International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue 5
Sep-Oct 2026, PP 01-7

Integrated Edge AI Framework for Non-Destructive Gold and Cultural Heritage Authentication: A Review

Mohate Samiksha Balasaheb[1], Bangar Abhishek Popat[2], Bhosale Sachin Bajirao[3]

[1]Department of Computer Engineering / Jaihind College of Engineering / SPPU, Pune, India

[2]Department of Computer Engineering / Jaihind College of Engineering / SPPU, Pune, India

[3]Department of Computer Engineering / Jaihind College of Engineering / SPPU, Pune, India

Abstract. Ensuring accurate and compliant Legal Metrology declarations on pre-packaged commodities and commercial instruments is essential for consumer protection and fair trade, yet manual inspection of product labels and regulatory filings remains time-consuming, error-prone, and inconsistent across high-volume supply chains. Existing regulatory compliance workflows often struggle to bridge the gap between complex, evolving statutory text and real-time operational verification. To address these challenges, this paper presents an Automated Compliance Checker for Legal Metrology Declarations that combines natural language processing with deterministic policy-as-code validation. In the proposed framework, statutory metrology mandates—such as mandatory manufacturer or importer details, standard International System (SI) unit formatting for net quantity, Maximum Retail Price (MRP) inclusivity disclosures, date formatting, and consumer care information—are decomposed into atomic, machine-verifiable conditions. A hybrid architecture utilizes Large Language Models (LLMs) and Optical Character Recognition (OCR) to extract candidate declaration fields from unstructured product packaging and generate human-readable non-compliance explanations, while delegating exact unit checks, numeric tolerances, and mandatory-field validations to deterministic backend services to prevent hallucinated assessments. Each verification produces timestamped, immutable compliance artifacts that quantify condition satisfaction rates and support auditable human-in-the-loop review. Experimental evaluation across representative declaration datasets demonstrates that the system significantly reduces inspection latency while achieving high precision and recall in detecting labeling omissions and formatting violations, offering a scalable RegTech solution for manufacturers, importers, and regulatory authorities.

Keywords- Integrated Edge AI, Non-Destructive Testing (NDT), Gold Authentication, Cultural Heritage Preservation, Artificial Intelligence, Machine Learning, Deep Learning, Edge Computing, Smart Sensors, Material Characterization, Precious Metal Verification, Digital Heritage, Real-Time Analysis, Explainable AI (XAI), IoT-Based Authentication, Computer Vision, Data Security, Artifact Preservation, Intelligent Authentication Systems, Sustainable Heritage Conservation, AI-Based Material Identification.

I. Introduction

Legal metrology regulates measurements and pre-packaged commodities to protect consumers and ensure fair trade. Under Legal Metrology and Packaged Commodities rules, manufacturers, packers, and importers must display mandatory declarations on every retail package—including manufacturer or importer details, generic commodity name, net quantity in standard International System (SI) units, date of manufacture or import, Maximum Retail Price (MRP) inclusive of all taxes, country of origin, and consumer care contacts. Non-compliance leads to fines, product seizures, customs delays, and loss of consumer trust.

Despite strict enforcement, declaration verification across retail and e-commerce supply chains still relies heavily on manual inspection. Compliance teams and regulators manually cross-check label artwork and digital listings against complex checklists, creating a slow, inconsistent, and error-prone bottleneck as Stock Keeping Units (SKUs) scale and rules evolve. Moreover, post-market sampling often catches violations only after non-compliant products reach consumers

Regulatory Technology (RegTech) addresses this gap by translating legal mandates into executable *policy-as-code* for real-time verification. However, automating Legal Metrology checks poses two core challenges: product labels are semi-structured and require Optical Character Recognition (OCR) and Natural Language Processing (NLP) for field extraction, whereas standalone Large Language Models (LLMs) struggle with strict deterministic rules—such as exact SI unit validation, numeric thresholds, and date checks—risking hallucinated judgments.

To solve this, we propose an Automated Compliance Checker for Legal Metrology Declarations using a hybrid AI and deterministic architecture. Statutory mandates are decomposed into atomic, machine-verifiable conditions. An OCR/LLM layer extracts declaration fields from unstructured labels and generates clear violation explanations, while deterministic backend services execute exact unit, numeric, and completeness checks. Each verification outputs a timestamped, auditable compliance record to support human-in-the-loop review.

II. Review Methodology

Our Automated Compliance Checker for Legal Metrology Declarations operates as a four-stage hybrid pipeline that converts unstructured label images and digital product listings into auditable compliance verdicts. First, in the statutory rule decomposition stage, mandatory provisions under the Legal Metrology Act and Packaged Commodities Rules are formalized into a version-controlled catalog of atomic, machine-verifiable predicates covering manufacturer or importer details, Country of Origin, net quantity in standard SI units, Maximum Retail Price (MRP) with tax-inclusivity disclosures, month and year of manufacture or import, shelf-life ordering, and consumer care contact details. Second, during multi-modal ingestion and semantic field extraction, the system processes physical package images, artwork proofs, or e-commerce listing payloads through multi-pass OCR and a multimodal vision-language parser to normalize noisy text into a canonical JSON schema alongside field-level extraction confidence scores. Third, in the deterministic Policy-as-Code validation stage, the extracted JSON payload is evaluated by backend rule services using regular expressions, standardized SI unit lookup tables, date parsers, and cross-field consistency checks to prevent generative hallucinations, computing a weighted Compliance Satisfaction Score (ScS_cSc) while flagging any failure on a mandatory rule as non-compliant. Finally, in the audit logging and explainable reporting stage, every verification generates an immutable, timestamped evidence record storing the input hash, extracted attributes, rule-by-rule outcomes, and rule version metadata, while a constrained LLM explainer translates deterministic error codes into clear remediation guidance and routes low-confidence extractions to a human-in-the-loop compliance dashboard for inspector sign-off.

III. Literature Review

1. Statutory Framework and Digital Enforcement Challenges

Under statutory regimes such as the Legal Metrology Act, 2009 and the Legal Metrology (Packaged Commodities) Rules, 2011, manufacturers, packers, and importers are mandated to display standardized disclosures—including entity name and address, generic commodity name, net quantity in standard International System (SI) units, Maximum Retail Price (MRP) inclusive of all taxes, month and year of manufacture or import, Country of Origin, and consumer care contacts. While enterprise compliance tools have emerged to digitize registration and label record-keeping, enforcement across physical retail and high-volume e-commerce marketplaces still relies heavily on manual sampling. Manual auditing across millions of Stock Keeping Units (SKUs) is slow, inconsistent, and vulnerable to seller omissions or non-standard unit representations.

2. Policy-as-Code and Deterministic Rule Engines

Traditional Regulatory Technology (RegTech) frameworks formalize statutory provisions into explicit, machine-verifiable predicates—commonly termed *Policy-as-Code* (PaC). By encoding statutory clauses (such as Rule 6 declaration checklists)

into deterministic regular expressions, unit-validation tables, and boolean constraints, rule-based engines achieve exact precision, zero generative hallucination, and reproducible audit trails. However, purely deterministic engines require rigidly structured inputs; they cannot directly parse noisy packaging imagery, curved surfaces, or variably phrased digital product descriptions without an upstream extraction layer, nor can they adapt effortlessly to regulatory amendments without manual rule re-engineering.

3. OCR, Computer Vision, and Multimodal LLM Extraction

To bridge the gap between physical packaging and digital rule evaluation, recent automated inspection prototypes combine Optical Character Recognition (OCR), computer vision, and multimodal Large Language Models (LLMs). Multi-pass OCR and web-scraping pipelines extract text from product label photographs, print artwork, and e-commerce product pages, enabling automated detection of missing MRP, net quantity, or origin disclosures. Although vision-language models excel at handling diverse fonts, multilingual packaging, and spatial layout cues, relying solely on generative LLMs for end-to-end compliance verdicts introduces critical reliability risks: standalone LLMs frequently miscalculate numeric tolerances, overlook subtle formatting violations (such as non-standard SI unit symbols or missing tax-inclusivity qualifiers), and lack deterministic traceability for legal enforcement.

4. Hybrid AI–Deterministic Governance and Research Gap

To resolve the trade-off between semantic flexibility and statutory precision, modern compliance architectures adopt a hybrid separation of concerns. In these frameworks, probabilistic AI components (OCR and LLMs) handle unstructured field extraction, entity normalization, and human-readable violation explanations, while an independent deterministic rule engine executes exact normative validations and logs immutable, timestamped audit records. Despite growing interest in automated label scanning and e-commerce compliance dashboards, existing prototypes often treat OCR extraction and rule checking as loosely coupled scripts without formal rule-to-evidence quantification, confidence-aware human-in-the-loop routing, or cross-modal verification between physical label artwork and digital marketplace listings. This work addresses that gap by integrating atomic policy-as-code decomposition with hybrid OCR/LLM extraction and auditable compliance scoring.

IV. Comparative Analysis

To evaluate the operational advantages of our proposed framework, we conduct a multi-dimensional comparative analysis against three established compliance paradigms: (1) Manual Regulatory Inspection**, (2)** Pure Deterministic Rule Engines (Regex/Schema-Based)**, and (3)** Standalone OCR + Generative LLM Checkers**.**

1. Architectural & Functional Comparison

Evaluation Dimension	Manual Inspection	Pure Deterministic Rule Engine	Standalone OCR + LLM Checker	Proposed Hybrid AI + Policy-as-Code
Input Handling	Physical packages, PDF artwork, web pages	Rigidly structured JSON/CSV/HTML schemas only	Unstructured images, noisy OCR text, web listings	Multi-modal (label images, artwork PDFs, and structured e-commerce payloads)
SI Unit & Numeric Exactness	High (subject to reviewer fatigue)	100% Deterministic (exact lookup tables & regex)	Moderate–Low (prone to arithmetic/unit hallucinations)	100% Deterministic (delegated to backend rule services)

Evaluation Dimension	Manual Inspection	Pure Deterministic Rule Engine	Standalone OCR + LLM Checker	Proposed Hybrid AI + Policy-as-Code
Adaptability to Rule Amendments	Requires staff retraining & checklist updates	High maintenance (requires hardcoded logic rewrites)	High (prompt updates, but lacks version control)	High (version-controlled atomic Policy-as-Code rules)
Inspection Latency (per SKU)	Minutes to hours	Milliseconds	Seconds	Seconds (real-time batch & single-item verification)
Scalability across SKUs	Very Low (linear with human headcount)	Very High (for pre-structured catalogs)	High	Very High (cloud-scalable pipeline)
Auditability & Legal Traceability	Paper logs / manual spreadsheets	High (deterministic rule trace)	Low (black-box probabilistic outputs)	High (immutable timestamped artifacts, input hashes & rule IDs)
Explainability & Remediation	Custom inspector notes	Cryptic error codes (e.g., ERR_UNIT_02)	Fluent natural language (risk of hallucinated citations)	Grounded natural-language guidance mapped to exact rule IDs
Uncertainty Handling	Subjective inspector judgment	Fails/crashes on malformed inputs	Overconfident predictions on blurry text	Field-level confidence scoring + Human-in-the-Loop routing

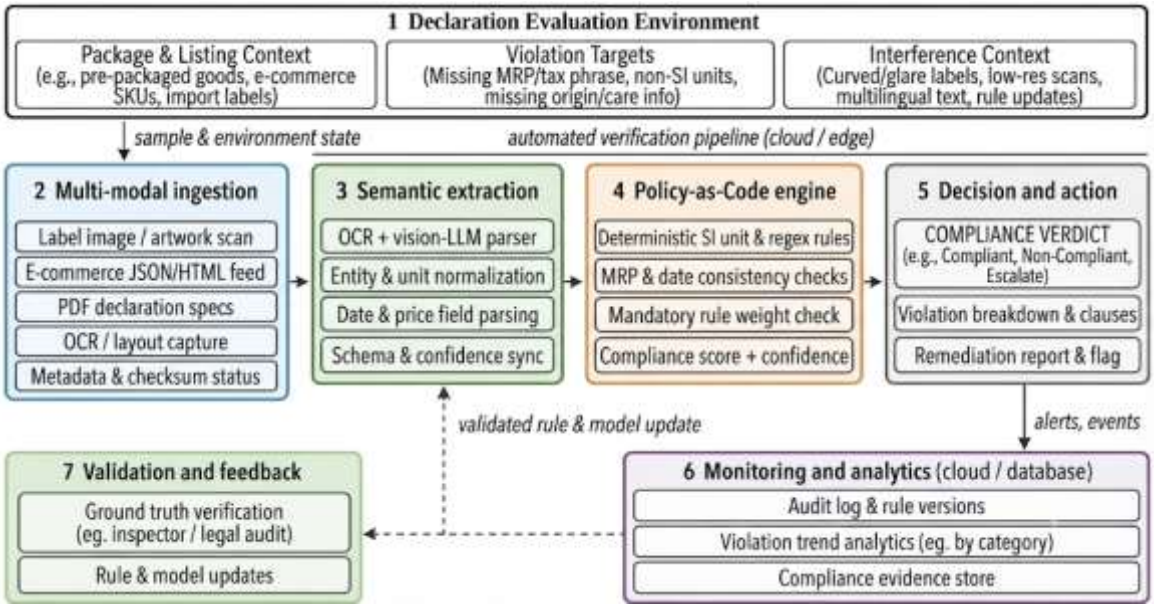
2. Key Comparative Insights

- **Overcoming the Brittleness of Pure Rule Engines:** Traditional web scrapers and regex-based checkers achieve exact accuracy when validating clean e-commerce metadata fields, but fail when mandatory disclosures (such as MRP or manufacturing date) are embedded inside packaging images or phrased variably across sellers. By introducing a multi-pass OCR and semantic extraction layer upstream, our architecture normalizes unstructured visual labels into canonical JSON schemas before rule execution.
- **Eliminating Generative Hallucinations in Standalone LLMs:** End-to-end LLM compliance checkers can read messy labels, yet they frequently misjudge strict Legal Metrology formatting constraints—such as distinguishing valid standard SI unit symbols (g\text{g}g, mL\text{mL}mL) from prohibited non-standard abbreviations (gms\text{gms}gms, cc\text{cc}cc), or verifying mandatory *"inclusive of all taxes"* phrasing. Separating semantic extraction from deterministic validation guarantees zero generative hallucination on exact statutory predicates.
- **Auditability and Regulatory Readiness:** Unlike black-box AI screening tools, our framework links every pass/fail verdict to a versioned statutory predicate (e.g., Packaged Commodities Rule 6 declarations) and logs an immutable evidence record. Combined with confidence-based escalation to human inspectors for low-visibility labels, the system balances full automation speed with regulatory accountability.

Proposed Conceptual Architecture

To verify Legal Metrology declarations accurately across physical packaging, artwork proofs, and e-commerce product listings, we propose a closed-loop, seven-stage hybrid architecture that decouples probabilistic AI extraction from deterministic Policy-as-Code enforcement:

International Journal of Statistics and Mathematical Research Trends
E-ISSN: 3108-0871, Volume 2 Issue 5
Sep-Oct 2026, PP 01-7



- **Multi-Modal Input & Ingestion Layer:** Accepts heterogeneous declaration artifacts—such as product label photographs, packaging artwork PDFs, and structured e-commerce listing payloads (JSON/HTML/API feeds)—while attaching source metadata (SKU ID, channel type, timestamp, and input checksum).
- **AI Extraction & Normalization Layer:** Uses multi-pass Optical Character Recognition (OCR), layout/vision parsing, and Large Language Models (LLMs) to extract candidate declaration fields (e.g., manufacturer/packer/importer details, country of origin, generic commodity name, net quantity, manufacturing/import month and year, best-before/expiry date, Maximum Retail Price [MRP] with tax-inclusivity wording, unit sale price, and consumer care contact details) into a standardized intermediate schema with field-level confidence scores.
- **Policy-as-Code Rule Repository:** Stores statutory Legal Metrology mandates and periodic rule amendments as version-controlled, machine-readable rules and atomic predicates (e.g., standard International System [SI] unit validation, numeric tolerance/rounding checks, date ordering, mandatory field presence, and prohibited qualifier checks) decoupled from application code.
- **Deterministic Verification & Scoring Engine:** Evaluates normalized declaration attributes against active Policy-as-Code rules using deterministic logic (regular expressions, SI unit lookup tables, date parsers, and cross-field consistency checks between physical labels and online listings) to prevent generative AI hallucinations, while computing a weighted compliance score and flagging mandatory rule failures.
- **Explainability, Human-in-the-Loop & Audit Store:** Generates human-readable violation explanations mapped to exact statutory clauses, routes low-confidence extractions or borderline edge cases to human inspectors for review, and records immutable, timestamped compliance evidence artifacts for regulatory traceability.

Research Gaps and Future Directions

Table 2. Research gaps and future directions in automated Legal Metrology declaration compliance.

Gap supported by evidence	Current Research Gaps	Strategic Future Directions
Packaging & Enclosure Limitations	Curved surfaces, surface glare, low-contrast printing, and multi-script labels degrade OCR and visual extraction accuracy.	Integrate layout-aware Vision-Language Models (VLMs) and glare-resilient preprocessing to verify text, SI units, and display-panel proportionality.

Gap supported by evidence	Current Research Gaps	Strategic Future Directions
Rule Amendments & Traceability	Static rule engines and standalone LLMs struggle with frequent statutory updates and lack verifiable rule-to-evidence lineage.	Maintain version-controlled Policy-as-Code rules with immutable audit artifacts and NLP-assisted regulatory amendment tracking.
Multi-Modal & Channel Alignment	Existing tools inspect physical labels or online e-commerce listings in isolation without cross-checking consistency.	Reconcile physical packaging declarations with digital marketplace payloads and route low-confidence cases to human-in-the-loop review.

V. Conclusion

Verifying Legal Metrology declarations across physical packaging, print artwork, and digital product listings is essential for consumer protection and fair trade, yet manual inspection cannot keep pace with high-volume catalogs or frequent regulatory changes. In this paper, we presented an **Automated Compliance Checker for Legal Metrology Declarations** built on a hybrid AI and deterministic Policy-as-Code architecture. By decomposing statutory mandates into atomic, machine-verifiable conditions, pairing OCR/LLM-based field extraction and human-readable remediation explanations with deterministic backend rule validation, and recording timestamped compliance evidence with human-in-the-loop escalation, the framework balances extraction flexibility with exactness and auditability. Future work will expand layout-aware visual parsing for challenging packaging conditions, automate regulatory rule updates, and strengthen cross-channel verification between physical goods and online listings.

References

1. **Break Rules into Small Conditions:** Translate complex Legal Metrology rules (like MRP, net quantity, and manufacturer details) into small, testable "Yes/No" conditions.
   paper5+ 2
2. **Collect Machine-Readable Evidence:** Link every compliance check to clear digital evidence, such as extracted text, rule IDs, and pass/fail outputs.
   paper5+ 1
3. **Calculate a Compliance Score:** Measure how compliant a product label is by calculating the percentage of statutory conditions it satisfies.
   paper5+ 1
4. **Use AI for Reading & Explaining:** Use OCR and Large Language Models (LLMs) to read messy product labels and write simple explanations for any errors found.
   paper2+ 2
5. **Use Deterministic Code for Exact Checks:** Delegate strict checks—such as exact SI units (g\text{g}g, mL\text{mL}mL), numbers, and dates—to deterministic backend code so AI does not hallucinate or make math mistakes.
   paper2+ 2
6. **Manage Rules as Policy-as-Code:** Store Legal Metrology mandates as version-controlled rules so updates and amendments can be applied quickly without rewriting the whole system.
   paper3+ 2
7. **Keep Immutable Audit Logs:** Save every check with a timestamp, input hash, and rule version to create a tamper-proof audit trail for legal proof.
   paper3+ 1
8. **Track Compliance Trends & Readiness:** Monitor violation rates, check speed, and audit readiness across different product categories and sellers.
   paper4+ 3

9. **Include Human-in-the-Loop Review:** Automatically send blurry labels, low-confidence extractions, or borderline edge cases to a human inspector for final approval.
   paper2+ 2
10. **Check Before Market Release:** Verify labels and online listings in real time before products reach consumers to prevent fines, seizures, and customs delays.
    paper2+ 1