PITON Global

EXECUTIVE WHITE PAPER

TRUST & SAFETY · PHILIPPINES

The policy-fidelity standard: the executive economics of trust & safety outsourcing to the Philippines.

An analysis of why cases closed is a volume vanity metric, how policy fidelity and program coverage — never case throughput — decide the true cost of a trust-and-safety operation once policy drift, uneven enforcement across harm areas, slow crisis response, and regulatory gaps are counted, and the vendor-selection discipline that runs the whole safety program to policy.

AUTHORS

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

MANILA · ESTABLISHED 2001

TRUST & SAFETY

Contents

– About the authors 03

01 Executive summary 04

02 The volume mirage: cases closed versus policy-faithful outcomes 05

03 The trust-and-safety contract: enforce to policy, cover the harm surface, respond to crisis 06

04 Cost and performance: cost-per-policy-faithful-case economics and benchmarks 07

05 Sourcing discipline: the 7-Step Vendor Vetting Framework, applied to trust & safety 08

06 Case study: a 140-seat trust-and-safety program re-based on policy fidelity 09

07 The executive playbook: governance for the first 180 days 10

– Methodology & notes 11

ABOUT THIS SERIES

Every trust-and-safety proposal is priced per case closed, because volume is easy to count. This volume is about the only program outcome a platform's safety and its regulators actually depend on: enforcement that is faithful to policy, consistent across every harm area, and fast when a crisis breaks. The gap between closing a case and running a policy-faithful program is where drift, coverage gaps, and regulatory exposure live. The full series lives at piton-global.com.

THE PITON-GLOBAL LEADERSHIP TEAM

About the authors

John Maczynski

Chief Executive Officer

John has bought trust-and-safety capacity for three decades, and learned that the fast program that closes every case and drifts from policy is the most expensive one a platform can hire — because drift is uneven enforcement a regulator will find, a coverage gap is a harm area nobody is watching, and a slow crisis response is a headline. His standing question ignores the cases-closed report: of the cases you closed, how many were faithful to policy, across every harm area, fast enough when it mattered? Applied at PITON-Global, that question is what clients get on every trust-and-safety 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 trust-and-safety programs, and he judges an operation by its policy fidelity and harm-surface coverage, not its cases-per-day: whether enforcement matched policy consistently across every harm area, and whether a crisis got a fast, correct response. The trust-and-safety contract in Section 3 codifies what he learned rebuilding programs that closed cases fast and drifted from policy. He now reads trust-and-safety providers the way he once ran them — from the policy-drift audit and the coverage-gap map 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 trust-and-safety operation's real policy fidelity and harm-coverage record rather than its cases-per-day 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

-58%
Fully-loaded cost per policy-faithful case vs. US onshore

98%+
Policy-fidelity rate across harm areas on vetted teams, up from 84–91%

-61%
Reduction in crisis response time on vetted teams

6.3×
First-year ROI, reconstructed case (Section 6)

A trust-and-safety engagement is bought on the price per case and paid for on the program's drift from policy. The vendor reports millions of cases closed at a fraction of the onshore cost; the platform finds the enforcement that drifted from written policy and now varies by shift, the harm areas nobody was really covering, the crisis that got a slow and improvised response, and the regulator asking for a consistency report the operation cannot produce. The operation closed its cases and the platform inherited policy drift, coverage gaps, and regulatory exposure across its whole safety program. The gap between closing a case and running a policy-faithful program is where the real cost of cheap trust-and-safety work hides, because a case closed off-policy is not throughput delivered — it is drift accrued and exposure taken on.

The Philippines — English-fluent, culturally fluent with Western platform norms, and with the deepest at-scale trust-and-safety talent pool offshore — is where trust and safety is delivered with the most policy-fidelity discipline, because that pool can run the consistent enforcement, full harm-surface coverage, and crisis readiness a program demands. Five findings:

1 **Cases closed is the industry's most flattering number.** Any program can close cases at volume. Vetted teams are measured on outcome: policy fidelity, harm-surface coverage, and crisis-response time. Weak providers quote a low per-case price the platform repays in drift, coverage gaps, and regulatory exposure. Section 2 shows the gap.

2 **The policy-faithful, well-covered program, not the closed case, is the product.** Safety that protects users and satisfies regulators comes from enforcing to written policy, consistently across every harm area, with a fast path when crisis hits — not from closing a case queue. Programs that stay faithful hold up to scrutiny; programs that close loosely drift and leave gaps, at a cost far above the per-case price.

3 **A program failure cuts two ways — drift and gap.** Enforcement that drifts from policy is inconsistent and indefensible; a harm area left uncovered is a safety failure waiting to surface. Vetted teams hold policy fidelity above 98% across the harm surface precisely because both cost — the inconsistency a regulator penalizes and the gap a bad actor exploits.

4 **The regulatory exposure and the crisis, not the case, are what cost.** A program that drifts costs the compliance finding, the remediation, and the crisis handled slowly — many times the per-case saving. Programs that stay faithful and cover the harm surface avoid the finding and answer the crisis fast, and that protected standing — not the offshore rate alone — is where the case study's return is built.

5 **Contract on the policy-faithful case, not the closed one.** The contracting failure of trust-and-safety is paying per case and hoping the program stays faithful and covered. Vetted deals price against policy-fidelity floors, harm-coverage targets, and crisis-response SLAs, making a faithful program the team's problem, which is where it belongs.

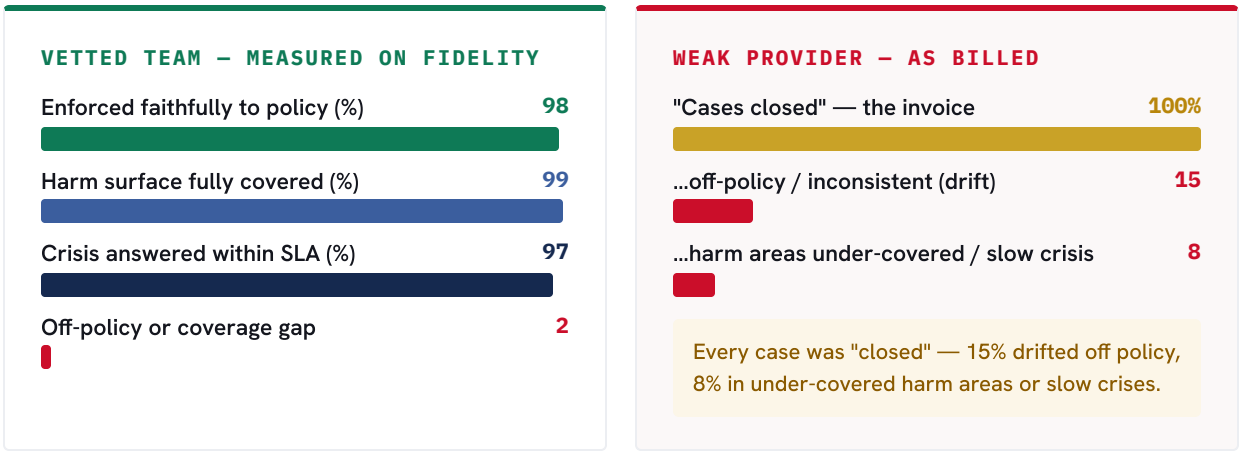
The intended reader is a head of trust and safety, chief safety officer, or COO buying trust-and-safety operations per case and paying for it in policy drift, coverage gaps, and regulatory exposure. Sections 2–4 separate the programs that stay policy-faithful from the operations that close cases; Sections 5–7, the method and the proof.

SECTION 02

The volume mirage: cases closed versus policy-faithful outcomes

Figure 1 shows one quarter of a platform's trust-and-safety program two ways: as the cases-closed report a weak vendor bills, and as the policy-fidelity reality the audit, the harm-coverage map, and the regulator experience. The cases-closed number is identical in spirit to a chatbot's containment rate — a figure engineered to look like output while policy drift, coverage gaps, and slow crisis response accumulate underneath.

FIG. 1 — ONE QUARTER OF THE PROGRAM: POLICY-FIDELITY QUALITY VS. REPORTED VOLUME



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

The economics follow the policy fidelity and coverage, not the cases count. On the left, ninety-eight percent of enforcement is faithful to policy and the harm surface is fully covered, so the program holds up to audit and answers a crisis fast; on the right, a cheap per-case price drifts on fifteen percent and under-covers eight, paying in inconsistency, coverage gaps, and slow crisis response at once. A low per-case price that drifts and leaves gaps is not a saving; it is regulatory exposure and safety risk, delayed and compounding. The policy-fidelity-and-coverage audit — fidelity rate, harm-surface coverage, crisis-response time — is the cheapest due-diligence instrument in this paper.

SECTION 03

The trust-and-safety contract: enforce to policy, cover the harm surface, respond to crisis

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

FIG. 2 — THE TRUST-AND-SAFETY CONTRACT, AS OPERATED ON VETTED TEAMS

ENFORCE TO POLICY

Every case is decided against the current written policy, held consistent across moderators, shifts, and time by calibration and audit, with policy updates propagated fast so enforcement never drifts. Fidelity to the written rule is what makes a program defensible to a regulator and fair to a user.

COVER THE HARM SURFACE

Every harm area in scope — from spam to the most severe categories — is staffed, trained, and measured, with no policy area quietly left thin, and specialist queues where severity demands them. Full coverage is how a program stops a harm nobody was assigned to watch.

RESPOND TO CRISIS

A defined escalation and surge path handles spikes, novel harms, and regulator or press events fast and correctly, with a documented evidence trail — and a fidelity-and-coverage report the platform can read. Crisis readiness is what separates a program that absorbs a shock from one that becomes the story.

↓ yields ↓

THE POLICY-FAITHFUL PROGRAM

98%+ POLICY FIDELITY · FULL COVERAGE · CRISIS-READY

Enforcement stays faithful to policy, the whole harm surface is covered, and crises get a fast correct response — so drift disappears, gaps close, and the program holds up to regulators and users alike. The operation gets cheaper per policy-faithful case and safer at the same time, which is the entire promise of outsourcing trust-and-safety to people who measure policy fidelity and coverage, not cases closed.

Operating parameters from PITON-Global trust-and-safety engagements, 2025–26. Wellness provisions for exposure-heavy queues reflect on-staff support and rotation.

In diligence, the contract is tested from the policy's side: pull closed cases across harm areas and re-decide a sample against written policy to measure true fidelity and consistency; map staffing and outcomes against the full harm surface for gaps; and run a crisis tabletop to measure response. 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 policy fidelity and coverage, not cases closed.

SECTION 04

Cost and performance: cost-per-policy-faithful-case economics and benchmarks

TABLE 1 — FULLY-LOADED COST PER POLICY-FAITHFUL CASE: THREE OPERATING MODELS (2026)

OPERATING MODEL	COST / CASE	VS. US BASELINE
US onshore in-house T&S team	$1.28	baseline
Low-cost offshore case mill (per-case)	$0.88	−31%
Vetted Manila team — policy-faithful, covered	$0.54	−58%

Basis: PITON-Global engagement pricing 2025–26; cost per policy-faithful case = fully-loaded operating cost (including policy management, calibration, and crisis-surge capacity) divided by cases closed faithfully to policy within a fully-covered harm surface (off-policy and gap cases excluded); the case mill looks cheap per case but drift, coverage gaps, and regulatory remediation lift its true cost per faithful case; US comparator fully-loaded in-house cost.

TABLE 2 — TRUST-AND-SAFETY BENCHMARKS, VETTED VS. BASELINE (2025–26)

METRIC	VETTED TOP TIER	BASELINE	EXECUTIVE READING
Policy-fidelity rate	98%+	84–91%	The only program number regulators follow
Harm-surface coverage	99%+	80–90%	Whole surface, or a blind spot
Enforcement consistency (inter-rater)	97%+	80–88%	"Enforce to policy" — held or drifting
Crisis-response time vs. baseline	−61%	baseline	Absorbed, or the story
Regulatory-reporting readiness	Audit-ready	Partial	Report on demand, or scramble

Source: PITON-Global trust-and-safety engagement data 2025–26; baseline = typical pre-engagement operations billed on cases closed.

The strategic read: the case mill's −31% per case is a false economy — drift, coverage gaps, and regulatory remediation lift its true cost per faithful case far above the per-case gap, and a compliance finding or a crisis mishandled is a cost no rate recovers. The vetted Manila team's −58% is real because it is measured

where the value is: the case enforced faithfully to policy, across a covered harm surface, crisis-ready. Buy on cases closed and you buy the right column of Figure 1; buy on cost-per-policy-faithful-case and you buy policy fidelity — 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 trust & safety

Trust-and-safety diligence adds one hard requirement to every step: a per-case price is real only when the policy fidelity behind it survives a re-decision against written policy across the full harm surface. The funnel below is representative of a 100–250-seat trust-and-safety program search.

FIG. 3 — THE 7-STEP FRAMEWORK AS A TRUST-AND-SAFETY FUNNEL

	STEP	WHAT IT ELIMINATES IN A TRUST-AND-SAFETY SEARCH	FIELD REMAINING
1	**Requirements definition**	Throughput-first scoping: fixes the policy-fidelity floor, the harm-coverage target, the crisis-response SLA, and the regulatory-readiness standard the operation must meet across the harm surface	1,000+
2	**Market mapping**	Case mills: operations that bill closed volume and treat drift and coverage gaps as the client's problem, and generalist floors with no policy-operations or crisis depth	~44
3	**Capability screening**	Teams without the machinery: audited policy-fidelity history, calibration and policy-management workflow, harm-area specialization, crisis-surge capacity, a real wellness program, and language coverage	~13
4	**Compliance & security audit**	Weak governance: user-data protection, regulatory alignment (DSA and peers), statement-of-reasons and reporting readiness, duty-of-care standards, SOC 2 scope covering the T&S stack	~8
5	**Operational diligence**	The fidelity read: re-decide cases across harm areas against written policy for true fidelity and consistency; map coverage for gaps; run a crisis tabletop; interview on how they propagate a policy change	4–6
6	**Competitive RFP & negotiation**	Per-case pricing: finalists bid against policy-fidelity floors, harm-coverage targets, and crisis-response SLAs — all carrying credits for off-policy outcomes or coverage gaps	2–3
7	**Transition & launch governance**	Cold start: the team calibrates against your policy set and harm taxonomy on a sample before it owns the program, with the fidelity floor and coverage map certified before steady state, PITON-Global embedded buyer-side	1

Field-remaining counts represent a 100–250-seat trust-and-safety program search, PITON-Global engagements 2024–26.

The standing guarantees hold: **vendor neutrality** — PITON-Global operates no trust-and-safety floors, closes no cases, and takes no placement economics, so the shortlist reflects audited policy-fidelity and coverage performance, not inventory; **right-sizing** — teams for whom your program 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 program on policy fidelity: a 140-seat trust-and-safety rebuild, reconstructed

The case engagement is a US digital platform — running roughly 58,000,000 trust-and-safety cases a year across a dozen harm areas — that had offshored its safety program to a per-case mill. The case price was excellent; the program was not: policy fidelity sat near 88%, enforcement varied by shift, two harm areas were thinly covered, crisis response was slow, and a regulator had asked for a consistency report the operation could not produce. The head of trust and safety wanted the scale without the drift and the exposure. Identifying details are anonymized under NDA.

SELECTION (WEEKS 0–8, FIDELITY READS)

Requirements fixed a 98% policy-fidelity floor, a 99% harm-coverage target, and a crisis-response SLA, with a mandatory calibration, policy-management, and surge workflow. Mapping produced 44 claimants; the machinery screen — fidelity history, calibration, harm-area specialization, crisis capacity, wellness program — left 13; the compliance-and-reporting audit left 8. Fidelity re-decisions across harm areas plus a crisis tabletop shortlisted 5. A Manila team won priced against policy fidelity and coverage, with both written into the reporting spec.

TRANSITION (MONTHS 2–7, CALIBRATED FIRST)

The team calibrated against the platform's policy set and harm taxonomy before owning the program, staffing the two thin harm areas and standing up the crisis-surge path; the fidelity floor was certified in month 3. The calibration-and-policy-management loop kept enforcement current continuously. By month 7 policy fidelity reached 98.1%, harm-surface coverage hit 99%, crisis response fell 61%, and the operation could produce a regulator-ready consistency report on demand. PITON-Global chaired the policy-fidelity review board to steady state.

TABLE 3 — CASE FIRST-YEAR VALUE BUILD-UP (US$)

VALUE COMPONENT	YEAR 1	BASIS
Labor arbitrage on the 140-seat program	**$3.62M**	262.5k productive hrs × ($23.30 − $9.50)
Regulatory exposure & remediation avoided	**$0.94M**	Compliance-finding and remediation cost removed
Crisis & coverage-gap harm avoided	**$0.40M**	Slow-crisis and blind-spot exposure removed
Rework & re-review reduction	**$0.22M**	Re-review of off-policy outcomes removed
Gross value created	**$5.18M**	
Less: engagement & transition costs	**($0.82M)**	Calibration, coverage build-out, crisis-path and reporting integration; advisory fee $0 (provider-paid model)
Net value ÷ cost = first-year ROI	**6.3×**	$5.18M ÷ $0.82M

Figures rounded; fidelity, coverage, and crisis effects audit-verified. Method in Methodology & notes.

Seventy percent arbitrage, thirty percent from avoided regulatory exposure, crisis and coverage-gap harm, and less re-review — a typical split for the series, earned because the program was re-based on policy fidelity rather than case volume. Every line traces to a diligence step: the fidelity floor and coverage target to Step 1, the fidelity-and-coverage pricing to Step 6, the calibration workflow to Step 3's contract. The head of trust and safety's line at the year-one review: "The case price barely moved. What moved is that enforcement is consistent now, nothing's uncovered — and when the regulator asked, we handed them the report."

SECTION 07

The executive playbook: governance for the first 180 days

A trust-and-safety transition succeeds when the team is measured and paid on policy fidelity and harm coverage — client in command throughout.

DAYS 0–30

Define fidelity & coverage, contract on them

Set the policy-fidelity floor, the harm-coverage target, and the crisis-response and regulatory-readiness thresholds with trust-and-safety leadership before launch, and document the policy set, harm taxonomy, and escalation paths. Contract against fidelity and coverage — never per case alone — with floors and SLAs carrying credits. Stand up the fidelity re-decision and coverage map as the program's arbiters of truth.

DAYS 30–90

Calibrate, close the coverage gaps

Calibrate the team against your policy set and harm taxonomy, staff every harm area to target, and stand up the crisis-surge path before it owns the program; certify the fidelity floor and coverage map. Switch on the policy-management and calibration loop from day one. Publish the program dashboard: policy fidelity, harm coverage, enforcement consistency, crisis-response time, reporting readiness.

DAYS 90–180

Own the program, certify steady state

Hand the program to the team only after it holds the fidelity floor and coverage target. Review the board monthly with the drift and coverage-gap trends and a live crisis-readiness check on the table. Certify steady state on two months at floor including a policy-update and a crisis-tabletop cycle; hand over on documented criteria; keep the quarterly fidelity re-decision standing.

THE FIVE FAILURE MODES WE ARE HIRED TO PREVENT

Paying per case while enforcement drifts from policy · closing cases instead of covering the harm surface · leaving a harm area thinly staffed · improvising a slow crisis response · owning the program before fidelity and coverage are proven. All five are settled at selection and contract.

The trust-and-safety argument, in one sentence: a closed case is activity and a policy-faithful program is consistent enforcement, a fully-covered harm surface, and a crisis answered fast, and a vetted Philippine team delivers −58% cost per policy-faithful case with 98%+ policy fidelity, 99%+ harm coverage, and crisis response cut by over half — for buyers who audit the fidelity and the coverage instead of counting cases closed, and compete the finalists through a fully transparent, fair, comprehensive, and highly competitive RFP. Anyone can close a case. We make sure you buy the team that runs the whole program to policy.

METHODOLOGY & NOTES

How the numbers were built

Fidelity and coverage figures (Fig. 1, Table 2) derive from engagement instrumentation in PITON-Global trust-and-safety engagements, 2025–26. Policy fidelity and enforcement consistency are measured by re-deciding cases across harm areas against written policy; harm-surface coverage, crisis-response time, and regulatory-reporting readiness are measured against the client's pre-engagement baseline. The weak-provider column is a composite of pre-engagement vendor reporting reviewed in diligence, shown for contrast.

Cost-per-policy-faithful-case figures (Table 1) are fully loaded — wages, benefits and statutory charges, facilities, technology and tooling, management and QA overhead including policy managers, calibration leads, crisis-surge capacity, and on-staff wellness support, and attrition-driven recruiting — divided by cases closed faithfully to policy within a fully-covered harm surface, with off-policy or coverage-gap cases excluded from the denominator; US comparators reflect fully-loaded in-house cost on the same basis. A policy-faithful case is enforced to written policy, within a covered harm surface, crisis-ready.

The case study is anonymized under NDA; volumes and dates are generalized, figures rounded. Fidelity, coverage, and crisis effects are audit-verified; the regulatory-exposure-avoided line reflects compliance-finding and remediation cost removed, and the crisis-harm-avoided line is a conservative estimate. 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; policy fidelity and harm-coverage performance re-verified May 2026). Moderator wellbeing is treated as both an ethical obligation and an operating fundamental; wellness provisions reflect on-staff support, exposure limits, and rotation. 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 case price survive a policy-fidelity audit?

Thirty minutes with John will tell you. Free to buyers, strictly vendor-neutral — a policy-fidelity-verified shortlist within two to three business days.

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

© 2026 PITON-GLOBAL, INC. · EXECUTIVE WHITE PAPER SERIES · TRUST & SAFETY · PHILIPPINES. This publication is provided for general information. Figures identified as estimates or models are PITON-Global analysis. Reproduction with attribution is permitted.