2025 6th International Conference on Computer Vision and Data Mining (ICCVDM)

# IoT Data Mining of Air Cargo for Autonomous Artificial Intelligent Raising of Import Duties by Revenue and Customs

Kareem Ali, Georgios Ioannou, Panos Lovieris, John Cosmas
*Brunel University of London*
Uxbridge, UK
kareem.ali@brunel.ac.uk

Marcus Taylor, Georgios Dimitrakakis
*Planet Computers Ltd*
Bromley, UK
marcus@planetcom.co.uk

Edward Chome, Sahin Mete, Taskin Kizil
*Bewell Technology*
Istanbul Turkey
edward.chome@bewelltech.com.tr

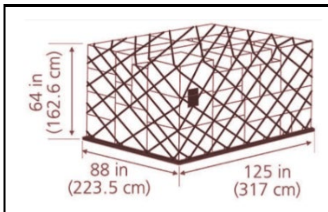
Semih Celik, Betul Bulut
*Vektora*
semih.celik@vektora.com

***Abstract*—The vision presented by OECD's Forum on Tax Administration proposes a digital transformation of tax administration by embedding tax collecting process into the everyday activities that taxpayers use to service their needs rather than being an add-on process. The ATTACC project demonstrates how this could be achieved through Generic Radio Access Network (GRAN)-IoT-enabled cross-border trading. This project seeks to reduce frictions in international trade that stem from data duplication and complex administrative processes to deliver an automated cross-border tax and customs compliance capability and apply AI to analyze large datasets to spot discrepancies between declared income and actual financial activity.**

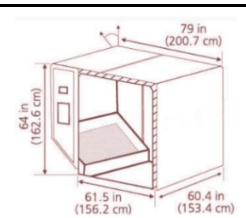
***Keywords*—*4G/5G, IoT, Artificial Intelligence, Revenue and Customs***

## I. Introduction

The aim is to track digitally recorded imported/exported goods loaded on Unit Load Devices (ULDs) as they pass through the border and on airstrips in airports with the use of Generic Radio Access Network – Internet of Things (GRAN-IoT) based localisation in order to perform faster customs clearance and to use AI to detect discrepancies in tax declarations.




Pallets with Nets and Straps




Containers open sided

Figure 1: Unit Load Devices (ULD)

GRAN-IoT tags are attached to pallets, in the form of ULDs or containers and the goods packages on them, as shown in Figure 1, which are attached to 4G/5G network enhanced via an IoT mesh network and used to announce the boarding/landing of ULDs at a GRAN situated at the point of exit/entry in the source/destination countries.

ULDs arrive/depart by airplane at airport terminals such as the ones in Heathrow Airport which consists of four terminals T2, T3, T4 and T5 from which focus on a hypothetical use case will be made on T5. Terminal 5 consists of a passenger terminal building supported by T5, T5B and T5C boarding gates at which numerous aircraft fingers connect the boarding gate in the terminal to the aircraft. 4G/5G network Access points must be able to serve UEs on tow vehicles towing ULD cargo trolleys within its coverage area, as shown in Figure 2.

Figure 2: Typical Tow Trucks for ULDs

Cargo on ULDs are then transported to bonded warehouse for processing and distribution such as the Dnata warehouse at Heathrow Airport, which is located in building 549 in Shorham Road East. IoT tags must be fixed securely on the ULD and have a battery life of 5 years or more so as such they must have a sleep mode during which very little battery energy is used. The IoT sensors must be able to transmit the data that it holds for up to 5 miles. The IoT sensors must have a unit cost of less than £1 and an IoT access node must have a unit cost of < £200.

## II. Overall System Design and Architecture

### A. System Design - George

The ATTACC system consists of a Logistic Traffic Service (LTS), Smart Data Service (SDS) and Automated Tax Compliance service (ATC), as shown in Figure 3.

The LTS tracks the location of ULDs and goods and relays this information to the SDS. The latter combines the localisation data with trade data (e.g. invoices, airway bills etc.) to determine the status of goods with reference to their end-to-end journey; it can identify if a certain shipment is loaded on a ULD, if it is undergoing customs checks at the point of origin, if it has landed at the point of destination etc. The ATC receives combined tracking and trade information from the SDS, to detect taxable events, calculate customs duties and trigger customs payments.

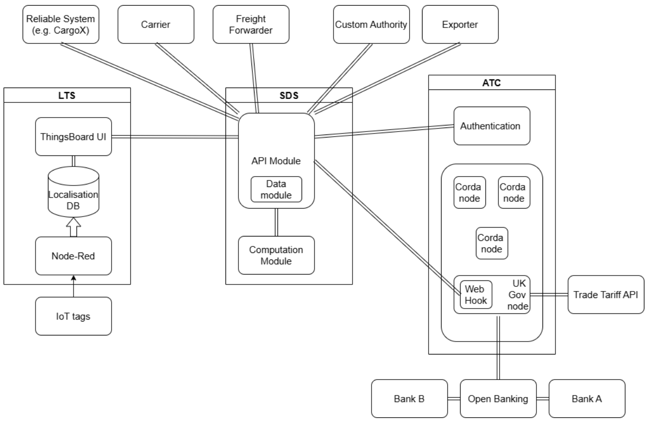



Figure 3: ATTACC System Design

## B. Logistics Traffic System

LTS consists of Fixed and Mobile Terminal Gateways (F/MTG), which are version 4 RPi with 32 GB memory, wireless networked Linux computers that are connected to a Neocortec (NC) Network Gateway Node and a mesh of NC Nodes each attached to a ULD, as shown in Figure 4.

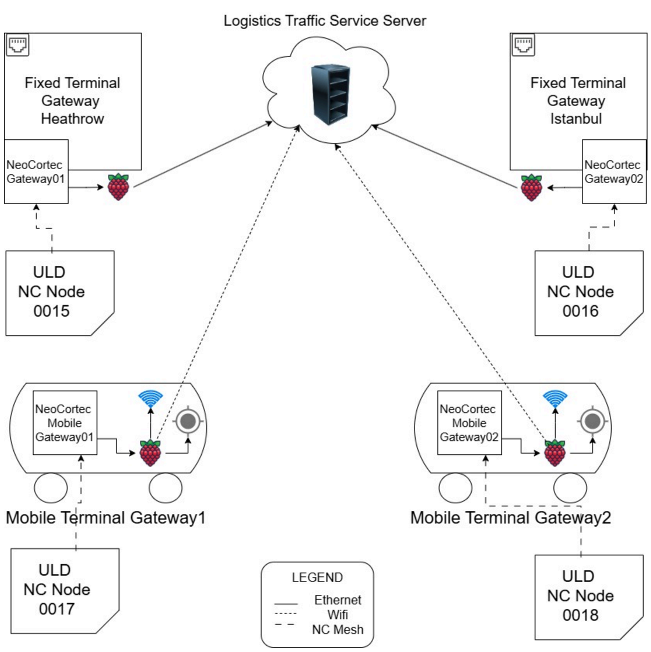



Figure 4: Logistics Traffic System

The FTGs are attached at fixed locations and are designed to monitor all the ULDs within a bonded warehouse. The MTGs are attached to Tow Trucks and are designed to monitor all the ULDs being loaded/unloaded from aircraft and towed from aircraft at boarding gates to bonded warehouses. The location of outdoor MTGs, are captured using a Satellite Global Position System (GPS) antenna, whilst the location of indoor MTGs, are captured using a 5G Time Difference of Arrival (TDoA) localization system to 5 cm accuracy. FTGs are attached to LTS Server with Ethernet cable, whereas MTGs are attached to LTS Server with 4G or 5G mobile network. Tailscale is used to manage access of F/MTGs to LTS Server so we can view all connected devices and secure our network from any attacks because it offers secure and easy-to-manage networking through a secure, private network connection between devices, streamlining remote access and simplifying network management [4]. Tailscale's ease of use, device compatibility, and zero-trust networking principles make it a suitable alternative to traditional VPNs and complex networking setups.

Node-RED is a flow-based development tool for visual programming developed by IBM and built on Node.js [4], which is used to program the RPis to capture and collate NodeID data at F/MTGs and communicate it to LTS server, as shown in Figure 5. Adding or subtracting code for monitoring ULD nodes is simply performed by adding/subtracting an extra finger of flow code with the credentials of the node, as shown in Figure 5. The data that is captured consists of: Timestamp, nodeId, Latitude, Longitude, GatewayId, GatewayType.

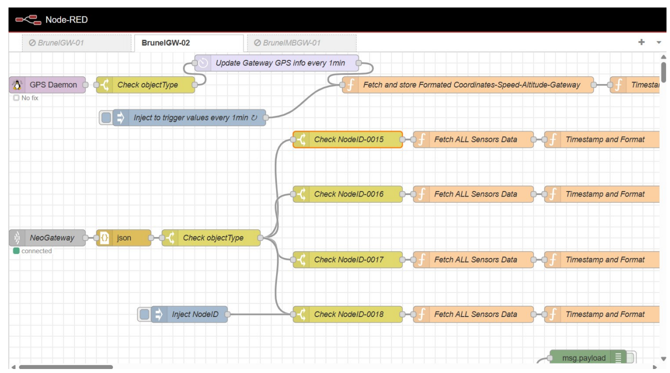



Figure 5: Node-Red program for F/MTGs

## C. Smart Data Service

This section presents the design and implementation of a Smart Data Services (SDS) architecture aimed at enabling secure, interoperable, and scalable logistics information exchange in alignment with the UK 2025 Border Strategy. The proposed architecture supports the Single Trade Window (STW) vision by enabling the “submit once, read many” principle across key stakeholders including importers, exporters, customs, and tax authorities.

The SDS architecture (Figure 6) is structured around three core modules: the Data Module, the API Module, and the Computation Module.

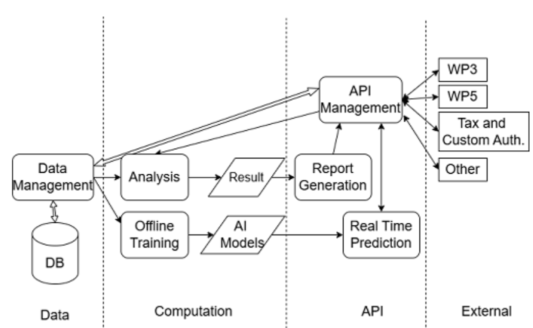



Figure 6: SDS Components

The Data Module is responsible for data acquisition, transformation, and harmonisation, ensuring consistent and standardized data representation across the system. The API Module manages all data exchange within internal modules as well as with external systems, enabling interoperability and integration. The Computation Module undertakes tasks such as anomaly detection, data validation, and analytics, providing intelligence and automation capabilities critical for real-time decision-making in cross-border logistics.

This framework aims to ensure that logistics tracking services are not only operationally efficient but also secure, privacy-aware, and aligned with evolving cross-border regulatory requirements. Moreover, by incorporating advanced data analytics and intelligent systems, the architecture enables proactive risk management, anomaly detection, and enhanced decision support throughout the supply chain.

For anomaly detection, a two-level architecture is designed. Prior to the detection phase, preprocessing and feature generation techniques are applied to enhance data quality and increase the expressiveness of input parameters. Standard feature engineering methods—such as computing ratios, differences, and applying Principal Component Analysis (PCA)—are employed to derive additional informative features. These enriched representations improve the system's capacity to identify both simple and complex anomalies.

In the first level of the detection architecture, univariate statistical methods are applied to individual parameters to detect localized deviations. Specifically, methods such as Z-score, Modified Z-score, Min-Max threshold, and Inter-quartile Range (IQR) are employed to identify outliers within each feature independently. These methods are deliberately chosen for their simplicity, interpretability, and computational efficiency, rather than relying on more complex univariate anomaly detection techniques as surveyed in [5]. This design choice allows for rapid deployment and easier integration into real-time systems, while still providing reliable first-level filtering of anomalous behavior.

In the second level, multivariate techniques are utilized to capture anomalies involving interdependencies among multiple parameters. These include density-based methods like DBSCAN, ensemble-based approaches such as Isolation Forest, and kernel-based models like One-Class SVM, along with deep learning-based anomaly detection algorithms for capturing more complex, non-linear patterns. All detection models are integrated using a weighted ensemble approach, where each model contributes to the final anomaly score based on its relevance and reliability. This hybrid structure enables both fine-grained inspection through single-parameter analysis and robust identification of complex patterns through advanced multivariate methods.

### D. *Automated Tax Compliance*

The Automated Tax Compliance service is based on Making Tax Smart (MTS), an HMRC validated and tested Distributed Ledger Technology-based infrastructure capability [6] for automated tax collection with a Central Bank Digital Currency. ATC is an extension of MTS which uses bank deposits, by incorporating the Open Banking standard [7]. Based on the R3 Corda Community Edition [8], the ATC enables the initiation, issuance and verification of customs payments in distributed payments ledger where information is shared on a need-to-know basis among the interested parties, such as the supplier, the importer, and the customs authority. All the interested parties are represented within the Corda network by means of a Corda account which is stored on a Corda node. The Corda node is maintained by a third-party entity such as a bank or a third-party payment processor, which offers financial services. The ATC assures the integrity of trade documents, provenance of payments and payment finality, which are enablers of financial assurance for customs processes within international trade.

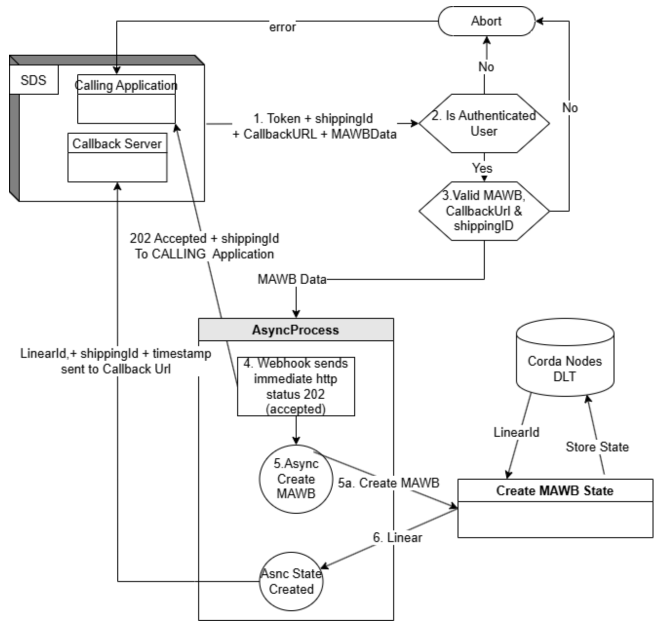



Figure 7: Automated Tax Compliance service and integration with the Smart Data Service

## III. Experimental Setup and Testing

### A. *Experimental Setups*

The network actors are: Network Managers and Network Users. Network Managers are responsible for setting up and managing any F/MTG. The Network Users are responsible to register the ULD Nodes to the RPi and the mesh network, [9].

All F/MTG devices were set up in advance with Tailscale, enabling secure, peer-to-peer network connections and remote system access. Each terminal gateway was assigned the shared Node ID 0x0001, acting as the primary hub for communication across the network. To differentiate between individual devices, ULD nodes were given unique Node IDs starting at 0x0015.

Each NeoCortec module includes a factory-set unique identifier, which was linked to its dedicated Raspberry Pi (RPi) during deployment. This ensured consistent tracking and accurate data attribution per node. Both the Raspberry Pis and NeoCortec nodes were configured to handle data reception, processing, and transmission according to the requirements of the network structure.

### B. *Tests*

User Testing is performed to test how the system interacts with the user against user requirements definition. Functional Testing is performed to test how the system functionally operates correctly against functional requirements definition. Technical Testing is performed to tests how the system performs against performance, reliability and availability technical requirements specifications

#### 1) *User Tests*

Network Manager

How easy was it to register the F/MTG RPi to the network?

How easy was it to view list of F/MTGs?

How easy it is to register a new RFID on a package or ULD?

How easy was it to view list of RFID packages at ULD node?

Network User

How easy it is to register a new ULD node?

How easy was it to view list of ULDs at Fixed/Mobile Terminals?

#### 2) *Functional Tests*

Does registry of a new ULD node operate correctly?

Does list of ULDs at Fixed/Mobile Terminals operate correctly?

Does registry of a new RFID on a package operate correctly?

Does viewing list of RFID packages at ULD node operate correctly?

#### 3) *Technical Tests*

How accurate was the location of the Fixed/Mobile Terminals?

How far can a ULD be to register with a Fixed/Mobile Terminals?

How accurate was the location of the Fixed/Mobile Terminals?

How far can a ULD be to register with a Fixed/Mobile Terminals?

How long does ULD node battery last?

How long does RFID battery last?

## IV. AI INTEGRATION FOR CUSTOMS AUTOMATION

As global trade intensifies and digital borders become more complex, customs authorities face mounting challenges in managing high volumes of transactions and ensuring compliance. The integration of Artificial Intelligence (AI) into the SmartFreight framework represents a transformative step toward automating revenue and customs decision-making.

AI-driven analytics can be utilized by the system to process large-scale datasets from IoT-enabled ULDs, including time-stamped geolocation, cargo type, weight, and declared value. Machine learning models trained on historical customs data can be used to identify anomalies, flag discrepancies between declared and inferred values, and recommend tariff adjustments autonomously. These intelligent models enable predictive tax collection and risk-based inspection, improving efficiency while reducing manual intervention.

Large Language Models (LLMs) further enhance the system's ability to interpret and reason over unstructured and semi-structured data sources. For example, LLMs can analyze ts to extract structured insights, assess compliance risks, and answer complex queries posed by customs officers or auditors in natural language. Additionally, LLMs can be fine-tuned to support multilingual translation and semantic harmonization of trade documentation, which is critical for cross-border operations.

Advanced techniques such as computer vision are also being explored to validate cargo against manifests using image data from camera-equipped drones or terminals. Natural Language Processing (NLP) and LLM-powered understanding further support intelligent automation of documentation analysis and fraud detection.

Integration with orchestration tools like Node-RED enables real-time decision-making by connecting AI modules with edge devices. Edge AI capabilities, supported by lightweight inference engines on Raspberry Pi gateways, allow local processing of AI tasks, reducing latency and maintaining operational resilience even in low-connectivity environments.

The convergence of IoT, AI, and LLMs under this framework supports the OECD's (Organisation for Economic Co-operation and Development) vision of "invisible tax administration," [2] where duties are computed, verified, and collected as an embedded function of digital trade flows, minimizing friction for compliant actors and enhancing detection of fraudulent declaration.

## V. CONCLUSIONS AND FUTURE WORK

This paper presents the design of a system for IoT Data Mining of Air Cargo which can be used for Autonomous Raising of Import Duties by Revenue and Customs and the identification of Artificial Intelligent predictive tax collection and risk-based inspection, which improves efficiency while reducing manual intervention.

## ACKNOWLEDGMENT

The authors gratefully acknowledge support of CELTIC NEXT Research Project ATTACC (AuTomated TAx Compliance for Cross-border Trading with GRAN-IoT) No C2023/2-23.

## REFERENCES

[1] Finsbury Robinson "HMRC: Using technology and AI to close the tax gap".online: HMRC: Using technology and AI to close the tax gap | Finsbury Robinson, 24-6-2024

[2] OECD Forum on Tax Administration, Tax Administration 3.0: The Digital Transformation of Tax Administration, OECD Publishing, Paris, 2020, pp. 6, 15, 23.

[3] Low-code programming for event-driven applications (no date) Node. Available at: https://nodered.org/ (Accessed: 15 June 2025).

[4] Tailscale · Best VPN service for Secure Networks (no date) Tailscale · Best VPN Service for Secure Networks. Available at: https://tailscale.com/ (Accessed: 15 June 2025).

[5] Sun, W., Li, H., Liang, Q. et al. On data efficiency of univariate time series anomaly detection models. J Big Data 11, 83 (2024). https://doi.org/10.1186/s40537-024-00940-7

[6] P. Louvieris, G. Ioannou, and G. White, "Making Tax Smart: Feasibility of Distributed Ledger Technology for Building Tax Compliance Functionality to Central Bank Digital Currency," Appl. Syst. Innov., vol. 7, no. 6, p. 106, 2024. [Online]. Available: https://doi.org/10.3390/asi7060106

[7] H. Wang, S. Ma, H. N. Dai, M. Imran, and T. Wang, "Blockchain-Based Data Privacy Management with Nudge Theory in Open Banking," Future Gener. Comput. Syst., vol. 110, pp. 812–823, 2019.

[8] Khan, C.; Lewis, A.; Rutland, E.; Wan, C.; Rutter, K.; Thompson, C. A Distributed-Ledger Consortium Model for Collaborative Innovation. Computer 2017, 50, 29–37.

[9] “The NeoMesh evaluation kit” Seen online 24-6-2025