Available at www.wcoomd.org/-/media/wco/public/global/pdf/topics/facilitation/activities-and-programmes/smart-customs/smart-customs_public_report_en.pdf?db=web.ra.

WCO, World Customs Organization (2024b). Unified File Format (UFF) 3.0.0 Standard: A technical description of the unified X-ray file format for Non-Intrusive Inspection (NII) devices. Available at www.wcoomd.org/-/media/wco/public/global/pdf/topics/facilitation/instruments-and-tools/tools/safe-package/nii-guidelines-2018/uff-300technical-specificationsen.pdf?db=web.

WCO/WTO, World Customs Organization and World Trade Organization (2022). WCO/WTO Study Report on Disruptive Technologies. Available at www.wcoomd.org/-/media/wco/public/global/pdf/topics/facilitation/instruments-and-tools/tools/wco-wto-joint-report/wco-wto-study-report-on-disruptive-technologies-en.pdf?db=web.

Williams, P. (2021). Intelligent automation: How combining RPA and AI can digitally transform your organization. Available at www.ibm.com/think/topics/ai-for-rpa.

WTO/WCO, World Trade Organization and World Customs Organization (2022). The role of advanced technologies in cross-border trade: A customs perspective. Available at www.wto.org/english/res_e/booksp_e/wcotech22_e.pdf.

Cross-border Research Association (CBRA)
Epalinges, Switzerland
www.cross-border.org