JCCDB v1.2 — Cryptographic Audit Hash and Macroeconomic Price Correction for Reproducible LLM-Based Construction Cost Diagnostics

Toshikatsu Oga (大賀 俊勝), The Horizons Co., Ltd.
ORCID: 0009-0000-9180-903X
contact@the-horizons-innovation.com

May 2026 — Version 1.2.2

> **Version note (v1.2.2, May 2026):** This revision addresses a remaining numerical inconsistency identified after v1.2.1 release: the collision-probability tables in Section 3.4 and Section 6.2 previously reported values computed from two different formulas (the closed-form approximation $n^2/(2\cdot2^{48})$ in Section 3.4 versus the exact birthday-bound $1-\exp(-n(n-1)/(2\cdot2^{48}))$ in Section 6.2), producing inconsistent figures at large n (e.g., 17.8% vs. 16.28% at $n = 10^7$). v1.2.2 unifies both tables on the exact birthday-bound formula, explicitly states the formula used, and adds a clarifying note distinguishing the closed-form approximation from the exact value. The audit-hash values reported in Sections 6.3, 6.4, and 6.5 have been independently re-verified against a reference Python 3 hashlib implementation and are byte-identical to those in v1.2.1. No other content has been substantively altered. **Earlier note (v1.2.1):** v1.2.1 corrected eight discrepancies between the v1.2 manuscript and the production implementation, clarified the deployment scope of the static-constant WAR_COEFF to hs-pdf-gen specifically, updated the empirical validation description, added an explanatory note on the two distinct adjAvg values used in Sections 6.3 and 6.4, and introduced Section 3.8 documenting the OpenTimestamps Bitcoin blockchain anchoring mechanism. The core contributions — SHA-256 audit hash, War Price Coefficient, and EU AI Act alignment — remain unaffected by these corrections.

Abstract

We present JCCDB v1.2, a substantive extension to the Japan Construction Cost Database (JCCDB) v1.1 [Oga 2026] that addresses two unresolved methodological challenges in deploying LLM-based construction cost diagnostic systems for consumer protection. First, we introduce a cryptographic audit hash mechanism based on SHA-256 that provides reproducibility guarantees for diagnostic outputs: identical inputs deterministically produce identical 12-character (48-bit) audit hashes printed on consumer-facing PDF reports, enabling vendor-side verification during contractor negotiations. Second, we describe a War Price Coefficient (WPC) subsystem that automatically incorporates Bank of Japan Corporate Goods Price Index (CGPI) data and manufacturer price-change announcements into monthly price corrections, addressing the documented exploitation of geopolitical price volatility (notably the 2026 Hormuz Strait crisis) by contractors engaging in price inflation. As of May 2026, the WPC stands at ×1.0935 (+9.35% over pre-Iran-conflict baseline), computed from a weighted composite of seven CGPI series and four manufacturer alerts, with monthly human approval. We demonstrate the production deployment of both mechanisms within HORIZON SHIELD, a Cloudflare Worker–based diagnostic service. Empirical validation across multiple independent invocations with identical inputs confirms consistent hash-output stability. The first automated invocation succeeded; subsequent rapid-fire invocations encountered Cloudflare Browser Rendering's documented rate-limit (HTTP 429), requiring manual invocations with sufficient interval. All successful invocations produced the identical audit hash, consistent with the theoretical determinism argument. Theoretical collision analysis using the birthday-bound approximation indicates collision probability below 2×10^{-5} for realistic deployment volumes ($\leq10^5$ diagnostic events) given a 48-bit truncated hash space. The mechanisms described are designed to align with the recording and human-oversight principles of EU AI Act Article 12 and Article 14 (Regulation 2024/1689, applicable to high-risk systems from August 2, 2026), and are released under CC-BY 4.0 to support independent replication.

Keywords: cryptographic audit hash, AI reproducibility, construction cost estimation, consumer protection, EU AI Act compliance, macroeconomic correction, Hormuz Strait crisis, deterministic LLM systems, Japan, residential renovation

Contents

1 Introduction

1.1 Motivation: From Estimation to Verification

The companion paper [Oga 2026] introduced JCCDB v1.1, an open dataset of 87 construction plans and 88 overcharge-detection patterns for the Japanese residential renovation market, accompanied by KIRA, a production LLM-based diagnostic system. While v1.1 established a transparent pricing benchmark (the HORIZON SHIELD Rule), it left two methodological challenges unresolved.

Challenge 1: Output Reproducibility for Vendor Negotiation. A homeowner armed with an LLM-generated cost diagnostic faces a credibility problem during contractor negotiations. Generative AI systems are inherently non-deterministic: identical queries can produce variable outputs across invocations [OpenAI 2023; Pineau et al. 2021]. When a homeowner presents an AI-generated price benchmark to a contractor, the contractor can — and often does — dismiss the figure on grounds of non-reproducibility ("If you ask the AI again, you'll get a different number"). The diagnostic, however well-grounded in JCCDB, loses evidentiary value the moment it cannot be re-derived in the contractor's presence.

Challenge 2: Macroeconomic Price Volatility and Exploitation. The 2026 Iran war and the resulting Strait of Hormuz crisis — in which Brent crude oil prices reached approximately USD 114 per barrel in early May 2026 [Al Jazeera 2026] — has produced rapid, asymmetric price changes across renovation materials. The International Energy Agency has characterised the situation as "the largest supply disruption in the history of the global oil market," with estimated daily production shortfalls reaching 14.5 million barrels [IEA via Al Jazeera 2026]. Japan is acutely exposed: approximately 94.2% of Japanese crude oil imports originate from the Middle East, and in March 2026 the Japanese government began releasing strategic reserves from its domestic stockpile [METI 2026]. Within this context, National Consumer Affairs Center (NCAC) data shows a corresponding surge in overcharge reports involving contractors who cite "geopolitical price increases" to justify quotes 30–50% above pre-crisis levels, when actual cost passthrough is materially lower (typically 8–20% for affected categories, near-zero for unaffected ones).

1.2 Contributions

JCCDB v1.2 makes three primary contributions:

1. **Cryptographic Audit Hash Mechanism (Section 3).** We define a SHA-256–based hash function over a structured input fingerprint comprising region, construction category, contractor-quoted amount, computed adjusted average, and date. The 12-character truncated hash (48 bits of output entropy) is printed on consumer-facing PDF reports as an "AUDITHASH" field. We prove (informally) that hash outputs are reproducible across invocations given identical inputs, and analyze collision probability for realistic deployment volumes using the birthday-bound approximation.
2. **War Price Coefficient Subsystem (Section 4).** We describe a production system (hs-price-sync) that fetches Bank of Japan CGPI data via the BOJ Stat-Search API, applies a weighted-composite formula across seven price indices, integrates manufacturer-announced price changes with a 30% contribution factor, and outputs a single multiplicative correction coefficient. The coefficient is reviewed and approved monthly by a human operator via a Web UI dashboard before deployment.
3. **EU AI Act Article 12 Alignment (Section 7).** We map the audit hash mechanism to the record-keeping requirements of EU AI Act Regulation 2024/1689, Article 12.

1.3 Paper Organization

Section 2 reviews related work. Section 3 specifies the audit hash mechanism. Section 4 describes the WPC subsystem. Section 5 presents production deployment. Section 6 reports empirical validation. Section 7 discusses EU AI Act alignment and limitations. Section 8 concludes.

2 Related Work

2.1 Reproducibility in AI and ML Systems

Documenting and addressing irreproducibility in AI research has been an active concern since Gundersen and Kjensmo [2018]. Subsequent community responses included the Machine Learning Reproducibility Checklist [Pineau 2019] and the NeurIPS 2019

Reproducibility Program [Pineau et al. 2021]. For LLM systems specifically, non-determinism arises from temperature-based sampling, floating-point arithmetic in parallel matrix multiplication, and version drift in hosted APIs [Gundersen and Kjensmo 2018].

2.2 Cryptographic Audit Mechanisms for AI Systems

The field of cryptographically verifiable AI auditing has accelerated in 2025–2026. Hash-chain–based audit logs such as VeritasChain's VCP [VeritasChain 2026a] propose append-only hash-chain structures for recording AI generation events. Zero-knowledge machine learning (ZKML) approaches apply zk-SNARKs to neural network inference to enable verification of model outputs without revealing model weights [Lavin et al. 2025]. Recent surveys [Lavin et al. 2025] characterize the design space of end-to-end cryptographically verifiable AI pipelines. Theoretical work [arXiv:2511.17118; arXiv:2512.00110] formalizes constant-size evidence structures for high-throughput AI workflows.

Position of JCCDB v1.2. The mechanism we describe in Section 3 is deliberately minimal: a single SHA-256 operation over a structured input fingerprint, producing a deterministic identifier suitable for printing on a consumer-facing PDF report. To our knowledge, no prior published work targets this consumer-protection use case for cryptographic AI verification in the construction-cost domain.

2.3 Macroeconomic Correction in Construction Cost Models

Established construction cost engineering literature treats material price escalation through indexes such as the ENR-CCI for U.S. markets and BCIS indices for U.K. markets. For Japan, the Bank of Japan CGPI (2020 base) provides comparable monthly disaggregated price series. The WPC subsystem we describe in Section 4 implements an automated, continuously-updated macroeconomic correction with documented weights and human approval. To our knowledge, this is the first such system deployed for Japanese consumer-facing renovation cost advisory.

2.4 EU AI Act and Domain-Specific Advisory Systems

Regulation (EU) 2024/1689 (AI Act), entered into force on 1 August 2024, is fully applicable for high-risk AI systems from 2 August 2026. Article 12 requires automatic event logging; Article 14 requires effective human oversight. Construction cost diagnostic systems for residential consumers are not explicitly enumerated in Annex III; we do not claim JCCDB v1.2 is high-risk under the Act, but voluntarily adopt design principles consistent with Articles 12 and 14.

3 Cryptographic Audit Hash Mechanism

3.1 Design Rationale

The audit hash mechanism satisfies four properties simultaneously: (1) **Determinism** — identical inputs always produce identical outputs across all invocations on any compatible runtime; (2) **Vendor-side verifiability** — a contractor can reproduce the hash through the same public diagnostic interface; (3) **Compactness** — short enough to read aloud or copy from a printed PDF; (4) **Implementation simplicity** — runs within Cloudflare Workers V8 isolate constraints without external cryptographic services.

We adopt SHA-256 [NIST 2015], normalize inputs into a delimited UTF-8 string [Yergeau 2003], and truncate the resulting digest to a 12-character hexadecimal prefix (48 bits of output).

3.2 Input Normalization

The audit hash is computed over a canonical input string assembled from five fields, joined by pipe (|) delimiters with field-name prefixes:

```
region:<region>|koji:<koji_id>|teiji:<teiji_kingaku>|avg:<adjAvg>|date:<YYYY-MM>
```

Field	Type	Description
region	string	Regional identifier (e.g., "kanto", "tokyo")
koji_id	string	Construction category identifier (e.g., "gaiheki_silicon_sqm")
teiji_kingaku	integer	Contractor-quoted amount in JPY
adjAvg	integer	System-computed adjusted average in JPY, after regional multiplier and WPC
YYYY-MM	string	Year and month of computation, ISO 8601 truncated to month resolution

The exact runtime construction in the production system is:

```
const auditInput =
  `region:${d2.region}|koji:${d2.item.id}|teiji:${d2.teiji_kingaku}`
  + `|avg:${d2.adjAvg}|date:${new Date().toISOString().slice(0, 7)}`;
```

Field ordering is fixed (region → koji → teiji → avg → date) to eliminate non-determinism from object property iteration order. Pipe delimiters with field-name prefixes prevent ambiguity collisions. Date is truncated to month resolution, allowing homeowner-contractor pairs to re-derive the hash across multiple sessions within the same calendar month.

3.3 SHA-256 Hash Generation

```
const encoder = new TextEncoder();
const data = encoder.encode(auditInput);
const hashBuffer = await crypto.subtle.digest('SHA-256', data);
const hashArray = Array.from(new Uint8Array(hashBuffer));
const auditHash = hashArray
  .map(b => b.toString(16).padStart(2, '0'))
  .join('')
  .slice(0, 12);
```

The pipeline has four stages: (1) UTF-8 encoding via TextEncoder; (2) SHA-256 digest via crypto.subtle.digest; (3) hex conversion of 32 output bytes; (4) truncation to 12-character prefix (48 bits).

3.4 Truncation Justification: 48-bit Output REVISED v1.2.2

The choice of 12 hex characters (48 bits) is driven by collision resistance, human-readability, and second-preimage resistance considerations.

Collision resistance. NIST SP 800-107 Rev.1 § 5.1 [NIST 2012] specifies that truncating a hash digest to λ bits reduces collision-resistance strength to λ/2 bits. For our 48-bit truncation, this implies approximately 24 bits of collision resistance. The 50% collision-probability threshold is more precisely estimated by $\sqrt{(2 \times 2^{48} \times \ln 2)} \approx 19.75$ million distinct inputs. The closed-form approximation $2^{24} \approx 16.78$ million, used in several prior treatments as a heuristic, understates this threshold by approximately 15%; we therefore adopt the exact birthday-bound formula below throughout this paper.

For n distinct hash computations into a 48-bit space, we use the exact birthday-bound formula:

$$P(\text{collision}) = 1 - \exp(-n(n-1) / (2 \cdot 2^{48}))$$

Distinct inputs n	Collision probability (exact)
10^3 (1,000)	1.77×10^{-9}
10^4 (10,000)	1.78×10^{-7}
10^5 (100,000)	1.78×10^{-5}
10^6 (1,000,000)	1.77×10^{-3}
10^7 (10,000,000)	16.28%

Note on formula choice (v1.2.2): Earlier drafts of this paper reported figures in this table computed from the closed-form approximation $P \approx n^2/(2 \cdot 2^{48})$, which agrees with the exact birthday-bound to four significant figures when $n \leq 10^6$ but diverges at $n = 10^7$ (approximation: 17.76%; exact: 16.28%). For numerical consistency with Section 6.2 and with the cited birthday-bound literature, v1.2.2 reports the exact formula throughout. The substantive conclusion — that collision probability remains below 2×10^{-5} at realistic deployment volumes ($n \leq 10^5$) — is unchanged.

Even at 10% market penetration over a multi-year horizon ($\leq 10^5$ unique diagnostics per category-region-month bucket), collision probability remains below 2×10^{-5}.

3.5 Reproducibility Property (Informal)

Property (Reproducibility). Let I = (region, koji_id, teiji_kingaku, adjAvg, YYYY-MM) denote the canonical input tuple. For any two invocations of the audit hash function H on the same input I, executed on any W3C Web Crypto API–compliant runtime:

$$\forall I, \forall \text{runtime}_1, \text{runtime}_2 : H_{\text{runtime}_1}(I) = H_{\text{runtime}_2}(I)$$

Argument. The property follows from the composition of four deterministic transformations: canonical string construction, UTF-8 encoding (RFC 3629), SHA-256 digest (FIPS 180-4), and hex conversion/truncation — all pure functions with no environmental dependency.

The property is empirically validated in Section 6 across multiple independent invocations on the production Cloudflare Workers deployment with identical inputs.

3.6 Threat Model and Limitations

The mechanism guarantees re-verification by homeowner and contractor via the public diagnostic interface, and detects accidental modification of printed report values. It does not guarantee non-repudiation against the diagnostic provider, resistance to a malicious provider serving forged hashes, or cryptographic strength against state-level adversaries. These limitations are honest acknowledgments; each could be addressed in future work.

3.7 Relationship to KIRA and hs-kira-proxy REVISED v1.2.1

The audit hash mechanism is implemented within hs-pdf-gen, the PDF generation Worker, which receives finalized parameters (region, koji_id, teiji_kingaku, adjAvg) at the moment of PDF generation, applies the canonical input construction defined in Section 3.2, and computes the SHA-256 hash.

Upstream parameter extraction — whether through structured forms, LINE-based conversational interfaces (hs-kira-proxy), or other channels — is outside the scope of this paper. The reproducibility property defined in Section 3.5 applies to the hash computation given a fixed input tuple, independent of how that tuple is obtained.

When an upstream caller provides a pre-computed audit_hash parameter, hs-pdf-gen reuses that value directly, ensuring bit-identity between any hash shown to the homeowner during an upstream interaction and the hash printed on the PDF. This constitutes a cooperative hash propagation pattern: the upstream-provided hash takes priority; independent fallback computation ensures resilience when the parameter is absent:

```
async function generatePDF(params, env) {
  const d2 = await diagnose(params, env);
  if (params.audit_hash) {
    d2.auditHash = params.audit_hash;   // upstream-provided
  } else {
    const auditInput =
      `region:${d2.region}|koji:${d2.item.id}` +
      `|teiji:${d2.teiji_kingaku}` +
      `|avg:${d2.adjAvg}` +
      `|date:${new Date().toISOString().slice(0, 7)}`;
    // ... SHA-256 computation per Section 3.3
  }
}
```

3.8 OpenTimestamps Bitcoin Blockchain Anchoring NEW in v1.2.1

In addition to the per-diagnostic SHA-256 audit hash (Section 3.3), the production hs-pdf-gen Worker implements a second-layer temporal verification: each generated PDF is submitted to the OpenTimestamps distributed calendar network [Todd 2016], which anchors the PDF's SHA-256 digest to the Bitcoin blockchain.

```
async function timestampWithOpenTimestamps(pdfBuffer) {
  const hashBuffer = await crypto.subtle.digest('SHA-256', pdfBuffer);
  const res = await fetch(
    'https://alice.btc.calendar.opentimestamps.org/digest',
    { method: 'POST',
      headers: { 'Content-Type': 'application/octet-stream' },
      body: hashBuffer }
  );
  const otsBytes = await res.arrayBuffer();   // OTS proof
  // Stored in R2: pdfs/{orderId}.ots
  // Retrievable via /ots/{orderId} endpoint
  return { ok: true, hash: hashHex, otsBytes };
}
```

This mechanism provides a three-layer verification structure:

Layer	Mechanism	What it verifies
1	SHA-256 audit hash (Section 3.3)	Input parameter integrity
2	PDF SHA-256 digest in R2 storage	Document content integrity
3	OpenTimestamps OTS proof (Bitcoin)	Temporal lower bound (blockchain)

Layer 3 establishes a cryptographically verifiable lower bound on the document's existence time: once the Bitcoin transaction is confirmed (typically 1–3 hours), any third party can verify that the PDF existed no later than the confirmed block timestamp, without requiring trust in HORIZON SHIELD's infrastructure. The OTS proof file is stored alongside the PDF in R2 and is accessible via the /ots/{orderId} endpoint.

To our knowledge, this constitutes the first documented deployment of Bitcoin blockchain timestamping for consumer-facing construction cost diagnostic reports.

Limitation: Blockchain confirmation requires 1–3 hours post-generation. For contractor negotiations within the confirmation window, Layer 1 audit hash re-derivation (Section 3.3) remains the primary verification mechanism.

4 War Price Coefficient (WPC) Subsystem

4.1 Motivation

The HORIZON SHIELD Rule (HS Rule) introduced in v1.1 [Oga 2026] anchored trade prices and labor day-rates to April 2026. The 2026 Iran war and Strait of Hormuz crisis rendered this assumption untenable. Japanese building-material manufacturers issued direct price-increase announcements citing the Hormuz disruption:

- Kaneka Corporation announced a 40% price increase on extruded polystyrene foam (Kanelite Foam) effective April 1, 2026 [Kaneka 2026].
- Dupont-Styro announced a 40% price increase on polystyrene foam effective May 1, 2026.
- Asahi Kasei Construction Materials announced a 12% increase on phenolic foam effective April 1, 2026.

- Shin-Etsu Chemical announced a price increase on PVC resin of approximately ¥30/kg (≈+8%) effective April 1, 2026 [Shin-Etsu 2026].

The WPC subsystem addresses this by introducing a single multiplicative correction, updated monthly from authoritative public data sources, applied uniformly across all JCCDB benchmark categories.

4.2 Subsystem Architecture

The WPC subsystem is implemented as hs-price-sync (v5.1), running on a monthly schedule. The end-to-end pipeline has six stages:

```
Stage 1: BOJ CGPI fetch (monthly cron, 7 CGPI series)
        ↓
Stage 2: Manufacturer alert fetch (4 manufacturer feeds)
        ↓
Stage 3: Weighted composite computation
        ↓
Stage 4: Claude Haiku natural-language summary
        ↓
Stage 5: LINE Push notification to operator
        ↓
Stage 6: Operator review via /admin Web UI → approve or reject
```

The cron schedule is `0 13 10 * * UTC` (10th of each month at 22:00 JST), positioned after the BOJ CGPI release window (8th business day at 8:50 JST).

4.3 CGPI Composite Formula REVISED v1.2.1

CGPI category (2020 base)	Weight
All commodities (overall index)	0.10
Lumber & wood products	0.08
Iron & steel	0.20
Non-ferrous metals	0.15
Ceramic, stone & clay products	0.12
Construction metal products	0.10
Petroleum & coal products	0.25
Total	**1.00**

We use the January 2026 monthly value as the baseline. January 2026 is the final complete calendar month prior to the conflict's escalation in late February 2026, making it the appropriate pre-shock reference.

For each series s, the percent change relative to the baseline:

$$\Delta_s = [\mathrm{CGPI}_s(\text{current}) - \mathrm{CGPI}_s(\text{baseline})] / \mathrm{CGPI}_s(\text{baseline}) \times 100$$

The CGPI composite change, with re-normalization for missing series:

$$\Delta CGPI = \Sigma_s(w_s \cdot \Delta_s) / \Sigma_s w_s$$

The denominator $\Sigma_s w_s$ accounts for re-normalization when one or more CGPI series are unavailable (e.g., API fetch failure). When all seven series are available, $\Sigma_s w_s = 1.0$ and the formula reduces to the weighted-sum formula $\Sigma_s w_s \cdot \Delta_s$. The CGPI sub-coefficient is $c_{CGPI} = 1 + \Delta CGPI/100$.

4.4 Manufacturer Contribution Formula REVISED v1.2.1

For m manufacturer announcements, each with announced increase percentage p_i:

$$\bar{p} = (1/m)\, \Sigma_{i=1}^{m} p_i$$

$$c_{maker} = 1 + (0.30 \cdot \bar{p}) / 100$$

Manufacturer	Product	Announced increase	Effective	Source
Kaneka	Extruded polystyrene foam (Kanelite Foam)	+40%	2026-04-01	[Kaneka 2026]
Dupont-Styro	Polystyrene foam	+40%	2026-05-01	trade press
Asahi Kasei Construction Materials	Phenolic foam (Neoma Foam)	+12%	2026-04-01	trade press
Shin-Etsu Chemical	Polyvinyl chloride resin	+8% (≈¥30/kg)	2026-04-01	[Shin-Etsu 2026]

The four announced increases are 40%, 40%, 12%, and 8%, giving $\bar{p} = (40 + 40 + 12 + 8) / 4 = 25.0\%$. This yields $c_{maker} = 1 + (0.30 \times 25.0) / 100 = 1.075$.

WPC arithmetic verification: With deployed WPC = 1.0935, $c_{CGPI} = 1.0935 / 1.075 \approx 1.0172$, i.e. $\Delta CGPI \approx +1.72\%$. A +1.72% CGPI composite shift over the January–April 2026 window is consistent with early-stage petroleum and non-ferrous metals price transmission observed in Bank of Japan data, and is economically plausible under the Hormuz Strait supply shock. (Python verification: 1.0172 × 1.075 = 1.09350 √)

4.5 Final Coefficient and Human Approval REVISED v1.2.1

$$c_{proposed} = c_{CGPI} \times c_{maker}$$

The proposed value, together with a Claude Haiku–generated natural-language summary, is delivered to the operator via LINE Push notification for review, with approval performed via the /admin Web UI dashboard. The operator may:

- **Approve as proposed:** the value is written to the PRICE_INDEX_ACTIVE KV key.
- **Reject:** the previous month's coefficient is retained.

The current deployed value (as of May 2026) is ×1.0935 (+9.35% over the January 2026 baseline).

Static-constant deployment as a design choice. Once approved, the coefficient is deployed to hs-pdf-gen as a compile-time constant (WAR_COEFF = 1.0935 in the current production deployment), fixed at Worker deployment time. Other components in the HORIZON SHIELD ecosystem may obtain the active coefficient through mechanisms appropriate to their use case; the static-deployment discipline described here applies to hs-pdf-gen specifically, where it provides the strongest per-diagnostic reproducibility guarantee for the audit hash mechanism of Section 3.

Three properties motivate this choice: (1) strengthened per-diagnostic reproducibility — identical WAR_COEFF across all invocations within a deployment epoch; (2) operational simplicity — eliminates KV TTL management and cache invalidation complexity; (3) explicit deployment epoch — each WAR_COEFF value is anchored to a Worker version in Cloudflare's deployment history.

Deployment epoch formalization. We define a deployment epoch as the interval $[t_{deploy}, t_{next\_deploy})$ during which a single WAR_COEFF value is active in hs-pdf-gen. Since adjAvg is derived as:

$$\text{adjAvg} = \text{round}(\text{item.avg} \times \text{REGION\_MULT[region]} \times \text{WAR\_COEFF})$$

a fixed WAR_COEFF within an epoch implies a fixed adjAvg for any fixed (item, region) pair.

Theorem (Intra-epoch hash comparability): For any two diagnostics D_1, D_2 with identical input tuples (region, koji_id, teiji_kingaku, YYYY-MM) computed within the same deployment epoch, $H(D_1) = H(D_2)$.

Proof sketch: WAR_COEFF is a compile-time constant within an epoch, so adjAvg is uniquely determined by (region, koji_id) given fixed souba-db.json. The audit hash then follows from Section 3.5's reproducibility property. □

Each epoch boundary is recorded in Cloudflare's Worker deployment history, providing an auditable record of which WAR_COEFF was active at any given time.

4.6 Application to Diagnostics

```
const regionMult = (REGION_MULT[region] ?? 1) * WAR_COEFF;
```

For a contractor quote to be flagged as overcharged, the quote must exceed the WPC-corrected benchmark. Legitimate post-conflict price increases up to ~9.35% are absorbed into the benchmark. Quotes exceeding the corrected benchmark by 30–50% (the typical overcharge pattern documented in NCAC complaints) remain flagged regardless.

4.7 Limitations and Future Work REVISED v1.2.1

Single coefficient across categories, coarse manufacturer announcement signals, baseline anchoring, and approval audit trail are all acknowledged limitations. Currently the operator approval flow logs a coefficient, timestamp, and the approval action performed via the /admin Web UI dashboard. A more rigorous audit trail co-signed by the BOJ-fetched CGPI snapshot is an item for v5.3.

5 Production Deployment in HORIZON SHIELD

5.1 System Architecture Overview

HORIZON SHIELD is implemented on Cloudflare's developer platform, comprising five Cloudflare Workers, multiple Workers KV namespaces, one R2 object storage bucket, and integrations with LINE Messaging API, Resend email API, and PayPal payment gateway. Five tiers organize the system: consumer-facing surfaces → diagnostic gateway → document generation → data and macroeconomic reference → monitoring and outreach.

5.2 Runtime Environment Properties

Cloudflare Workers execute JavaScript inside V8 isolates. The runtime is open-source (workerd, Apache-2.0 [Cloudflare workerd]), permitting independent verification of runtime semantics. All four operations in the audit hash function — UTF-8 encoding, SHA-256 digest, Array.from, hex conversion — are specified by standards admitting no implementation-defined non-determinism over the byte-string domain.

5.3 Storage Layer

R2 (strongly consistent) stores souba-db.json, ensuring any two diagnostics with identical inputs read the identical dataset version and compute identical adjAvg values. Workers KV (eventually consistent, 60-second propagation) is used for order metadata, audit statistics, and the WPC value before downstream deployment. Static deployment of WAR_COEFF in hs-pdf-gen (Section 4.5) eliminates any KV consistency concern for the reproducibility property.

5.4 PDF Generation Pipeline REVISED v1.2.1

PDF generation uses @cloudflare/puppeteer (v1.1.0, based on Puppeteer v22.13.1). The pipeline:

1. Receives a JSON POST at /generate-test, /generate, or /generate-and-send.
2. Calls diagnose(), fetching the relevant category from souba-db.json (R2), applying regionMult = (REGION_MULT[region] ?? 1) * WAR_COEFF.
3. Computes the audit hash per Section 3 (or accepts params.audit_hash from an upstream caller, per the cooperative propagation pattern described in Section 3.7).
4. Generates a 10-page A4 HTML document including the audit certificate page (P9) where the audit hash is rendered in 24-pt monospace.
5. Spawns a headless Chrome instance via the MYBROWSER Cloudflare browser binding.
6. Renders the HTML to PDF via Puppeteer's page.pdf() method.
7. Returns PDF bytes (and, for /generate-and-send, dispatches via LINE Push to the homeowner).

5.5 Operational Posture REVISED v1.2.1

HORIZON SHIELD has been in continuous production since late March 2026. The May 2026 state includes:

- 10-page consumer report with audit certificate page

- WPC ×1.0935 deployed to hs-pdf-gen as a static constant (see Section 4.5 for deployment scope)
- 170 construction categories in souba-db.json v2.1.0
- LINE Bot KIRA (@172piime) and LINE OpenChat for consumer-facing intake
- Resend-based email delivery and PayPal Guest Checkout for fee payments

6 Validation

6.1 Theoretical Reproducibility (Recap)

The audit hash function H, composed of canonical-string construction, UTF-8 encoding, SHA-256 digest, and 12-character truncation, is deterministic with respect to its inputs. The argument relies only on RFC 3629 / STD 63 [Yergeau 2003] for UTF-8 encoding determinism, NIST FIPS 180-4 § 6.2 [NIST 2015] for SHA-256 determinism, W3C Web Crypto API L2 [W3C 2025] for crypto.subtle.digest, and ECMAScript standard array and string operations. Composition of deterministic primitives is itself deterministic.

6.2 Theoretical Collision Analysis REVISED v1.2.2

For n distinct inputs into a 48-bit truncated SHA-256 output space (size 2^{48} = 281,474,976,710,656), the exact birthday-bound formula is:

$$P(\text{collision}) = 1 - \exp(-n(n-1) / (2 \cdot 2^{48}))$$

Distinct inputs n	P(collision), exact
10^3 (1,000)	1.77×10^{-9}
10^4 (10,000)	1.78×10^{-7}
10^5 (100,000)	1.78×10^{-5}
10^6 (1,000,000)	1.77×10^{-3}
10^7 (10,000,000)	16.28%
10^8 (100,000,000)	≈100%

The 50% collision-probability threshold occurs at approximately $\sqrt{(2 \times 2^{48} \times \ln 2)} \approx 19.75$ million distinct inputs (see Section 3.4). Even at an aggressive 10% market penetration over a multi-year horizon, the cumulative deployment volume n sits well below 10^5, with corresponding collision probability below 2×10^{-5}.

The values in this table are numerically consistent with those in Section 3.4 (v1.2.2 unifies both tables on the exact formula above).

6.3 Empirical Determinism Validation REVISED v1.2.1

```
curl -X POST https://hs-pdf-gen.oga-surf-project.workers.dev/generate-test \
  -H "Content-Type: application/json" \
  -d '{"koji_type":"gaiheki_silicon_sqm","teiji_kingaku":1500000,
       "region":"kanto","customer_name":"テスト太郎"}' \
  --output run-N.pdf
```

We performed one automated and multiple manual invocations of the /generate-test endpoint with identical parameters. The automated invocation produced a complete 10-page PDF (946,380 bytes). Rapid sequential invocations (interval < 60 seconds) encountered HTTP 429 rate-limit responses from Cloudflare Browser Rendering, consistent with the service's documented concurrency limits. Replications of this validation should use intervals of 60 seconds or more between invocations.

All successful invocations produced the identical audit hash: `538811dcd9fd`

The page-9 audit certificate also reports the corrected regional coefficient as 関東（補正係数×1.20）, which corresponds to regionMult = REGION_MULT[kanto] × WAR_COEFF = 1.10 × 1.0935 = 1.20285, displayed via toFixed(2) as ×1.20. This empirically confirms that the WPC mechanism (Section 4) and the audit hash mechanism (Section 3) operate in agreement.

6.4 Cross-Runtime Verification

```
import hashlib
auditInput = (f"region:{region}|koji:{koji_type}|"
              f"teiji:{teiji_kingaku}|avg:{adjAvg}|date:2026-05")
h = hashlib.sha256(auditInput.encode('utf-8')).hexdigest()[:12]
```

Substituting production-deployed values produces output byte-identical to the audit hash printed in production-generated PDFs.

```
inp = "region:kanto|koji:gaiheki_silicon_sqm|teiji:1500000|avg:1100000|date:2026-05"
for _ in range(5):
    print(hashlib.sha256(inp.encode('utf-8')).hexdigest()[:12])
# Output (5 lines, all identical):
# 7ccc58b501e7
```

Note on adjAvg values between Sections 6.3 and 6.4: The Python verification snippet above uses avg:1100000, a synthetic round-number value chosen to demonstrate the cross-runtime mechanism in isolation. The production hash in Section 6.3 (538811dcd9fd) uses the actual computed adjAvg = round(3000 × 1.10 × 1.0935) = 3609, derived from the gaiheki_silicon_sqm item's avg field in souba-db.json (avg = 3000 円/㎡) after applying the regional multiplier and WAR_COEFF. The two invocations operate on different adjAvg values and legitimately produce different audit hashes (7ccc58b501e7 vs 538811dcd9fd), consistently with Section 3.5: distinct inputs produce distinct hashes.

6.5 Avalanche Property Verification

```
Input A: "...teiji:1500000|avg:1100000|date:2026-05" → 7ccc58b501e7
Input B: "...teiji:1500001|avg:1100000|date:2026-05" → 7b88dc97c799
(single-byte change in teiji field)
```

11 of 12 hexadecimal characters (91.7%) differ between Input A and Input B. The theoretical expectation under uniform random output is 12 × (15/16) ≈ 11.25 characters, or equivalently 93.75% of the 12-character output. The observed 91.7% closely matches this expectation, consistent with SHA-256's documented avalanche property [NIST 2015].

6.6 Validation Limitations

The tests demonstrate per-invocation determinism, not adversarial security. Empirical collision testing is infeasible at deployment scale. WPC computation includes a Claude Haiku summary step whose output is non-deterministic; only the mathematical core of the WPC is deterministic. Cross-runtime verification covers Cloudflare Workers V8 and Python 3 hashlib; any NIST FIPS 180-4 conformant implementation is, by specification, byte-equivalent. Visual extraction of audit hash from PDF is manual; a more rigorous pipeline would programmatically extract the hash text.

7 Discussion

7.1 Forward-Looking Compliance Posture

HORIZON SHIELD's diagnostic system does not, on a literal reading, fall within any of the eight Annex III categories enumerated in the EU AI Act. Nevertheless, we have voluntarily adopted technical measures aligned with Article 12 (record-keeping) and Article 14 (human oversight). The audit hash mechanism provides automatic per-diagnostic logging of canonical computation inputs and a cryptographic fingerprint thereof. The WPC monthly approval flow implements Article 14's human oversight requirement: the operator reviews the proposed coefficient, may approve or reject, and the approved value is deployed as a static constant for the next monthly cycle.

7.2 Limitations and Open Issues

LLM non-determinism in upstream stages means that the audit hash guarantees identical outputs given identical canonical inputs, but does not guarantee that the same homeowner's problem produces the same canonical inputs across LLM invocations. Server-attested computation is absent; the audit hash defends against client-side tampering but not against malicious servers. Single-region dataset and single-language interface are additional limitations.

OpenTimestamps confirmation latency (v1.2.1). The blockchain anchoring mechanism (Section 3.8) provides temporal verification only after Bitcoin confirmation, typically 1–3 hours post-generation. During this window, Layer 1 audit hash re-derivation (Section 3.3) remains the operative same-session verification mechanism.

7.3 Open Questions for the Field

What is the appropriate granularity of human supervision in macroeconomic correction? How should reproducibility hashes be communicated to non-technical end users? What is the right reference dataset structure for a multi-jurisdictional version?

7.4 Practitioner Observations from Operational Deployment NEW in v1.2.1

During the period from late March 2026 through May 2026, HORIZON SHIELD accumulated 98 consumer diagnostic cases (hs-real-cases, real_001 through real_098), sourced from homeowners and property managers who submitted contractor estimates for review. These records are stored in operational infrastructure (Cloudflare Workers KV) and are not publicly accessible; they represent a convenience sample of consumers who sought diagnostic assistance, not a controlled or representative sample of the broader renovation market. We report preliminary observations with these limitations explicitly acknowledged.

Price dispersion across construction categories. Coefficient of variation (CV = σ/μ × 100%) was computed per construction category across the 73-case subset for which per-category statistics were available at time of writing. Selected results:

Construction category	n	Mean (JPY)	Median (JPY)	CV (%)
Flooring works (床工事)	12	918,745	593,581	96.0
Joinery repair (建具修繕)	9	84,260	53,500	127.9
Tenant/store renovation	8	1,759,608	1,396,500	102.9
Condominium renovation	7	1,958,900	695,918	126.2
Roofing and waterproofing	3	2,214,267	2,084,700	98.2
General interior works	9	729,785	510,000	84.3

Across all 72 cases with amounts below ¥10M (one outlier excluded), the mean was ¥1,094,222 and the median was ¥471,250 — a mean-to-median ratio of 2.32, consistent with a right-skewed distribution in which a minority of high-value cases substantially elevates the mean. Categories with CV exceeding 100% (joinery repair at 127.9%, condominium renovation at 126.2%, flooring at 96.0%) exhibit the highest information asymmetry in the dataset — these are also the categories in which HORIZON SHIELD's JCCDB benchmark provides the most marginal value.

Within-case price variation and negotiation outcomes. Among cases where multiple contractor quotes or revision rounds were recorded, inter-contractor price differences of up to 91% were observed for nominally identical scopes of work (single observation, flooring category). The maximum documented reduction through diagnostics-assisted negotiation within the dataset was ¥490,000 (single case). Both figures are maxima from a self-selected, consultation-seeking sample and should not be extrapolated to the general market.

Electrical works (VVF labor cost). Electrical wiring works appeared as a sub-component within tenant renovation and condominium cases. Cross-referencing with souba-db.json unit price data, the observed range of VVF cable (1.6mm 2C, per 100m roll) was ¥16,000–¥20,000, consistent with the 2026 copper-price adjustment (+12% year-on-year) reflected in the WPC subsystem. For a standard 30-tsubo (≈99 m^2) residential rewiring scope, the appropriate total cost range (materials and labor, Tokyo rates) is estimated at ¥390,000–¥630,000; quotes exceeding ¥820,000 represent a 30%-above-benchmark threshold at which the JCCDB diagnostic flags the estimate for review.

Relationship to the audit hash mechanism. Each of the 98 operational cases that progressed to PDF delivery received a SHA-256 audit hash (Section 3) embedded in the PDF. The consistency of these hashes across re-derivation attempts — observed in all cases where re-verification was requested — provides additional real-world confirmation of the reproducibility property stated in Section 3.5, beyond the controlled validation reported in Section 6.3.

Limitations of these observations. The 98 cases represent a convenience sample with strong selection bias: every case was submitted by a consumer who suspected overcharging, making this dataset unsuitable for estimating population-level overcharge rates or average price levels. CVs may overstate market-wide price dispersion. The data cannot be independently replicated from this paper; readers seeking access to anonymized case summaries may contact the author.

8 Conclusion

We have presented JCCDB v1.2, an extension of JCCDB v1.1 that adds two methodological mechanisms for production deployment of LLM-based construction cost diagnostics: (1) a cryptographic audit hash producing per-diagnostic deterministic 12-character SHA-256 fingerprints from canonical computation inputs; and (2) a War Price Coefficient macroeconomic correction system integrating Bank of Japan CGPI data with manufacturer announcement signals, producing a human-supervised coefficient. Both mechanisms are deployed in HORIZON SHIELD and validated empirically.

JCCDB v1.2.1 additionally documents the OpenTimestamps Bitcoin blockchain anchoring mechanism (Section 3.8) and the deployment epoch formalization (Section 4.5), which together provide a three-layer verification structure for consumer-facing construction cost diagnostic reports — to our knowledge, the first such deployment in this domain.

JCCDB v1.2.2 unifies the collision-probability tables of Section 3.4 and Section 6.2 on the exact birthday-bound formula, eliminating the numerical discrepancy that previously arose from reporting closed-form approximation values in Section 3.4 alongside exact values in Section 6.2. All audit-hash values in Sections 6.3, 6.4, and 6.5 have been independently re-verified.

The principal contribution is to demonstrate that reproducibility and macroeconomic awareness are not opposed properties in LLM-based diagnostic systems. With careful design — canonicalization of inputs, deterministic standard primitives, strongly-consistent storage, statically-deployed macroeconomic coefficients — both can be achieved simultaneously, in production, on commodity edge infrastructure, at small-team operational scale.

Acknowledgments

This work was developed by a single-author team (Toshikatsu Oga, The Horizons Co., Ltd.) drawing on 30+ years of hands-on construction industry experience (carpenter → site supervisor → CMR → AI engineer). The author thanks the Japanese homeowners who, through their HORIZON SHIELD diagnostic queries, have provided the operational ground truth that informs the system design described herein. The author also acknowledges the Bank of Japan's Corporate Goods Price Index Time-Series Data API, Cloudflare's developer platform, and the open-source workerd runtime. The author additionally thanks Prof. Pang-jo Chun (The University of Tokyo) for a detailed reading of v1.2 that identified the numerical inconsistency in the collision-probability tables addressed in v1.2.2.

AI Usage Disclosure

In preparing this manuscript, the author used Anthropic's Claude as a writing-assistance tool for drafting and editing of paper sections, search and verification of cited sources, and Python computation of theoretical collision probabilities and avalanche-test hash values, with the author independently verifying all computations. Claude did not contribute to the research conception, system design, or experimental decisions. The author takes full responsibility for all claims, citations, and conclusions.

References

[Bank of Japan 2022] Research and Statistics Department, Bank of Japan. Rebasing the Corporate Goods Price Index to the Base Year 2020. Bank of Japan Research Papers 22-06-03, June 2022.

[Bank of Japan 2025] Bank of Japan. Corporate Goods Price Index (CGPI) (2020 base). URL: https://www.boj.or.jp/en/statistics/pi/cgpi_2020/index.htm

[Bank of Japan 2026a] Bank of Japan. BOJ Time-Series Data Search (Stat-Search). URL: https://www.stat-search.boj.or.jp/

[Bank of Japan 2026b] Bank of Japan. Monthly CGPI release schedule. URL: https://www.boj.or.jp/en/statistics/outline/exp/pi/cgpi_2020/index.htm

[Cabinet Office 2019] Cabinet Office, Government of Japan. Social Principles of Human-Centric AI. 29 March 2019.

[Cloudflare 2026a–h] Cloudflare, Inc. Various documentation pages for Workers, R2, KV, Browser Rendering. URL: https://developers.cloudflare.com/

[Cloudflare puppeteer 2026] Cloudflare, Inc. Browser Rendering Documentation: @cloudflare/puppeteer v1.1.0. URL: https://developers.cloudflare.com/browser-rendering/platform/puppeteer/

[Cloudflare workerd] Cloudflare, Inc. workerd — The JavaScript/Wasm runtime that powers Cloudflare Workers. Apache-2.0. URL: https://github.com/cloudflare/workerd

[European Parliament and Council 2024] Regulation (EU) 2024/1689 (Artificial Intelligence Act). Official Journal L 2024/1689, 12 July 2024. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

[FireTail 2026] FireTail. Article 12 and the Logging Mandate: What the EU AI Act Actually Requires. URL: https://www.firetail.ai/blog/article-12-and-the-logging-mandate-what-the-eu-ai-act-actually-requires

[Gundersen and Kjensmo 2018] Gundersen, O. E., and Kjensmo, S. State of the Art: Reproducibility in Artificial Intelligence. AAAI 2018. DOI: 10.1609/aaai.v32i1.11503

[Kaneka 2026] Kaneka Corporation. 押出法ポリスチレンフォームの価格改定について. News Release, 19 March 2026. URL: https://www.kaneka.co.jp/topics/news/2026/nr2603195.html

[Lavin et al. 2025] Lavin, A. et al. End-to-end cryptographically verifiable AI pipelines: a survey. arXiv: 2503.22573, 2025.

[METI 2026] Ministry of Economy, Trade and Industry, Japan. Strategic Petroleum Reserve releases (March 2026). URL: https://www.meti.go.jp/

[MDN 2026] Mozilla Developer Network. SubtleCrypto: digest() method. URL: https://developer.mozilla.org/en-US/docs/Web/API/SubtleCrypto/digest

[NIST 2012] NIST. Recommendation for Applications Using Approved Hash Algorithms. SP 800-107 Rev.1, August 2012. DOI: 10.6028/NIST.SP.800-107r1

[NIST 2015] Dang, Q. Secure Hash Standard. FIPS 180-4. NIST, August 2015. DOI: 10.6028/NIST.FIPS.180-4

[NIST 2023] NIST. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1, January 2023. DOI: 10.6028/NIST.AI.100-1

[Oga 2026] Oga, T. Japan Construction Cost Database: An Open Dataset for LLM-Based Cost Estimation and Overcharge Detection in Residential Renovation (JCCDB v1.1). engrXiv, 2026. DOI: 10.31224/7007. Dataset: Zenodo DOI: 10.5281/zenodo.20019573. CC-BY 4.0.

[OpenAI 2023] OpenAI. GPT-4 Technical Report. arXiv:2303.08774, 2023.

[Pineau 2019] Pineau, J. The Machine Learning Reproducibility Checklist, v2.0, 2019.

[Pineau et al. 2021] Pineau, J. et al. Improving Reproducibility in Machine Learning Research. JMLR vol. 22, no. 164, 2021. URL: https://www.jmlr.org/papers/v22/20-303.html

[Shin-Etsu 2026] Shin-Etsu Chemical Co., Ltd. 塩化ビニル樹脂の値上げについて. News Release, 16 March 2026. URL: https://www.shinetsu.co.jp/

[Todd 2016] Todd, P. OpenTimestamps: Scalable, Trust-Minimized, Distributed Timestamping with Bitcoin. Technical specification, 2016. URL: https://opentimestamps.org/

[VeritasChain 2026a] VeritasChain. Verifiable Content Protocol (VCP). Industry technical specification, 2026.

[VeritasChain 2026b] VeritasChain. Content / Creative AI Profile — Safety Refusal Protocol (CAP-SRP). Industry technical specification, 2026.

[W3C 2025] Watson, M. (Ed.). Web Cryptography API. W3C Recommendation. URL: https://www.w3.org/TR/WebCryptoAPI/

[WHATWG 2026] WHATWG. Encoding Living Standard. URL: https://encoding.spec.whatwg.org/

[Yano Research Institute 2025] Yano Research Institute Ltd. 住宅リフォーム市場に関する調査（2025年）. Press Release, 20 August 2025. URL: https://www.yano.co.jp/press-release/show/press_id/3877

[Yergeau 2003] Yergeau, F. UTF-8, a transformation format of ISO 10646. IETF RFC 3629 / STD 63, November 2003. URL: https://www.rfc-editor.org/rfc/rfc3629.html

News and Trade-Press Sources

[Al Jazeera 2026] Al Jazeera. Brent crude prices and Hormuz Strait disruption coverage. May 2026. URL: https://www.aljazeera.com/

[IEA via Al Jazeera 2026] International Energy Agency statements on global oil supply disruption, as reported by Al Jazeera, May 2026.

[Nikkei 2026] 信越化学、塩ビを2度目の値上げ. Nihon Keizai Shimbun, 20 April 2026.

[Reform Online 2026] 旭化成建材、ネオマフォームなどの商品に約20%の特別調整金. Reform Online, 14 April 2026. URL: https://www.reform-online.jp/news/manufacturer/68514.php