Insurance Sector Cybersecurity Psychology Framework (IS-CPF v1.0):

The Empathy Factor and Distributed Agent Networks as Amplifiers of Core Vulnerability Categories

Giuseppe Canale, CISSP
Independent Researcher
g.canale@cpf3.org
URL: cpf3.org
ORCID: 0009-0007-3263-6897

December 2025

Abstract

The insurance sector presents a distinctive psychological vulnerability profile shaped by two defining characteristics: the centrality of empathy in claims operations and the decentralized structure of agent distribution networks. This paper presents the Insurance Sector Cybersecurity Psychology Framework (IS-CPF v1.0), demonstrating that sector-specific risks—including Agent Network Deference, Claims Surge Collapse, Empathy Exploitation, and Insurtech Automation Blindness—constitute *calibrated manifestations* of the established Core 10 CPF taxonomy rather than novel psychological categories. By mapping insurance-specific phenomena to Categories 1, 2, 4, and 9, we preserve the mathematical integrity of the Implementation Companion's Bayesian network architecture while enabling precise detection and intervention in insurance environments. The framework addresses the unique challenge of securing organizations where professional effectiveness requires emotional openness that simultaneously creates exploitable vulnerabilities. We present sector-specific detection functions, intervention strategies adapted to agent network resistance patterns, and a case study of the "Deepfake Life Claim" demonstrating convergent exploitation of AI bias and affective vulnerabilities.

Keywords: insurance cybersecurity, claims processing, agent networks, empathy exploitation, insurtech, deepfake fraud, psychological vulnerabilities, CPF implementation

1 Introduction

1.1 The Insurance Sector Threat Landscape

The insurance industry occupies a unique position in the cybersecurity threat landscape, serving simultaneously as a repository of extraordinarily sensitive personal data, a financial institution processing billions in claims payments, and a distributed network of semi-autonomous agents operating beyond direct organizational control. This configuration creates a threat environment distinct from both traditional financial services and healthcare, combining elements of each while introducing vulnerabilities unique to the insurance operating model.

1.1.1 Data Sensitivity and Attack Surface

Insurance organizations maintain data portfolios of exceptional sensitivity and breadth. Life insurers hold detailed medical histories, genetic information, and beneficiary designations. Health

insurers possess comprehensive treatment records and diagnostic information. Property and casualty insurers maintain asset inventories, security system specifications, and behavioral patterns derived from telematics. This data concentration makes insurers prime targets for ransomware operations, with the 2021 CNA Financial attack demonstrating that even major carriers face existential threats from data-encrypting adversaries (CNA Financial, 2021).

The regulatory environment compounds exposure. HIPAA requirements for health data, state insurance commissioner mandates, and emerging privacy regulations (GDPR, CCPA) create compliance obligations that may conflict with operational efficiency. Breach notification requirements ensure that successful attacks produce both operational disruption and reputational damage.

1.1.2 The Legacy System Problem

Insurance industry consolidation has produced organizations operating heterogeneous technology environments accumulated through decades of acquisitions. A typical large insurer may maintain policy administration systems from multiple vendors, claims platforms of varying vintages, and agent portals reflecting the technology choices of acquired companies. This technical debt creates security gaps at system boundaries, authentication inconsistencies across platforms, and monitoring blind spots where legacy systems lack instrumentation capability.

The COBOL systems still processing policies written decades ago cannot be easily replaced—they encode business logic accumulated over generations that no current employee fully understands. Security controls designed for modern architectures cannot be retrofitted to these systems, creating permanent vulnerability pockets within the organizational technology landscape.

1.1.3 The Agent Network Challenge

Unlike vertically integrated financial institutions, insurers operate through distributed networks of agents who maintain varying degrees of organizational affiliation. Independent agents represent multiple carriers, captive agents work exclusively for single insurers, and broker networks operate as intermediaries. Each configuration creates distinct security challenges:

- Independent agents access carrier systems from environments outside carrier security control
- Agent offices maintain local data stores with inconsistent protection
- High-performing agents acquire informal authority that inhibits security enforcement
- Agent compensation structures create incentives misaligned with security requirements

1.1.4 The Empathy Vulnerability

Most distinctively, insurance claims operations require employees to engage empathetically with individuals experiencing significant life stressors: accident victims, bereaved beneficiaries, disaster survivors, and patients confronting serious diagnoses. This empathy requirement is not incidental but central to claims function—adjusters who cannot connect with claimants cannot effectively assess claims or manage claimant expectations.

Yet empathy creates exploitable vulnerability. The claims adjuster trained to respond supportively to distress signals becomes susceptible to social engineering that manufactures distress. The professional skill of believing claimants' narratives becomes a security weakness when adversaries construct compelling false narratives. The insurance sector thus faces a fundamental tension: the emotional openness required for professional effectiveness simultaneously creates psychological attack surface.

1.2 The Sector Adaptation Rationale

This paper extends the Cybersecurity Psychology Framework to insurance contexts while maintaining strict compatibility with the established theoretical and mathematical architecture. Following the methodology established in FS-CPF v2.0 (Canale, 2025d), we demonstrate that insurance-specific phenomena represent *manifestations* of Core 10 categories requiring calibration rather than novel categories requiring architectural extension.

This approach preserves:

(1) **Bayesian Network Integrity**: The joint probability distribution $P(I_1, \ldots, I_{100}) = \prod_{i=1}^{100} P(I_i \mid \text{parents}(I_i))$ remains valid with recalibrated conditional probabilities

(2) **Detection Algorithm Compatibility**: OFTLISRV implementations require parameter adjustment, not structural modification

(3) **SOC Integration**: Organizations implementing standard CPF can deploy IS-CPF through configuration rather than redevelopment

(4) **Theoretical Completeness**: The Core 10 taxonomy's claim to capture all psychologically-relevant vulnerability dimensions remains validated

1.3 Document Structure

Section 2 maps insurance-specific phenomena to Core 10 categories with mathematical calibration specifications. Section 3 presents CPIF intervention methodology adapted for insurance organizational dynamics, with particular attention to agent network resistance. Section 4 provides OFTLISRV technical implementation details. Section 5 presents the Deepfake Life Claim case study. Section 6 concludes with validation requirements and deployment roadmap.

2 Sector-Specific Manifestations: Mapping Insurance Phenomena to the Core 10 Taxonomy

The insurance sector does not introduce novel psychological vulnerabilities; rather, it creates environmental conditions that activate, amplify, and combine existing vulnerability categories in distinctive configurations. This section maps four critical sector-specific phenomena to their foundational CPF categories.

2.1 Category 1 Manifestation: Agent Network Deference

2.1.1 Theoretical Foundation

Category 1 (Authority-Based Vulnerabilities) encompasses patterns of deference to perceived authority figures, originally grounded in Milgram (1974) obedience research. Standard indicators address authority figure impersonation, executive exception normalization, and authority gradient effects inhibiting security reporting.

In insurance contexts, the relevant "authority" extends beyond formal organizational hierarchy to include *economic authority*—the power derived from revenue generation rather than positional rank. High-performing agents, while formally subordinate to corporate security functions, exercise de facto authority through their contribution to organizational revenue.

2.1.2 Manifestation Characteristics

Agent Network Deference describes the systematic inhibition of security enforcement against revenue-generating agents:

(1) **Top Performer Immunity**: Agents generating significant premium volume receive implicit exemption from security requirements that would impede their sales processes. Security teams learn not to escalate agent violations when the agent's production ranking exceeds informal thresholds.

(2) **Revenue-Based Authority Gradient**: The authority gradient inhibiting upward reporting (Indicator 1.6) operates inversely in agent contexts—security personnel hesitate to report violations by agents whose revenue contribution exceeds the security team's perceived organizational value.

(3) **Termination Threat Dynamics**: Agents, particularly independent agents representing multiple carriers, implicitly threaten to redirect business to competitors if security requirements become "too burdensome." This economic coercion produces compliance behavior in security teams analogous to authority-based compliance in hierarchical contexts.

(4) **Exception Cascade**: Security exceptions granted to top performers establish precedents that propagate through the agent network. "If Agent X doesn't need MFA, why do I?" The exception becomes the norm through social proof amplification (Category 3 interaction).

2.1.3 Mathematical Mapping

Agent Network Deference maps to indicators 1.4, 1.6, and 1.8 with insurance-specific calibration:

Indicator 1.4 (Convenience-Based Bypassing) - Insurance Calibration:

The convenience bypass ratio requires agent-specific formulation:

$$CBR^{INS}(a,t) = \frac{E_{agent}(a,t)}{E_{standard}(t)} \cdot \frac{R_{premium}(a)}{R_{threshold}} \tag{1}$$

Where:

- $E_{agent}(a,t)$ = security exceptions granted to agent a in period t
- $E_{standard}(t)$ = standard exception rate in period t
- $R_{premium}(a)$ = premium revenue attributed to agent a
- $R_{threshold}$ = revenue threshold for "top performer" designation

Detection triggers when $CBR^{INS} > 2.0$ (exceptions more than double standard rate) AND correlation $\rho(R_{premium}, E_{agent}) > 0.6$ (revenue predicts exceptions).

Indicator 1.6 (Authority Gradient Effects) - Insurance Calibration:

The authority gradient function requires economic dimension:

$$AG^{INS}(i,a) = \alpha \cdot \frac{H_a - H_i}{H_{max}} + \beta \cdot \frac{R_{premium}(a)}{R_{max}} \tag{2}$$

Where the economic term $\beta \cdot \frac{R_{premium}(a)}{R_{max}}$ captures revenue-based authority independent of hierarchical position. Empirical calibration suggests $\beta > \alpha$ in insurance contexts (economic authority exceeds positional authority).

Indicator 1.8 (Executive Exception Normalization) - Insurance Calibration:

$$EEN^{INS}(t) = \frac{\sum_{a \in \text{TopPerformers}} E_a(t)}{\sum_{a \in \text{AllAgents}} E_a(t)} \cdot N_{agents} \tag{3}$$

Detection triggers when top performers (top 10% by revenue) account for more than 30% of total exceptions.

2.1.4 Conditional Probability Update

$$P^{INS}(1.8|1.4) = 0.85 \quad \text{(versus base } P(1.8|1.4) = 0.65) \tag{4}$$

The elevated conditional probability reflects that convenience bypasses for agents strongly predict exception normalization in insurance contexts.

2.2 Category 2 Manifestation: Claims Surge Collapse

2.2.1 Theoretical Foundation

Category 2 (Temporal Vulnerabilities) addresses the interaction between time pressure and cognitive capacity. Standard indicators assume baseline temporal pressures with episodic intensification. Insurance claims operations face a distinctive temporal pattern: long periods of manageable volume punctuated by catastrophic surge events that overwhelm processing capacity.

Natural disasters, mass casualty events, and pandemic conditions produce claims volumes that exceed normal capacity by orders of magnitude. Organizations face impossible choices: maintain security controls and fail claimants during crisis, or disable controls to meet humanitarian obligations.

2.2.2 Manifestation Characteristics

Claims Surge Collapse describes the systematic degradation of security controls during high-volume claims events:

(1) **MFA Suspension**: Multi-factor authentication requirements are suspended to enable rapid claims processor deployment. Temporary workers and redeployed staff from other functions cannot be provisioned with MFA tokens within operationally acceptable timeframes.

(2) **Verification Threshold Reduction**: Normal claims verification procedures (documentation requirements, identity confirmation, loss validation) are reduced or eliminated to accelerate payment. The humanitarian imperative to "get money to victims" overrides security concerns.

(3) **Approval Authority Expansion**: Claims approval authorities are temporarily expanded, enabling individual adjusters to approve larger payments without supervisory review. This reduces bottlenecks while eliminating segregation of duties.

(4) **Third-Party Access Expansion**: Catastrophe response often requires engaging third-party adjusting firms, restoration contractors, and temporary staffing agencies. These entities receive system access under emergency provisions that bypass normal vetting.

(5) **Audit Trail Suspension**: Documentation requirements are relaxed to accelerate processing. Claims approved during surge periods may lack the audit trail required for subsequent fraud detection or regulatory examination.

2.2.3 Mathematical Mapping

Claims Surge Collapse maps to indicators 2.1, 2.3, and 2.6 with catastrophe-event calibration:

Indicator 2.1 (Urgency-Induced Bypass) - Insurance Calibration:
The urgency index requires claims-volume integration:

$$U_i^{INS}(t) = \frac{\Delta t_{normal} - \Delta t_{surge}}{\Delta t_{normal}} \cdot \frac{V_{claims}(t)}{V_{baseline}} \tag{5}$$

Where $\frac{V_{claims}(t)}{V_{baseline}}$ amplifies urgency proportional to claims volume above baseline. During catastrophe events, this multiplier may exceed 10x.

The Poisson regression for bypass rate incorporates catastrophe state:

$$\lambda^{INS} = e^{\beta_0 + \beta_1 \cdot \text{volume\_ratio} + \beta_2 \cdot \text{CAT\_declared} + \beta_3 \cdot \text{media\_pressure}} \tag{6}$$

Where CAT_declared is a binary indicator for declared catastrophe status and media_pressure captures external attention intensity.

Indicator 2.6 (Temporal Exhaustion) - Insurance Calibration:

Standard circadian modeling requires surge-duration extension:

$$E^{INS}(t) = E_0 \cdot \left(1 + A \cdot \sin\left(\frac{2\pi(t-\phi)}{24}\right)\right) \cdot e^{-\lambda_{exhaust} \cdot D_{surge}} \tag{7}$$

Where D_{surge} = duration of surge conditions in days. The exponential decay term captures cumulative exhaustion that baseline circadian patterns cannot represent.

2.2.4 Conditional Probability Update

$$P^{INS}(5.x|2.x) = 0.92 \quad \text{(versus base } P(5.x|2.x) = 0.70) \tag{8}$$

Claims surge conditions produce cognitive overload with near-certainty, exceeding even HFT environments in the FS-CPF calibration.

2.3 Category 4 Manifestation: Empathy Exploitation

2.3.1 Theoretical Foundation

Category 4 (Affective Vulnerabilities) addresses the influence of emotional states on security-relevant decision-making. Standard indicators address fear paralysis, anger-induced risk-taking, and emotional contagion. These indicators assume that emotional states are *incidental*—conditions that arise and must be managed, but not requirements of professional function.

Insurance claims work inverts this assumption. Empathy is not incidental but *constitutive*—claims adjusters who cannot engage empathetically with claimants cannot perform their professional function. The emotional openness required for effective claims handling is the same emotional openness that social engineers exploit.

2.3.2 Manifestation Characteristics

Empathy Exploitation describes the systematic targeting of claims personnel through manufactured emotional appeals:

(1) **Distress Signal Manipulation**: Social engineers construct scenarios triggering trained empathetic responses. The adjuster conditioned to respond supportively to bereaved beneficiaries cannot easily distinguish genuine bereavement from performed bereavement.

(2) **Narrative Coherence Exploitation**: Claims adjusters evaluate claim legitimacy partly through narrative assessment—does the claimant's story "make sense"? Sophisticated adversaries construct narratively coherent false claims that satisfy the adjuster's pattern-matching while concealing fraud indicators.

(3) **Urgency-Empathy Compound**: Adversaries combine urgency claims ("I need this payment for my mother's funeral") with emotional distress displays, activating both Category 2 (temporal) and Category 4 (affective) vulnerabilities simultaneously.

(4) **Helper Identity Exploitation**: Claims adjusters often self-identify as "helpers"—professionals whose purpose is assisting people through difficult circumstances. Social engineers exploit this identity by framing compliance with fraudulent requests as "helping."

(5) **Guilt Induction**: Adversaries who encounter resistance induce guilt through statements implying that security requirements cause claimant suffering: "My children haven't eaten because you won't release this payment." Adjusters trained to prioritize claimant welfare experience cognitive dissonance when security requirements conflict with apparent welfare.

2.3.3 Mathematical Mapping

Empathy Exploitation maps to indicators 4.3, 4.5, 4.6, and 4.10 with claims-specific calibration:

Indicator 4.3 (Trust Transference) - Insurance Calibration:

The trust transference function requires claimant-relationship integration:

$$TT^{INS}(c,a) = T_{baseline} + \alpha \cdot \text{distress\_signals}(c) + \beta \cdot \text{narrative\_coherence}(c) + \gamma \cdot \text{interaction\_duration}(c,a) \tag{9}$$

Where:

- distress_signals(c) = detected emotional distress indicators in claimant communication
- narrative_coherence(c) = assessed coherence of claim narrative
- interaction_duration(c, a) = cumulative interaction time between claimant and adjuster

Detection triggers when TT^{INS} exceeds thresholds AND claim exhibits fraud indicators that should have reduced trust.

Indicator 4.6 (Guilt-Driven Overcompliance) - Insurance Calibration:

$$GDO^{INS}(a,t) = \frac{N_{exceptions}(a,t)}{N_{interactions}(a,t)} \cdot S_{guilt}(a,t) \tag{10}$$

Where $S_{guilt}(a,t)$ = guilt sentiment score derived from adjuster communication analysis. High guilt sentiment coupled with elevated exception rate indicates guilt-driven overcompliance.

Indicator 4.10 (Emotional Contagion) - Insurance Calibration:

Claims environments exhibit distinctive contagion patterns:

$$EC^{INS}(t) = \rho(\bar{S}_{claimant}(t), \bar{S}_{adjuster}(t+\tau)) \cdot \frac{V_{claims}(t)}{V_{baseline}} \tag{11}$$

Where $\bar{S}_{claimant}$ and $\bar{S}_{adjuster}$ represent average sentiment scores. Volume amplification captures the finding that contagion effects intensify during high-volume periods.

2.3.4 Conditional Probability Update

$$P^{INS}(4.x|3.x) = 0.88 \quad \text{(versus base } P(4.x|3.x) = 0.60) \tag{12}$$

Social influence tactics (Category 3) produce affective vulnerability with high probability in claims contexts, reflecting the professional requirement for emotional openness.

2.4 Category 9 Manifestation: Insurtech Automation Blindness

2.4.1 Theoretical Foundation

Category 9 (AI-Specific Bias Vulnerabilities) addresses human-AI interaction patterns including automation bias, anthropomorphization, and algorithm aversion. Insurance has aggressively adopted AI for underwriting, claims triage, fraud detection, and customer service, creating extensive human-AI interaction surfaces.

The insurtech movement has particularly emphasized "instant" or "touchless" claims processing—claims resolved entirely by AI without human review. While efficient for straightforward claims, this automation creates blind spots for sophisticated fraud that AI systems cannot detect.

2.4.2 Manifestation Characteristics

Insurtech Automation Blindness describes the systematic failure to maintain human oversight of AI-processed claims:

(1) **Straight-Through Processing Trust**: Claims processed "straight through" by AI receive no human review regardless of characteristics. Adversaries who understand AI decision boundaries can construct claims that satisfy AI approval criteria while concealing fraud indicators visible only to human reviewers.

(2) **Deepfake Vulnerability**: AI systems trained on historical fraud patterns cannot detect novel attack vectors. Deepfake technology enabling synthetic video/audio evidence represents a category of fraud for which historical training data provides no preparation. Human reviewers might detect "uncanny valley" artifacts that AI systems miss.

(3) **Fraud Detection Automation Bias**: When AI fraud detection systems flag claims, human reviewers increasingly defer to AI judgment. When AI systems *fail* to flag claims, humans assume legitimacy without independent assessment. The AI becomes both first and last line of defense.

(4) **Speed-Security Tradeoff Pathology**: Insurtech competitive positioning emphasizes processing speed. Security controls that delay instant claims processing face removal pressure. The market expectation of instant gratification creates structural pressure against security-enhancing friction.

(5) **Training Data Poisoning Vulnerability**: AI systems trained on claims data can be manipulated through coordinated submission of fraudulent claims designed to shift decision boundaries. Over time, the AI "learns" that certain fraud patterns are legitimate.

2.4.3 Mathematical Mapping

Insurtech Automation Blindness maps to indicators 9.2, 9.6, and 9.7:

Indicator 9.2 (Automation Bias Override) - Insurance Calibration:

$$\text{Override}_{rate}^{INS} = \frac{N_{\text{human\_review}}}{N_{\text{AI\_approved}}} \cdot W_{complexity} \tag{13}$$

Where $W_{complexity}$ weights by claim complexity (high-value claims, complex loss scenarios should receive higher review rates).

Detection triggers when $\text{Override}_{rate}^{INS} < 0.02$ for complex claim categories (indicating dangerous under-review).

Indicator 9.7 (AI Hallucination Acceptance) - Insurance Calibration:

For claims AI, "hallucination" manifests as confident approval of fraudulent claims:

$$\text{HAR}^{INS} = \frac{N_{\text{AI\_approved} \cap \text{later\_fraud}}}{N_{\text{AI\_approved}}} \cdot \frac{1}{\text{AI\_confidence}_{avg}} \tag{14}$$

The hallucination acceptance rate increases when high-confidence AI approvals correlate with subsequent fraud detection.

Novel Metric: Deepfake Detection Gap

$$DDG = P(\text{fraud}|\text{synthetic\_evidence}) - P(\text{detect}|\text{synthetic\_evidence}) \tag{15}$$

When $DDG > 0.3$, the organization faces significant deepfake vulnerability—synthetic evidence successfully deceives at rates exceeding detection capability.

2.4.4 Conditional Probability Update

$$P^{INS}(9.x|5.x) = 0.82 \quad (\text{versus base } P(9.x|5.x) = 0.55) \tag{16}$$

Cognitive overload strongly predicts automation bias in insurance contexts, as overwhelmed adjusters increasingly defer to AI recommendations without independent assessment.

3 CPIF Intervention Strategy in Insurance

The Cybersecurity Psychology Intervention Framework (CPIF) provides methodology for translating vulnerability assessment into organizational change (Canale, 2025c). Insurance sector application requires adaptation to distinctive organizational structures and resistance patterns.

3.1 Phase 1: Readiness Assessment in Insurance Contexts

3.1.1 Distributed Authority Structures

Insurance organizations exhibit distributed authority across multiple dimensions:

- Corporate headquarters versus regional operations
- Underwriting versus claims versus distribution functions
- Carrier versus agent network relationships
- Legacy organizational units from acquired companies

Readiness assessment must map authority distribution and identify veto points where intervention can be blocked. The readiness function:

$$R_{org} = \min_{u \in \text{Units}} R_u \cdot \prod_{u} \text{alignment}(u, u') \tag{17}$$

Where minimum unit readiness constrains organizational readiness, and cross-unit alignment multipliers capture coordination requirements.

3.1.2 Agent Network Readiness

Agent networks present distinctive readiness challenges:

- Independent agents may refuse participation in carrier security initiatives
- Agent associations may mobilize collective resistance
- Competitive carriers may exploit security requirements as market differentiation

Agent readiness assessment:

$$R_{agent} = \frac{\sum_a R_a \cdot W_{revenue}(a)}{\sum_a W_{revenue}(a)} \tag{18}$$

Revenue-weighted agent readiness captures the practical reality that high-revenue agent resistance has disproportionate impact.

3.1.3 Regulatory Readiness Alignment

Insurance regulators increasingly mandate cybersecurity controls (NAIC Model Law, state-specific requirements). Regulatory mandates can substitute for internal readiness by creating external compliance pressure. The effective readiness:

$$R_{effective} = \max(R_{internal}, R_{regulatory}) \tag{19}$$

Where regulatory pressure can overcome internal resistance.

3.2 Phase 2: Vulnerability-Intervention Matching

3.2.1 Agent Network Deference Interventions (Category 1 Manifestation)

Interventions must address economic authority dynamics:

(1) **Revenue-Neutral Security Requirements**: Design security controls that do not impede sales processes. If security adds friction, agents will resist; if security is invisible, resistance diminishes.

(2) **Contractual Security Requirements**: Embed security requirements in agent contracts with explicit performance standards. Convert security from discretionary to contractual obligation.

(3) **Security as Competitive Advantage**: Position security compliance as market differentiator. Agents who demonstrate security excellence gain access to preferred products, higher commissions, or exclusive territories.

(4) **Collective Accountability**: Implement agent network-level security metrics that create peer pressure. When agent group security performance affects collective benefits, high performers pressure low performers.

3.2.2 Claims Surge Interventions (Category 2 Manifestation)

Interventions must prepare for surge conditions before they occur:

(1) **Pre-Positioned Surge Protocols**: Develop and test security-degradation protocols *before* catastrophe events. Decisions made under crisis pressure are worse than decisions made in advance.

(2) **Tiered Security Degradation**: Define explicit tiers of security relaxation with clear triggers and time limits. Tier 1 might suspend MFA for existing employees; Tier 2 might extend to temporary staff; Tier 3 might reduce verification requirements. Each tier has defined duration and restoration protocol.

(3) **Surge Capacity Investment**: Maintain trained surge capacity that can be deployed without security degradation. The cost of maintaining excess capacity may be less than the fraud losses enabled by security suspension.

(4) **Post-Surge Forensics**: Implement enhanced retrospective review of claims processed during surge periods. Security controls suspended during crisis can be partially compensated through post-event analysis.

3.2.3 Empathy Exploitation Interventions (Category 4 Manifestation)

Interventions must preserve empathic capability while building resistance to exploitation:

(1) **Empathy-Security Integration Training**: Train adjusters to recognize that security protects genuine claimants by preventing fraud that increases premiums. Frame security as empathy toward the broader policyholder community, not opposition to individual claimants.

(2) **Manipulation Recognition Training**: Provide specific training on social engineering tactics that exploit empathy. Adjusters can maintain genuine empathy while recognizing manufactured distress signals.

(3) **Structured Verification Protocols**: Implement verification procedures that operate independently of emotional state. When verification is procedural rather than discretionary, emotional manipulation cannot bypass it.

(4) **Team-Based Claims Processing**: Assign high-risk claims to teams rather than individuals. Team processing distributes emotional load and provides multiple perspectives on manipulation attempts.

(5) **Psychological Support Programs**: Provide support for adjusters experiencing emotional exhaustion. Emotionally depleted adjusters are more vulnerable to manipulation; organizational support for emotional wellbeing is security investment.

3.2.4 Insurtech Automation Blindness Interventions (Category 9 Manifestation)

(1) **Mandatory Human Review Thresholds**: Establish claim characteristics that mandate human review regardless of AI recommendation. High-value claims, novel loss scenarios, and claims with synthetic media indicators require human assessment.

(2) **Deepfake Detection Investment**: Deploy and continuously update deepfake detection capabilities. As synthetic media technology evolves, detection capabilities must evolve correspondingly.

(3) **AI Confidence Calibration**: Require AI systems to report confidence intervals and mandate human review when confidence falls below thresholds. Prevent AI systems from expressing false confidence.

(4) **Adversarial Testing Programs**: Regularly test AI systems with adversarial inputs designed to exploit decision boundaries. Red team exercises for claims AI should include synthetic media attacks.

3.3 Phase 3: Resistance Navigation in Insurance Cultures

3.3.1 Agent Network Resistance

Agent resistance to security requirements follows predictable patterns requiring specific navigation strategies:

Resistance Pattern: "Security Complicates Sales"

Agents perceive security requirements as friction that impedes customer acquisition and policy issuance.

Navigation Strategy: Reframe security as customer protection. Agents who position themselves as protecting customers from fraud and identity theft convert security from obstacle to selling point. Provide agents with customer-facing security messaging that enhances rather than impedes sales conversations.

The value reframe function:

$$V_{security} = V_{protection} \cdot P(\text{customer\_values\_protection}) - C_{friction} \cdot P(\text{customer\_abandons}) \quad (20)$$

When protection value exceeds friction cost, agents will adopt security as sales advantage.

Resistance Pattern: "I've Been Doing This for 30 Years"

Experienced agents resist security requirements by asserting experiential authority superior to corporate security expertise.

Navigation Strategy: Acknowledge experience while introducing novel threat information. "Your experience is invaluable, and the threats have evolved. Here's what we're seeing that didn't exist five years ago." Position security as protecting the agent's book of business and career legacy.

Resistance Pattern: "My Competitor Doesn't Require This"

Agents threaten to redirect business to carriers with less stringent security requirements.

Navigation Strategy: Collective industry action through agent associations and regulatory requirements neutralizes competitive disadvantage. Individual carrier action should emphasize regulatory compliance ("we're all going to be required to do this") and liability protection ("when a breach occurs, security-compliant agents have defense").

3.3.2 Claims Operations Resistance

Claims personnel resistance reflects different dynamics:

Resistance Pattern: "This Slows Down Helping People"

Claims staff perceive security as impediment to their helping mission.

Navigation Strategy: Frame security as protecting the helping mission. Fraud consumes resources that could help genuine claimants. Security enables sustainable helping by preserving organizational capacity. Provide data on fraud losses and their impact on claimant service levels.

Resistance Pattern: "I Can Tell Who's Legitimate"

Experienced adjusters believe their judgment superior to procedural verification.

Navigation Strategy: Acknowledge judgment value while presenting evidence of sophisticated fraud that defeated experienced adjusters. "Your judgment is excellent for 95% of claims. These procedures protect you from the 5% designed to defeat even excellent judgment."

3.4 Phase 4: Working Through in Insurance Cycles

Insurance operates on annual policy cycles with quarterly financial reporting. Intervention timing should align:

- **Q1 (January-March)**: Post-renewal assessment and design. Annual policy renewals complete in January; organizational attention available for security initiatives.
- **Q2 (April-June)**: Pilot implementation. Avoid June-quarter-end financial close.
- **Q3 (July-September)**: Evaluation and refinement. Hurricane season may trigger surge conditions requiring intervention pause.
- **Q4 (October-December)**: Scaled deployment. Complete before year-end renewal season demands organizational attention.

Catastrophe events override intervention scheduling. Intervention activities pause during declared catastrophes and resume when surge conditions normalize.

4 Technical Implementation: OFTLISRV Schema for Insurance

4.1 Data Source Integration

Insurance organizations possess distinctive telemetry sources:

Table 1: Insurance Sector Data Source Mapping

Data Source	CPF Categories	Integration Method
Policy Admin Systems	1.x, 2.x	API/Database query
Claims Management Systems	2.x, 4.x, 9.x	Real-time streaming
Agent Portal Logs	1.x, 3.x	SIEM integration
Call Center Recordings	4.x, 3.x	Speech analytics pipeline
Fraud Detection Systems	9.x, 10.x	Alert correlation
Document Management	9.x (deepfake)	Media analysis pipeline
Catastrophe Declarations	2.x, 10.x	Event trigger integration

4.2 Detection Logic: Sentiment-Behavior Correlation

The Empathy Exploitation detection requires correlation between communication sentiment and security-relevant behavior:

$$D_{4.x}^{INS}(a,t) = \rho(S_{call}(a,t), B_{exception}(a,t+\delta)) \cdot V_{claims}(t) \tag{21}$$

Where:

- $S_{call}(a,t)$ = sentiment score from adjuster a's calls in period t
- $B_{exception}(a,t+\delta)$ = exception/bypass behavior in subsequent period
- δ = lag parameter (typically 1-4 hours)
- $V_{claims}(t)$ = claims volume amplification factor

Elevated correlation indicates that emotional communication content predicts security bypass behavior—the signature of empathy exploitation.

4.3 Mahalanobis Distance Application

The Implementation Companion's Mahalanobis distance formulation applies to insurance observables:

$$A_i^{INS} = \sqrt{(\mathbf{x}_i^{INS} - \boldsymbol{\mu}_i^{INS})^T (\boldsymbol{\Sigma}_i^{INS})^{-1} (\mathbf{x}_i^{INS} - \boldsymbol{\mu}_i^{INS})} \tag{22}$$

For Category 4 (Empathy Exploitation):

$$\mathbf{x}_{4.x}^{INS} = \begin{pmatrix} S_{distress} \\ T_{interaction} \\ R_{exception} \\ N_{verification\_bypass} \end{pmatrix} \tag{23}$$

4.4 Convergence Index: Catastrophe Calibration

The convergence index for insurance incorporates catastrophe-state amplification:

$$CI^{INS} = \prod_{i=1}^{n}(1 + v_i^{INS}) \cdot C(t) \tag{24}$$

Where:

$$C(t) = \begin{cases} 1.0 & \text{normal operations} \\ 1.5 & \text{elevated claims volume} \\ 2.5 & \text{declared catastrophe} \\ 3.5 & \text{catastrophe + media attention} \end{cases} \tag{25}$$

4.5 Response Protocol: Insurance Adaptations

The graduated response function requires insurance-specific thresholds:

$$R^{INS}(s,c,t) = \begin{cases} \text{automatic} & \text{if } s \cdot c > 0.75 \text{ AND NOT CAT\_declared} \\ \text{semi\_auto} & \text{if } s \cdot c > 0.75 \text{ AND CAT\_declared} \\ \text{semi\_auto} & \text{if } 0.5 < s \cdot c \leq 0.75 \\ \text{manual} & \text{if } s \cdot c \leq 0.5 \end{cases} \tag{26}$$

During declared catastrophes, automatic responses convert to semi-automatic to prevent security controls from impeding legitimate surge operations.

5 Case Study: The Deepfake Life Claim

5.1 Incident Overview

In [Date Redacted], a life insurance claim for $2.3 million was submitted to [Carrier Redacted] for an insured who had purportedly died abroad. The claim included video evidence of local officials confirming the death and authorizing body repatriation. The claim was processed through the carrier's AI-assisted claims system and approved within 72 hours. Subsequent investigation revealed the insured was alive and the video evidence was entirely synthetic.

5.2 Attack Sequence

5.2.1 Phase 1: Target Selection (T-90d to T-60d)

Adversaries identified vulnerable claim characteristics:

- High-value policy ($2.3M death benefit)
- Insured with international travel history
- Beneficiary relationship enabling coordination
- Policy in force long enough to avoid contestability

5.2.2 Phase 2: Synthetic Evidence Creation (T-60d to T-30d)

Adversaries created:

- Deepfake video of purported local officials
- Synthetic death certificate with authentic-appearing stamps

- Fabricated hospital records
- Coordinated social media presence suggesting travel and illness

5.2.3 Phase 3: Claim Submission During Surge (T-0)

Claim submitted during Q4 open enrollment period when:

- Claims volume elevated 40% above baseline
- AI approval rate increased to manage volume
- Human review threshold raised from $500K to $1M
- Senior adjusters diverted to enrollment support

The convergence index at submission: $CI^{INS} = 2.8$ (elevated but below critical threshold of 3.0).

5.2.4 Phase 4: AI Processing (T+0 to T+48h)

The AI claims system:

- Verified policy status and beneficiary designation
- Assessed documentation completeness (satisfied)
- Ran fraud detection models (no flags—synthetic evidence not in training data)
- Assigned confidence score of 0.87 (above 0.75 threshold)
- Routed for expedited processing

5.2.5 Phase 5: Human Review Under Pressure (T+48h to T+72h)

The assigned adjuster:

- Carried caseload 2.3x normal (surge conditions)
- Received AI recommendation with 0.87 confidence
- Reviewed video evidence (duration: 47 seconds of 3-minute video)
- Noted "nothing unusual" in review notes
- Approved claim within 18 minutes of assignment

Override rate during this period: 0.018 (well below safe threshold of 0.05).

5.2.6 Phase 6: Payment and Discovery (T+72h to T+180d)

Payment processed. Discovery occurred when:

- Routine reinsurance audit flagged international death documentation inconsistencies
- Investigation revealed insured alive in different country
- Synthetic media analysis confirmed deepfake video
- Beneficiary and insured were co-conspirators

5.3 CPF Analysis

The attack exploited the following category convergences:

Table 2: Deepfake Life Claim: Category Mapping

Category	Manifestation	Exploitation
2.x	Claims Surge	Submission during volume spike
4.x	Empathy (indirect)	Beneficiary distress narrative
5.x	Cognitive Overload	Adjuster caseload 2.3x normal
9.x	Automation Blindness	AI unable to detect synthetic media
9.x	Override Failure	Human review cursory (47 seconds)
10.x	Convergent State	Multiple vulnerabilities aligned

5.4 Lessons for IS-CPF Implementation

(1) **Synthetic Media Detection**: Implement dedicated deepfake detection for claims with video/audio evidence, particularly international death claims.

(2) **Volume-Adjusted Thresholds**: Human review thresholds should decrease, not increase, during surge conditions for high-value claims.

(3) **Review Quality Metrics**: Monitor not just review occurrence but review duration and depth. 47-second review of $2.3M claim indicates inadequate assessment.

(4) **Override Rate Monitoring**: Alert when human override rates fall below category-specific safe thresholds.

(5) **AI Training Data Currency**: Continuous update of fraud detection training data to include emerging attack vectors (synthetic media, coordinated documentation fraud).

6 Conclusion

6.1 Theoretical Contribution

This paper demonstrates that insurance sector cybersecurity psychology does not require framework extension but framework calibration. Agent Network Deference, Claims Surge Collapse, Empathy Exploitation, and Insurtech Automation Blindness represent sector-specific manifestations of the Core 10 CPF taxonomy. This finding further validates the theoretical completeness of the original framework while enabling precise sector adaptation.

The insurance sector presents a distinctive challenge: professional effectiveness requires emotional openness that simultaneously creates exploitable vulnerability. The IS-CPF addresses this challenge by providing detection and intervention capabilities that preserve empathic function while building resistance to exploitation.

6.2 Mathematical Integrity

By mapping insurance phenomena to existing categories, IS-CPF preserves the Implementation Companion's Bayesian network architecture. Detection functions require recalibration (particularly for sentiment-behavior correlation in empathy exploitation), but not restructuring. Organizations implementing standard CPF can deploy IS-CPF through parameter adjustment.

6.3 Operational Resilience

The insurance sector requires frameworks that maintain security during surge conditions when standard approaches fail. The IS-CPF provides:

- Pre-positioned surge protocols for predictable security degradation
- Catastrophe-adjusted thresholds that adapt to operational reality
- Post-surge forensic capabilities that compensate for real-time control suspension

6.4 Validation Roadmap

Future work will empirically validate IS-CPF calibrations through:

(1) Pilot implementations with partner carriers

(2) Correlation analysis between IS-CPF scores and claims fraud rates

(3) Catastrophe-period analysis comparing protected and unprotected operations

(4) Agent network security metric development and validation

The insurance sector's combination of sensitive data, distributed operations, empathy requirements, and surge dynamics presents comprehensive testing of CPF's adaptability. Successful deployment validates the framework's applicability across diverse organizational contexts.

Note on AI-Assisted Composition

This manuscript presents the original theoretical framework and intellectual contributions of the author. In the composition process, the author utilized a large language model as an auxiliary tool for stylistic refinement and formatting consistency. The core ideas, the IS-CPF architecture, the theoretical integration, and the strategic analysis are solely the product of the author's expertise. The author is entirely responsible for the accuracy and integrity of the published content.

Acknowledgments

The author acknowledges the foundational work in insurance operations research, claims psychology, and fraud detection upon which IS-CPF builds.

References

Bion, W. R. (1961). *Experiences in groups*. London: Tavistock Publications.

Bowlby, J. (1969). *Attachment and Loss: Vol. 1. Attachment*. New York: Basic Books.

Canale, G. (2025a). The Cybersecurity Psychology Framework: A Pre-Cognitive Vulnerability Assessment Model. *CPF Technical Report Series*.

Canale, G. (2025b). Operationalizing the Cybersecurity Psychology Framework: A Systematic Implementation Methodology. *CPF Technical Report Series*.

Canale, G. (2025c). The Cybersecurity Psychology Intervention Framework: A Meta-Model for Addressing Human Vulnerabilities. *CPF Technical Report Series*.

Canale, G. (2025d). Financial Services Cybersecurity Psychology Framework (FS-CPF v2.0). *CPF Technical Report Series*.

CNA Financial Corporation. (2021). *Form 8-K: Cybersecurity Incident Disclosure*. SEC Filing.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Klein, M. (1946). Notes on some schizoid mechanisms. *International Journal of Psychoanalysis*, 27, 99-110.

Menzies Lyth, I. (1960). A case-study in the functioning of social systems as a defence against anxiety. *Human Relations*, 13, 95-121.

Milgram, S. (1974). *Obedience to authority*. New York: Harper & Row.