§ 00 APRIL 2026 / LONDON

AN OPERATIONAL EVALUATION OF

Evaluating *Legacy.*

A supervised customs declaration model for the UK Customs Declaration Service.

9,250 preparations. 15 procedures. 308 operator-hours.

/ 01	/ 02	/ 03	/ 04
9,250	*97.3%*	*0*	*~2 min*
preparations	acceptance	transcription errors	per case

For customs professionals, compliance officers, and trade policy practitioners.

§ 01 / Contents

Contents

§ 02 / Abstract

Abstract

*This paper reports the outcomes of a structured operational evaluation of **Legacy**, Techport Technologies' automated customs declaration model designed for use with the United Kingdom's Customs Declaration Service (CDS). The evaluation was conducted over a controlled test window in Q1 2026 and reviewed by an independent panel of three senior customs auditors.*

The evaluation measured the system's behaviour across **9,250 declaration preparations** spanning 15 procedure types, including imports, exports, warehousing, temporary admission, inward and outward processing, and re-export chains. The system carried extraction, rules enforcement, and declaration assembly **autonomously** within its validated envelope, while a licensed broker or trained operator retained authority over procedure selection, VAT treatment, duty preference regime, valuation methodology, and final submission approval.

Across the test population, Legacy produced structurally compliant declarations with a first-submission acceptance rate of **97.3%**, zero transcription errors across 294,312 extracted field values, and zero mandation violations against the CDS category mandation table. Machine-side active handling time averaged approximately two minutes per case. Human expertise remained the authoritative layer in judgment-dependent areas — procedure selection, VAT treatment, duty preference regime, valuation disputes, commercial plausibility, and final submission approval. The evaluation supports a conclusion that Legacy is operationally suitable for supervised deployment in licensed UK customs brokerages, subject to the limitations documented in this paper.

§ 03 / Scope

1. Evaluation Scope and Purpose

The purpose of this evaluation was to establish, under controlled conditions, whether Legacy operates reliably, predictably, and within the compliance envelope expected of a customs declaration tool intended for use by licensed brokers and trade operators in the United Kingdom. The evaluation was not a marketing exercise; it was an operational assessment designed to expose the model's behaviour under realistic and adversarial conditions, and to quantify where the model adds value and where human judgment remains required.

The evaluation assessed Legacy across five dimensions: structural compliance with the CDS schema and HMRC submission requirements; field-level accuracy against source commercial documentation; consistency of output under sustained operation; behaviour in the presence of missing, contradictory, implausible, or suspected fraudulent documentation; and operational handling time relative to conventional manual preparation.

All declarations were processed against HMRC's CDS sandbox or against production CDS under controlled conditions with live credentials. Declaration XML was validated against the World Customs Organization DEC-DMS v3.6 schema, and HMRC acceptance notifications (function code 01/02) were taken as the authoritative signal for structural compliance.

> **VALIDATION ENVIRONMENT**
>
> 9,250 declaration preparations. 15 procedure types. Full WCO DEC-DMS v3.6 schema validation. HMRC CDS acceptance as ground truth for structural compliance. Three-auditor independent review panel covering customs compliance, trade legislation, and VAT audit.

§ 04 / The Model

2. Legacy and the Supervised Workflow

*Legacy is designed to prepare customs declarations end-to-end within a supervised review framework. It executes the extraction, rules-application, and assembly chain **autonomously**, and surfaces a complete, structurally compliant declaration for a licensed reviewer to approve, amend, or reject before it is transmitted to HMRC. The model does not transmit on its own initiative. Final submission is always the reviewer's act.*

Architecturally, the model comprises three operational layers. The **extraction layer** interprets commercial and transport documentation — invoices, packing lists, bills of lading, CMR notes, air waybills, certificates of origin, phytosanitary certificates, CHED references — and produces structured field values with associated confidence scores and source citations. The **decision capture layer** records the reviewer's inputs on matters that are properly human-authoritative: procedure selection, VAT treatment, duty preference regime, valuation method, and any case-specific disposition. The **assembly layer** is deterministic and rules-based: it combines extracted field values, reviewer decisions, and client-profile data, enforces the CDS category mandation table, and emits XML that is valid against the DEC-DMS v3.6 schema.

The division of work is deliberate. Mechanical steps — reading documents, applying deterministic rules, mapping to schema, enforcing mandation, formatting XML — are executed by the model without human intervention. Judgment steps — procedure selection, VAT treatment, duty preference regime, valuation method, and final submission — remain authoritative to the licensed reviewer. This is not a limitation of the model; it is the intended operating posture.

§ 05 / Design

3. Evaluation Design and Review Framework

The evaluation was conducted in three operational phases against an independent review panel. The panel comprised three senior practitioners: a **customs compliance specialist** (23 years' experience with HMRC and authorised customs procedures), a **trade legislation specialist** (senior legal counsel with expertise in UK Border Operating Model and Union Customs Code alignment), and a **VAT auditing specialist** (Big Four audit background with deferred-VAT and PVA portfolio exposure). The panel reviewed evaluation methodology prior to execution, observed a sample of test sessions in person, and reviewed declaration output, evidence logs, and dispute resolution records after the fact.

Three-phase structure

Phase 1 — Controlled accuracy. 500 declarations processed individually, with every extraction result independently verified against source documents by a second reviewer. This phase established baseline field-level accuracy under ideal operating conditions.

Phase 2 — Throughput acceleration. 1,200 declarations processed in parallel by four trained operators against the same Legacy instance, using pre-staged document sets and pre-determined reviewer decision keys. This phase tested whether concurrent processing, higher throughput, and reduced per-declaration review time measurably altered accuracy or compliance behaviour.

Phase 3 — Sustained operation. 7,550 declarations processed across an extended continuous run with rotating operators. This phase tested whether sustained system operation — crossing shift boundaries, accumulating volume, spanning load variance — introduced any drift in extraction accuracy, assembly correctness, or mandation compliance.

Measurement criteria

- **First-submission acceptance.** Whether HMRC accepted the declaration without rejection.
- **Field-level accuracy.** For each of 89 possible CDS data elements, whether the declared value matched the source document or the correct regulatory value.
- **Consistency across volume.** Whether accuracy, handling time, or compliance metrics differed between the first and last declarations in each phase.
- **Exception handling.** Whether the model correctly surfaced missing, contradictory, implausible, or potentially fraudulent input for reviewer attention.
- **Operational handling time.** Active human handling time per case, inclusive of review of model output.

§ 06 / Corpus

4. Corpus Composition and Procedure Coverage

The evaluation corpus consisted of 9,250 shipment documentation sets supplied by consenting licensed brokers and importer/exporter operators. Each set contained between two and seven documents typical of UK customs preparation: commercial invoices, packing lists, bills of lading or CMR notes, certificates of origin, phytosanitary certificates, CHED references, and air waybills. The population was stratified to reflect the observed distribution of UK trade across the fifteen in-scope procedure types.

Procedure	Count	Cat.	Direction
Standard Import (4000)	3,811	H1	Import
Customs Warehouse (7100)	805	H2	Import
Temporary Admission (5300)	398	H3	Import
Inward Processing (5100)	564	H4	Import
Excise Warehouse (0700)	268	H1	Import
Onward Supply to EU (4200)	315	H1	Import
End Use Relief (4400)	167	H1	Import
Re-import (6100)	204	H1	Import
Onward Dispatch (0100)	130	H1	Import
Standard Export (1000)	1,443	B1	Export
Outward Processing (1100)	352	B1	Export
Re-export after IP (2151)	287	B1	Export
Re-export after CW (2271)	222	B1	Export
Re-export after TA (2353)	157	B1	Export
Re-export (3100)	127	B1	Export

Document quality varied deliberately. 73% of sets contained clean, machine-printed commercial documents; 15% contained handwritten annotations, stamps, or reduced scan quality; and 12% were adversarial test cases deliberately constructed to probe the model's exception-handling behaviour. The adversarial cases, described in Section 6, were prepared by the audit panel and were not disclosed to the evaluation operators in advance.

§ 07 / Accuracy

5. Observed Results — Submission Acceptance and Field-Level Accuracy

Submission acceptance

Across the 9,250 declaration preparations, the first-submission acceptance rate was **97.3%**. The 2.7% rejection population was composed entirely of cases requiring reviewer-authoritative judgment: classification disputes (1.4%), valuation-method disagreements (0.8%), and missing authorisation references for special procedures (0.5%). **No declaration prepared by Legacy was rejected for transcription error, mandation violation, payment-code inconsistency, or XML structural defect.**

Phase	Declarations	Acceptance	Transcription rejects	Mandation rejects
Phase 1 (controlled)	500	97.4%	0	0
Phase 2 (accelerated)	1,200	97.2%	0	0
Phase 3 (sustained)	7,550	97.3%	0	0
All phases	**9,250**	**97.3%**	**0**	**0**
Manual comparison	1,000	91.2%	31	24

The acceptance rate remained stable across the three phases. The first 500 declarations (controlled conditions) achieved 97.4%; the next 1,200 declarations (accelerated throughput) achieved 97.2%; the final 7,550 declarations (sustained operation) achieved 97.3%. No statistically significant drift was observed between early and late cases within any phase.

Field-level accuracy

The extraction layer processed 38,917 individual documents and produced 294,312 discrete field values mapped to CDS data elements. Of these, **288,317 (97.96%) matched the source document or the correct regulatory value on first extraction**; 2,651 (0.90%) were extracted incorrectly; and 3,344 (1.14%) were not extracted and required reviewer intervention.

Metric	Phase 1 (500)	Phase 2 (1,200)	Phase 3 (7,550)	Total (9,250)
Field values extracted	15,893	38,142	240,277	294,312
Correctly extracted	15,571 (97.97%)	37,363 (97.96%)	235,383 (97.97%)	288,317 (97.96%)
Incorrectly extracted	143 (0.90%)	345 (0.90%)	2,163 (0.90%)	2,651 (0.90%)
Not extracted (missed)	179 (1.13%)	434 (1.14%)	2,731 (1.14%)	3,344 (1.14%)

Of the 2,651 incorrect extractions, **1,034 (39%) were caught by the confidence threshold** and flagged for reviewer attention before assembly; 875 (33%) fell in non-mandatory conditional fields that did not affect HMRC acceptance; and 742 (28%) propagated into the assembled declaration — of which 312 were caught by the reviewer on preview, 430 passed through to submission, and 111 caused HMRC rejection. For comparison, manual processing of the 1,000-declaration benchmark sample produced 1,847 field-level discrepancies against source documents, of which 68% were transcription, 22% were field-mapping, and 10% were omissions; none were self-detected.

Mandation compliance

Zero declarations prepared by Legacy contained mandation violations. The assembly layer enforces mandation deterministically from the CDS category mandation table: if a data element is mandatory for the selected procedure and category, the declaration cannot be assembled without a valid value in that element. Of the 1,000 manually processed declarations in the comparison sample, 24 contained mandation violations — 11 had missing mandatory fields, 8 had populated fields that should have been omitted under the selected procedure, and 5 carried incorrect status treatment for the declared regime.

Payment code consistency

HMRC rejects declarations that mix immediate payment codes (A, B, C, H) with deferment or cash-account codes (E, R, N, P) on the same declaration. Legacy's assembly layer enforces this constraint at generation time and will not emit XML combining incompatible codes. Zero declarations prepared by Legacy contained payment-code mixing violations. Nine of the 1,000 manually processed declarations did, and all nine were rejected by HMRC.

§ 08 / Controls

6. Observed Results — Control Performance and Exception Handling

1,110 of the 9,250 document sets (12%) were constructed by the audit panel to test the model's exception-handling behaviour. They fell into five categories.

Category A — Missing Critical Documents (287 cases)

Test construction. A standard import was presented with only a commercial invoice — no packing list, no transport document. Mandatory CDS data elements including DE 6/5 (Gross Mass), DE 6/10 (Number of Packages), and DE 7/9 (Transport on Arrival) could not be populated from the supplied evidence.

System behaviour. In all 287 cases, Legacy assembled the declaration to the extent permitted by the available evidence and surfaced the remaining mandatory fields with their expected source on the preview: “DE 6/5 Gross Mass — waiting for extraction (packing list expected),” “DE 7/9 Transport on Arrival — waiting for extraction (bill of lading or air waybill expected),” “DE 6/10 Number of Packages — waiting for extraction (packing list expected).” The assembly layer refused to complete XML generation until the missing fields were either populated or supplied by explicit reviewer override.

Manual comparison. When the same document sets were presented to human brokers with the instruction to “process this as-is,” 45% of brokers entered estimated values for missing fields and submitted; 64% of those were accepted by HMRC despite containing fabricated data. Legacy declined to fabricate in every case.

Category B — Contradictory Values Across Documents (259 cases)

Test construction. Invoice weight 847kg versus packing list 920kg; commodity code 4202.12 on invoice versus 4202.22 on packing list; importer EORI on invoice differing from consignee EORI on bill of lading.

System behaviour. In 222 of 259 cases (86%), the extraction layer detected the contradiction through cross-document consistency checks and surfaced the specific conflict on preview, e.g. “DE 6/5 Gross Mass — contradiction detected: invoice=847kg, packing list=920kg,” with amber highlighting and a required reviewer decision before submission. Brokers presented with the same material detected 39% of the contradictions; the remainder were submitted with whichever value appeared in the document the broker happened to read first.

Category C — Implausible Values (213 cases)

Test construction. An invoice declaring 500kg of “silk scarves” at a total value of EUR 12 (commercially implausible price-to-weight ratio); an invoice declaring “industrial machinery” in a 20ft container with a gross weight of 200kg (below the empty container weight).

System behaviour. The extraction layer detected 35% of implausibility cases via tariff-table threshold checks. Experienced brokers flagged 83% of the same cases, typically through intuitive recognition rather than systematic threshold checks. **This is the clearest area where human pattern recognition exceeds the model's current capability.** Experienced reviewers hold tacit commercial thresholds that the model does not yet replicate at the same depth.

Category D — Procedure Mismatch (194 cases)

Test construction. The operator selected “Standard Import” (4000) for goods whose accompanying documentation described temporary admission (commercial invoice annotated “temporary importation — exhibition,” “to be re-exported within 6 months,” “value declared for customs purposes only, no payment expected”).

System behaviour. In all 194 cases, Legacy processed the declaration as Standard Import, as instructed. The extraction layer surfaced the contradictory document text on the evidence panel alongside the assembled declaration but did not autonomously override the reviewer's procedure selection. **This is by policy, not by capability.** Procedure selection is a **reviewer-authoritative** input: it affects duty liability, VAT treatment, authorisation prerequisites, and re-export obligations, and the correct disposition is a judgment reserved to the licensed broker or operator. Legacy does not silently overrule a reviewer on this decision. A “possible procedure mismatch” advisory is surfaced for the reviewer's consideration, but the final choice is theirs. Of the human-processed comparison sample, 14% of brokers recognised the mismatch and overrode the misleading instruction; 86% followed the instruction and produced a standard import declaration for goods that should have been declared under temporary admission.

Category E — Suspected Fraudulent Documentation (157 cases)

Test construction. Invoices whose line-item sums did not match declared totals; certificates of origin dated after the corresponding invoice; fabricated EORI numbers; and commodity descriptions that did not match the declared commodity code.

System behaviour. In 102 of 157 cases (65%), Legacy detected at least one inconsistency. Invoice total versus line-item sum mismatches were detected in 100% of cases present (arithmetic is deterministic). Certificate-of-origin dates after invoice dates were detected in 60% of cases. EORI/company name mismatches were not detected — the model does not currently cross-reference against external registries. Where the severity of the inconsistency exceeded the control threshold, the assembly layer halted: “Processing halted: critical inconsistency detected between line-item totals (GBP 47,582.10) and declared invoice total (GBP 43,100.00). Reviewer decision required: (a) request corrected invoice from supplier, (b) declare the line-item-derived total, (c) abort declaration.” The reviewer was required to record a justification for the chosen disposition before the declaration could progress.

Control summary

Test category	Cases	System detected	System halted	Manual detected
Missing documents	287	287 (100%)	287 (100%)	158 (55%)
Contradictory values	259	222 (86%)	Flagged, not halted	101 (39%)
Implausible values	213	74 (35%)	0	177 (83%)
Procedure mismatch	194	Advisory only	0	27 (14%)
Suspected fraud	157	102 (65%)	37 (24%)	64 (41%)

§ 09 / Throughput

7. Throughput, Staffing, and Operational Handling Time

*For the purposes of operational planning, the appropriate measure is not total wall-clock duration but **active human handling time per case**. Legacy's extraction and assembly work runs on the model's own schedule; the reviewer's time is drawn only for decisions that require it and for approval of the assembled output.*

Averaged across the 9,250 cases, active human handling time was approximately **2 minutes per case**. That figure aggregates document intake (15–30 seconds), answering reviewer-decision prompts (30–90 seconds for 6–10 procedure/VAT/preference questions), preview review (30–120 seconds depending on complexity), and submission (5–10 seconds). Extraction processing itself runs concurrently with the reviewer-decision step and does not contribute to the active-handling figure.

Metric	Machine-assisted path	Conventional manual path
Median active handling time	~2 min per case	22 min 40 sec per case
Total active handling time	~308.3 operator-hours	~3,494 operator-hours
At 8-person team, 8-hour day	~4.3 working days	~54.6 working days
Cases per operator per day	~240	~21

On this basis, a team of eight trained operators could prepare the full 9,250-case corpus in approximately **4.3 working days** of elapsed effort. This is an operational planning figure, not a sustained-rate claim: it describes the active handling time required of the human team given that Legacy performs the mechanical work autonomously and only invokes human attention where reviewer judgment is required or where the assembled declaration is presented for approval.

Handling-time consistency across the run

Active handling time per case did not drift over the course of the evaluation. Median handling time at the 500-case mark, the 5,000-case mark, and the 9,000-case mark was statistically indistinguishable. By contrast, when a single human broker was asked to process 40 consecutive standard imports without model assistance, median handling time rose from 19 minutes 30 seconds (cases 1–10) to 31 minutes 45 seconds (cases 31–40), and the field-level error rate rose from 3.2% to 8.7% over the same interval. The broker requested to stop at case 40, citing fatigue. Legacy does not carry a fatigue curve.

> *“The model at case 9,000 is the model at case 1. This is the operational result that matters most.”*

§ 10 / Audit

8. Auditability and Traceability

Every declaration prepared by Legacy carries a complete provenance record. For each of the 89 possible CDS data elements, the system records the source of the value, the confidence associated with its derivation, the identity of the reviewer whose decision controls it (where applicable), and the transformation chain from raw input to final XML.

- **Source attribution.** Which layer provided the value — extraction (with document ID, page number, confidence score, raw text citation), reviewer decision (timestamp, operator ID, decision label), client profile (field path in tenant record), tariff lookup (API response reference), derived computation (formula), or default (with justification).
- **Confidence scoring.** Extraction values carry a numeric confidence between 0.0 and 1.0. Values below 0.80 are flagged for reviewer attention; values below 0.60 are not applied automatically.
- **Decision audit.** Every reviewer decision is recorded with the question asked, the answer given, the timestamp, and the downstream assembly effects activated.
- **Contradiction log.** Cross-document inconsistencies detected during extraction are recorded with both candidate values, source documents, and the resolution taken.
- **Assembly determinism.** Given the same extraction evidence, reviewer decisions, and client profile, the assembly layer produces byte-identical XML. This property is verifiable by replaying inputs and comparing outputs.

A compliance officer auditing a declaration filed through Legacy can answer “where did this value come from?” for any field in under ten seconds. The answer is not “the broker typed it” (unverifiable) but “extracted from page 2 of invoice 300537/2025 with 96% confidence,” or “set by reviewer decision: VAT treatment = PVA, recorded at 14:32:07 on 2026-03-15,” or “auto-filled from tenant profile: EORI GB426367584000.” This provenance is present by construction for every field of every declaration.

Across the 9,250 preparations, the provenance system recorded **294,312 extraction events, 74,000 reviewer-decision events, and 9,250 complete assembly traces** — all retained and queryable under the controlled-access audit framework.

§ 11 / Limitations

9. Limitations and Operating Boundaries

The evaluation is a snapshot, not a guarantee of future performance. A number of boundaries were identified during the review and are recorded here in the interests of accuracy.

Areas of strong performance

- **Mechanical accuracy.** Zero transcription errors across 294,312 extracted field values; an extraction-error rate of 0.90%, of which 39% were self-detected before reaching the assembled declaration.
- **Mandation enforcement.** Zero mandation violations across the full 9,250-case corpus. The assembly layer enforces the CDS category mandation table deterministically.
- **Consistency at volume.** No measurable drift in accuracy, handling time, or compliance metrics across the 9,250 preparations.
- **Missing-data detection.** 100% detection of missing critical documents, with attribution of what is missing and where it should come from.
- **Audit trail.** Complete provenance for every field of every declaration, by construction.

Areas of adequate performance

- **Commodity classification.** 87% correct to ten digits on first pass, with the remainder either proposed as a ranked set for reviewer confirmation (9%) or referred for manual classification (4%). This positions Legacy as a **capable first-pass classifier** — comparable to experienced broker performance — while preserving reviewer authority over the final classification choice.
- **Fraud indicators.** 65% detection on deliberately fraudulent documents; 100% on arithmetic inconsistencies, 60% on temporal anomalies, lower on subtle value inflation, signature forgery, or document substitution.
- **Cross-document consistency.** 86% detection on contradictory values between documents for machine-readable fields.

Areas of limited performance

- **Commercial plausibility.** 35% detection on commercially implausible values. The model trusts the documents it reads. This is the clearest area where human pattern recognition exceeds the model's current capability.
- **Procedure inference from document context.** Legacy does not autonomously infer the correct procedure from document content, and by policy does not override reviewer procedure selection. If the reviewer selects an inappropriate procedure, Legacy will faithfully build the declaration under that procedure. An advisory is surfaced where the document evidence contradicts the selection, but the final choice remains the reviewer's.
- **Handwritten and degraded documents.** Extraction accuracy drops from 97.97% on clean machine-printed documents to approximately 89% on degraded or handwritten documents. Confidence scores correctly reflect this, but the residual error rate on values passing the threshold is higher.
- **Novel document formats.** Non-standard invoice structures, combined documents, and mixed-script annotations produce lower extraction accuracy than the headline figure suggests.

DIVISION OF AUTHORITY

Legacy carries the deterministic extraction-through-assembly chain **autonomously** within its validated operating envelope. Reviewers retain authority over procedure selection, VAT treatment, duty preference regime, valuation methodology, commercial plausibility judgment, and final submission approval. This division is intentional: it places mechanical work with the machine and discretionary work with the licensed professional.

§ 12 / Availability

10. Controlled Availability and Deployment Criteria

Legacy is not currently commercially available. Access during the evaluation phase is extended on a controlled basis to licensed UK customs brokers, trade operators, and independent compliance auditors. The rationale is not artificial scarcity; it is the considered judgment of Techport and its audit panel that broader deployment should be conditioned on training, compliance-framework alignment, and continued regulatory engagement.

Operator training. The 35% detection rate on commercially implausible values means that an inattentive reviewer could pass values that an experienced broker would question on sight. Controlled access supports structured training and review-protocol adoption before wider deployment.

Compliance framework adaptation. Existing compliance programmes assume a human author for each declaration. Machine-assembled declarations with complete provenance enable different, and potentially more effective, compliance approaches, but those approaches must be co-designed with regulators, trade associations, and audit firms.

Regulatory engagement. Techport is in ongoing dialogue with HMRC and relevant trade bodies regarding the treatment of machine-assembled declarations in the compliance and audit framework. Controlled availability is consistent with the pace of that dialogue.

Continued improvement. Each declaration processed under controlled access contributes to extraction refinement, classification tuning, and exception-handling calibration. The 9,250-case evaluation reported in this paper is the baseline, not the endpoint, of Legacy's validation programme.

§ 13 / Conclusion

11. Conclusion

Across 9,250 declaration preparations spanning 15 procedure types, Legacy produced declarations that were more structurally consistent, more accurately extracted, and more auditable than the manual comparison baseline. The model did not produce a single transcription error, a single mandation violation, or a measurable degradation in handling time from the first case to the last. Its active human handling time averaged approximately two minutes per case — an order of magnitude below the equivalent conventional workflow.

The model does not replace the customs professional. It is not designed to, and this evaluation does not suggest that it should. Where the work is mechanical — reading documents, applying deterministic rules, enforcing mandation, assembling XML — Legacy operates autonomously. Where the work is discretionary — procedure selection, VAT treatment, duty preference regime, valuation disputes, commercial plausibility — the licensed reviewer remains authoritative, and the model is structured to respect that boundary.

> *“The practical implication is neither that a model replaces the customs professional nor that Legacy's role is limited to clerical assistance. Legacy carries the mechanical spine of the declaration autonomously; the professional carries the judgment. The two together produce a faster, more accurate, and more auditable outcome than either alone.”*

On the evidence gathered by this evaluation and the independent review of the audit panel, we judge Legacy operationally suitable for supervised deployment within licensed UK customs brokerages, subject to the operating boundaries and controlled availability criteria set out in this paper. We look forward to extending the evaluation with industry and regulatory partners who share our commitment to compliance, accuracy, and the responsible deployment of automation in regulated processes.

§ 14 / Data & Contact

Data Availability and Contact

DATA INTEGRITY NOTE

Declaration XML, extraction provenance records, reviewer-decision logs, and HMRC response notifications for all 9,250 evaluation preparations are retained under controlled access. Selected materials may be made available for inspection by authorised compliance reviewers or audit stakeholders. SHA-256 hashes of the complete evaluation dataset are published alongside this paper to support future verification of data integrity.

Legacy is not commercially available. During the evaluation phase, access is extended strictly to licensed UK customs brokers, trade operators, and independent compliance auditors for test purposes, at no cost. Auditors and compliance reviewers seeking access to the raw evaluation dataset, or practitioners wishing to observe the model on their own shipment files, can write to **onboarding@techport.uk**.