PITON Global

EXECUTIVE WHITE PAPER · WP-10

DATA ENTRY OUTSOURCING · PHILIPPINES

The clean-record standard: the executive economics of data entry outsourcing to the Philippines.

An analysis of why records keyed is a volume vanity metric, how accuracy at volume and fields correct at the source — never throughput — decide the true cost of a data operation once every downstream correction is counted, and the vendor-selection discipline that keeps the database trustworthy.

AUTHORS

John Maczynski — Chief Executive Officer
Ralf Ellspermann — Chief Strategy Officer

AUGUST 2026
MANILA · ESTABLISHED 2001

Contents

ABOUT THIS SERIES — VOLUME 10

Every data-entry proposal is priced per thousand records, because volume is the easiest thing to meter. This volume is about the only thing that matters once the records are in the system: whether every field is right, because a database no one trusts is worse than no database at all. The gap between keying records and producing clean ones is where the real cost hides. The full series lives at piton-global.com.

THE PITON-GLOBAL LEADERSHIP TEAM

About the authors

John Maczynski

Chief Executive Officer

John has bought data-entry and data-operations capacity for three decades, and learned that a wrong record is more expensive than a missing one — because the business acts on it before anyone knows it is wrong. His standing question ignores the throughput report: of the records you keyed, how many were right in every field the first time? Applied at PITON-Global, that question is what clients get on every data sourcing decision — personally, with no account layer in between.

j.maczynski@piton-global.com
+1 402-598-8740

Ralf Ellspermann

Chief Strategy Officer

Ralf has spent twenty-five years running Philippine data-operations floors, and he judges a data team by its field-level error rate under load, not its records-per-hour: what validation runs at the keystroke, and whether accuracy holds when volume spikes. The data contract in Section 3 codifies what he learned rebuilding teams that keyed fast and populated databases with errors. He now reads data providers the way he once ran them — from the field-error rate backward.

r.ellspermann@piton-global.com
Manila, Philippines

PITON Global

ABOUT PITON-GLOBAL

PITON-Global is a buyer-side advisory practice, Manila-based since 2001, that has never operated a delivery floor or taken a placement commission. Its stock-in-trade is knowing the Philippine provider market at the operating level — 1,000+ mapped operations, re-verified continuously, including each data team's real field-level accuracy under load rather than its records-per-hour headline — and converting that knowledge into a shortlist of six to ten genuinely qualified partners per engagement. Buyers from Tripadvisor, Garmin, TSYS, HealthPartners, Yetipay, Workrise, and Norman Disney & Young have run their searches through the practice.

SECTION 01

Executive summary

-62%
Fully-loaded cost per clean record vs. US onshore

99.95%+
Field-level accuracy on vetted teams, held under volume

40×
Cost multiple of a bad record caught downstream vs. at the keystroke

6.2×
First-year ROI, reconstructed case (Section 6)

A data-entry operation is bought on the price per thousand records and paid for on the records it gets wrong. The vendor reports a million records keyed at a fraction of the onshore cost; the business finds the wrong addresses that bounced the shipments, the mistyped amounts that skewed the report, and the duplicate customers that corrupted the analysis — each one acted on before anyone knew it was wrong. The operation hit its volume number and the business inherited a database it can no longer trust. The gap between keying records and producing clean ones is where the real cost of cheap data entry hides, because a wrong record is more expensive than a missing one: it is used before it is caught.

The Philippines is where data entry is delivered with the most accuracy discipline, because a deep, precise, detail-cultured talent pool can hold field-level accuracy at volume through the validation an accurate database demands. Five findings:

1 **Records keyed is the industry's most flattering number.** Any team can hit volume. Vetted teams are measured on the clean record: field-level accuracy, held under load, verified at the source. Weak providers quote a low per-record price the business repays in downstream corrections. Section 2 shows the gap.

2 **The double-key validation gate is the product.** Accuracy at volume comes from keying critical fields twice and reconciling the difference, plus rule-based validation at the keystroke — not from spot-checking a sample after the fact. Teams that gate every record catch the error before it enters the database; teams that gate nothing ship it downstream.

3 **A wrong record is worse than a missing one.** A blank field gets flagged; a wrong field gets used. The shipment goes to the wrong address, the report sums the wrong number, the model trains on the wrong data. Vetted teams hold field-level accuracy above 99.95% precisely because the tail — the wrong record acted on — is where the damage is.

4 **The downstream correction, not the keystroke, is what costs.** A record fixed at entry costs a second keystroke; the same error caught downstream costs the investigation, the correction, and the damage already done — many times over. Teams that hold accuracy drive the escaped-error rate under 0.05%, and that avoided downstream cost — not the offshore rate alone — is where the case study's return is built.

5 **Contract on the clean record, not the keyed one.** The contracting failure of data entry is paying per record and hoping they are right. Vetted deals price per clean record with field-level accuracy floors, escaped-error ceilings, and double-key gates, making an accurate database the team's problem, which is where it belongs.

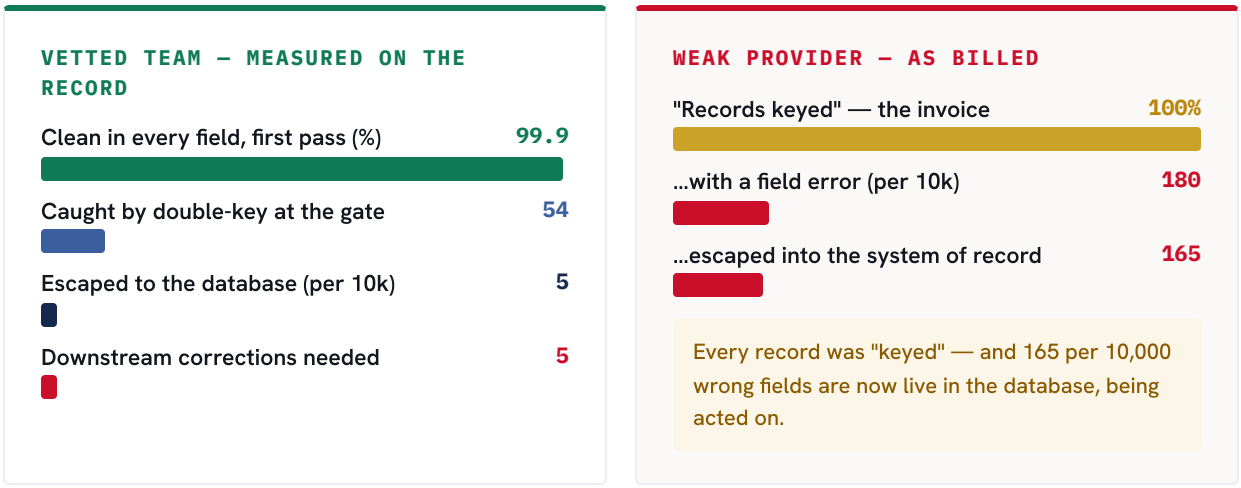
The intended reader is an operations, data, or finance leader buying data entry per record and paying for it in downstream corrections and a database no one trusts. Sections 2–4 separate the teams that produce clean records from the teams that key volume; Sections 5–7, the method and the proof.

SECTION 02

The volume mirage: records keyed versus records clean at the source

Figure 1 shows one batch of 10,000 records two ways: as the volume report a weak provider bills, and as the clean-record reality the database inherits. The records-keyed number is identical in spirit to a chatbot's containment rate — a figure engineered to look like productivity while the errors quietly populate the system of record.

FIG. 1 — ONE 10,000-RECORD BATCH: CLEAN-RECORD QUALITY VS. REPORTED VOLUME



Left: PITON-Global engagement instrumentation, 2025–26, field-level audited. Right: composite of pre-engagement vendor reporting reviewed in diligence.

The economics follow the escaped-error rate, not the volume. On the left, the double-key gate catches fifty-four errors before they enter and only five per ten thousand escape; on the right, a cheap per-record price ships 165 wrong fields per ten thousand straight into the system of record, each one to be discovered — expensively — when something built on it breaks. A low per-record price that populates the database with errors is not a saving; it is a downstream correction bill, delayed and multiplied. The field-level accuracy audit — clean-record rate, escaped errors, correction cost — is the cheapest due-diligence instrument in this paper.

SECTION 03

The data contract: the double-key gate, field-level QA, and the clean-at-source close

Figure 2 lays out the three clauses of a working data contract — the machinery a floor read verifies before any per-record price is believed.

FIG. 2 — THE DATA CONTRACT, AS OPERATED ON VETTED TEAMS

DOUBLE-KEY THE CRITICAL

Critical fields are keyed independently by two operators and reconciled on mismatch, so a slip by one is caught by the other before the record commits. The database never sees a critical field that only one pair of eyes confirmed.

VALIDATE AT THE KEYSTROKE

Format, range, and cross-field rules fire as the record is entered — a date that cannot exist, an amount out of range, a code that fails its checksum is rejected on the spot. Errors are prevented at the source, not inspected out of a sample later.

CLOSE IT CLEAN AT SOURCE

Each batch is released only when its field-level accuracy clears the floor, with an escaped-error log fed back into the rules. Records leave the operation clean, so nothing downstream has to re-clean them — which is the only economical place to fix data.

↓ yields ↓

THE TRUSTED DATABASE 99.95%+ FIELD ACCURACY | HELD UNDER VOLUME | <0.05% ESCAPED

Every field is right at the source, accuracy holds when volume spikes, and downstream systems can trust the data without re-checking it — so the database becomes an asset instead of a liability. The operation gets cheaper per clean record and more trustworthy at the same time, which is the entire promise of outsourcing data entry to people who measure the record, not the volume.

Operating parameters from PITON-Global data-entry engagements, 2025–26.

In diligence, the contract is tested at the field level: seed a batch with known errors and measure how many the double-key gate and validation rules catch; check that accuracy holds when volume doubles; review the escaped-error log for how each was root-caused. Perhaps one team in ten survives it cleanly. Steps 3 and 5 of the framework exist to find that one before the contract does — because everything in Table 3 of Section 6 depends on field-level accuracy, not records keyed.

SECTION 04

Cost and performance: cost-per-clean-

record economics and benchmarks

TABLE 1 — COST PER 1,000 CLEAN RECORDS: THREE OPERATING MODELS (US$, 2026)

OPERATING MODEL	COST / 1,000 CLEAN	VS. US BASELINE
US onshore in-house data-entry team	$430	baseline
Low-cost offshore keying mill (per-record)	$255	−41%
Vetted Manila team — double-key gated, validated	$163	−62%

Basis: PITON-Global engagement pricing 2025–26; clean record = one record correct in every field at the source, net of re-keys; the keying mill looks cheap per record but escaped errors cost far more downstream; US comparator fully-loaded in-house cost per clean record.

TABLE 2 — DATA-ENTRY BENCHMARKS, VETTED VS. BASELINE (2025–26)

METRIC	VETTED TOP TIER	BASELINE	EXECUTIVE READING
Field-level accuracy	99.95%+	98.0–99.0%	The only data number the business feels
Accuracy held at 2× volume spike	yes	degrades	Quality under load, or only when quiet
Escaped-error rate (to database)	<0.05%	1–2%	"Close it clean" — caught or shipped
Double-key on critical fields	100%	sampled	Every critical field, or a lucky few
Turnaround SLA adherence	99%+	88–94%	Clean and on time, or one of the two

Source: PITON-Global data-entry engagement data 2025–26; baseline = typical pre-engagement providers billed on records keyed.

The strategic read: the keying mill's −41% per record is a false economy — the escaped errors cost the business far more downstream than the per-record gap saves, and a corrupted database can cost more than it saved outright. The vetted Manila team's −62% is real because it is measured where the value is: records correct in every field at the source. Buy on records keyed and you buy the right column of Figure 1; buy on cost-per-clean-record and you buy a database you can trust — the same discipline every volume in this series keeps arriving at from a different door.

SECTION 05

Sourcing discipline: the 7-Step Vendor Vetting Framework, applied to data entry

Data-entry diligence adds one hard requirement to every step: a per-record price is real only when the field-level accuracy behind it survives a seeded-error audit under load. The funnel below is representative of a 25–60-seat data-entry search.

FIG. 3 — THE 7-STEP FRAMEWORK AS A DATA-ENTRY FUNNEL

	STEP	WHAT IT ELIMINATES IN A DATA-ENTRY SEARCH	FIELD REMAINING
1	**Requirements definition**	Volume-first scoping: fixes the field-level accuracy floor, the escaped-error ceiling, the double-key requirement, and the turnaround SLA the records must meet	1,000+
2	**Market mapping**	Keying mills: operations that bill volume and treat accuracy as the client's problem, and generalist floors with no double-key or validation infrastructure	~48
3	**Capability screening**	Teams without the machinery: audited field-level accuracy under load, a double-key gate on critical fields, keystroke validation rules, and a root-caused escaped-error log	~13
4	**Compliance & security audit**	Weak data governance: source-document and PII protection, access controls on the data store, retention and disposal discipline, SOC 2 scope covering the data-entry stack	~9
5	**Operational diligence**	The seeded-error read: plant known errors in a live batch and measure catch rate; double the volume and re-measure accuracy; interview the team on how they root-cause an escaped error	4–6
6	**Competitive RFP & negotiation**	Per-record pricing: finalists bid per clean record with field-level accuracy floors, escaped-error ceilings, and double-key gates — all carrying credits for records that reach the database wrong	2–3
7	**Transition & launch governance**	Cold cutover: the team calibrates validation rules and double-key on a sample before it owns the feed, with the accuracy floor certified at target volume before steady state, PITON-Global embedded buyer-side	1

Field-remaining counts represent a 25–60-seat data-entry search, PITON-Global engagements 2024–26.

The standing guarantees hold: **vendor neutrality** — PITON-Global operates no data-entry floors, keys no records, and takes no placement economics, so the shortlist reflects audited field-level accuracy under load, not inventory; **right-sizing** — teams for whom your volume is a flagship account. Six to ten genuinely qualified providers, delivered within 48–72 hours, free to the buyer, competing through a fully transparent, fair, comprehensive, and highly competitive RFP.

SECTION 06 · ANONYMIZED CASE STUDY

Re-basing the operation on accuracy: a 40-seat data rebuild, reconstructed

Engagement DE-035 is a US logistics and distribution company — keying roughly 9 million shipment, address, and catalog records a year — that had offshored data entry to a per-record vendor. The record price was excellent; the data was not: a 1.8% field-error rate meant mis-shipments, returned freight, and a product catalog the analytics team had stopped trusting. The VP of operations wanted the arbitrage and a database the business could act on. Identifying details are anonymized under NDA.

SELECTION (WEEKS 0–7, SEEDED-ERROR READS)

Requirements fixed a 99.95% field-level accuracy floor, a <0.05% escaped-error ceiling, and a double-key requirement on all critical fields, with accuracy to hold at 2× volume. Mapping produced 47 claimants; the machinery screen — audited accuracy under load, double-key gate, validation rules — left 13; the data-governance audit left 9. Seeded-error reads on each finalist under a doubled batch shortlisted 5. A Manila team won at a $0.019 blended per-record rate priced per clean record, with a field-level accuracy audit written into the reporting spec.

TRANSITION (MONTHS 2–6, RULES CALIBRATED FIRST)

The team calibrated validation rules and double-key on a sample feed before owning the live one; the accuracy floor was certified at target volume in month 2. The escaped-error log fed the rules from day one, and a peak-season volume spike was run without an accuracy drop. By month 6 field-level accuracy held at 99.96%, escaped errors fell from 1.8% to 0.04%, and the analytics team put the catalog back into production reporting. PITON-Global chaired the data-quality review board to steady state.

TABLE 3 — DE-035 FIRST-YEAR VALUE BUILD-UP (US$)		
VALUE COMPONENT	YEAR 1	BASIS
Labor arbitrage on the 40-seat data operation	$1.32M	75.0k productive hrs × ($26.10 − $8.50)
Mis-shipment & returned-freight collapse (errors 1.8% → 0.04%)	$0.52M	Avoided re-ships, freight, and correction handling
Downstream data-cleanup avoided	$0.24M	Analytics and ops hours no longer spent re-cleaning records
Catalog restored to decision use (margin recovered)	$0.14M	Recovered pricing/assortment decisions on trusted catalog, conservative
Gross value created	**$2.22M**	
Less: engagement & transition costs	($0.36M)	Rule calibration, double-key setup, source-system integration; advisory fee $0 (provider-paid model)
Net value ÷ cost = first-year ROI	**6.2×**	$2.22M ÷ $0.36M

Figures rounded; accuracy, escaped-error, and downstream effects audit-verified. Method in Methodology & notes.

Fifty-nine percent arbitrage, forty-one percent from accuracy, avoided mis-shipments, and a restored catalog — the series' familiar split, earned because the operation was re-based on field-level accuracy rather than volume keyed. Every line traces to a diligence step: the accuracy floor to Step 1, the per-clean-record pricing to Step 6, the double-key gate to Step 3's contract. The VP's line at the year-one review: "The record price barely moved. What moved is that we trust the data again."

SECTION 07

The executive playbook: governance for the first 180 days

A data-entry transition succeeds when the team is measured and paid on field-level accuracy at volume — client in command throughout.

DAYS 0–30

Define clean, contract per clean record

Set the field-level accuracy floor, the escaped-error ceiling, and the double-key requirement before launch, and specify the validation rules per field. Contract per clean record — never per record keyed — with accuracy floors and escaped-error ceilings carrying credits. Stand up the field-level accuracy audit as the program's arbiter of truth.

DAYS 30–90

Calibrate rules, certify under load

Calibrate validation rules and double-key on a sample before the team owns the live feed; certify the accuracy floor at target volume and again at a spike. Switch on the escaped-error log feeding the rules from day one. Publish the data-quality dashboard: field accuracy, escaped-error rate, accuracy-under-load, turnaround SLA.

DAYS 90–180

Own the feed, certify steady state

Hand the live feed to the team only after it holds the accuracy floor through a volume spike. Review the data-quality board monthly with the escaped-error log on the table. Certify steady state on two months at floor including one peak-volume period; hand over on documented criteria; keep the quarterly seeded-error audit standing.

THE FIVE FAILURE MODES WE ARE HIRED TO PREVENT

Paying per record while errors reach the database · sampling QA instead of double-keying critical fields · accuracy that degrades the moment volume spikes · an escaped-error rate no one root-causes · owning the live feed before accuracy is proven under load. All five are settled at selection and contract.

The data-entry argument, in one sentence: a keyed record is activity and a clean record is a database you can trust, and a vetted Philippine team delivers −62% per clean record with 99.95%+ field-level accuracy held under load — for buyers who audit the field-level accuracy instead of counting records keyed, and compete the finalists through a fully transparent, fair, comprehensive, and highly competitive RFP. Anyone can key a record. We make sure you buy the team that keys it right.

METHODOLOGY & NOTES

How the numbers were built

Accuracy and escaped-error figures (Fig. 1, Table 2) derive from engagement instrumentation in PITON-Global data-entry engagements, 2025–26. Field-level accuracy is measured against seeded-error controls and downstream reconciliation; accuracy-under-load is tested at 2× the baseline batch. The weak-provider column is a composite of pre-engagement vendor reporting reviewed in diligence, shown for contrast.

Cost-per-clean-record figures (Table 1) are fully loaded — wages, benefits and statutory charges, facilities, technology and data-capture tooling, management and QA overhead including double-key operators, attrition-driven recruiting — net of re-keys; US comparators reflect fully-loaded in-house cost on the same basis. A clean record is one correct in every field at the source.

The case study (DE-035) is anonymized under NDA; volumes and dates are generalized, figures rounded. Accuracy, escaped-error, and downstream effects are audit-verified; the restored-catalog line is a conservative estimate, not a booked saving. ROI = net first-year value ÷ all engagement-related client costs. PITON-Global's advisory fee is provider-paid and appears at $0 in the client cost base; neutrality safeguards are documented at piton-global.com.

Market context draws on IBPAP reporting and PITON-Global's proprietary provider mapping (1,000+ operations; data-entry field-level accuracy under load re-verified May 2026). Estimates and projections are PITON-Global analysis and should be cited as such. The full series is available at piton-global.com.

Would your provider's record price survive a seeded-error audit?

Forty-five minutes with John will tell you. Free to buyers, strictly vendor-neutral — an accuracy-verified shortlist within two to three business days.

j.maczynski@piton-global.com
+1 402-598-8740