IRMMF Ultimate v5
Methodology Paper (Extended)

Structural Integrity, Control Lag & Friction Model
Multi-Role Evidence-Gated Assessment with Advanced Scoring, Diagnostics, and Advisory Triggers

Version: v5.1
Date: January 08, 2026
Document type: Methodology specification for implementation
Audience: Product, engineering, data science, and advisory teams

Executive Summary

IRMMF Ultimate v5 measures organizational insider-risk readiness as a defensible, multi-perspective system rather than a checklist. It produces a 9-axis security fingerprint and three headline metrics: Operational Reality (OR), Structural Integrity (SI), and Confidence (CI). The scoring engine is evidence-gated, scope-aware, and penalizes weak links through harmonic aggregation and critical-control floor penalties. Two additional drivers of insider risk—Operational Friction (F) and Risk Velocity (W)—are operationalized using measurable indicators, including a Control Lag Index during organizational change. The methodology also quantifies cross-functional disagreement as a fracture signal and provides advisory triggers and roadmaps.

- OR answers: How strong is day-to-day execution (regardless of policy)?
- SI answers: Does the organization do what it claims (alignment of reality to posture)?
- CI answers: How much should we trust the scores (evidence, coverage, and agreement)?
- Fractures highlight: misalignment, control lag, friction-driven bypass, and weak-link control failures.

1. Scope, Non-Goals, and Design Principles

1.1 Scope

This methodology specifies how to design, administer, score, and report a multi-role insider risk assessment aligned to an IRMMF capability model. It is suitable for tool automation and for advisory-led assessments with evidence collection.

1.2 Non-Goals

- Predicting individual insider behavior or intent.
- Replacing legal advice, HR disciplinary procedures, or formal audit attestations.
- Providing exact loss predictions; cost-of-inaction outputs are scenario-based ranges.

1.3 Design principles

- Evidence before elevation: high scores require operational proof or independent verification.
- Multi-role truth: critical conclusions require more than one perspective; disagreement is a risk signal.
- Conservative aggregation: weak links materially reduce scores; do not rely on simple averages.
- Explainability: advanced math remains internal; reports must stay interpretable.
- Ethical legality first: monitoring must be proportionate, transparent, and defensible.

2. Core Constructs and Terminology

Gap: A missing capability or control expected for the organization’s objectives and risk profile.

Fracture: A misalignment likely to fail under stress (e.g., strong policy but weak execution; strong execution but weak defensibility; high role disagreement; high change with long adaptation delay).

Stated Posture (SP): Declared and evidenced posture of mandate, guardrails, and targeting (primarily Governance, Legal/Privacy, Visibility).

Operational Reality (OR): Evinced operational strength in day-to-day practice (primarily Execution, Technical Orchestration, Resilience).

Structural Integrity (SI): Alignment ratio of OR to SP.

Confidence Index (CI): Trustworthiness of results based on evidence tiers, recency, scope coverage, respondent coverage, and completion depth.

Misalignment (MIS): Measured disagreement between roles; reported as a fracture signal.

3. Measurement Model: 9 Axes

Each question maps to at most two axes to avoid double counting. Axis definitions are intentionally tight:

Axis	Label	Operational definition (what is measured)
G	Governance - The Mandate	Board sponsorship, funding, decision rights, operating model, accountability, oversight cadence.
E	Execution - The Habit	Operational cadence and adherence; processes happen by default and are enforced.
T	Technical Orchestration - The Scale	Automation/integration: telemetry pipelines, control enforcement, detection engineering, case workflow tooling (AI/ML optional).
L	Legal & Privacy - The Guardrail	Defensibility: lawful basis, DPIA/LIA, transparency, minimization, retention, access auditability, proportionality.
H	Human Sentiment - The Heart	Psychological safety, speak-up climate, perceived fairness, manager intervention readiness, trust.
V	Visibility - The Target	Crown jewels, data flows,

		critical role mapping, risk concentration, insider attack paths.
R	Resilience - The Recovery	Containment speed, investigation readiness, continuity for sabotage, recovery playbooks, learning loops.
F	Operational Friction - The Efficiency	Measured friction and bypass pressure: exceptions, access lead time, workarounds, shadow IT drivers, control usability.
W	Risk Velocity - The Stability (Control Lag lens)	Change rate multiplied by adaptation delay: how fast risk changes vs controls/policies/monitoring adapt.

4. Data Collection Architecture

4.1 Role Discovery

The assessment begins with a Role Discovery module to determine which functions exist and are in scope, including outsourced responsibilities. The tool uses this to assign role-specific packs and to compute Role Coverage for CI.

- Roles present vs outsourced (e.g., SOC, HR, DPO, IAM, Data Governance).
- Operating model (centralized, federated, hybrid).
- Tool ownership (who runs monitoring, who owns case management).
- Investigation authority (who can initiate, who can approve).

4.2 Multi-role Packs

Packs are targeted modules assigned to domain experts. Standard v5 packs:

- Exec Pack: G, V, W
- Security Pack: E, T, V, R
- HR Pack: H, E, F
- Legal/Privacy Pack: L, G
- Data/Business Pack: V, F, W (crown jewels and operational reality)
- IT/IAM Pack: E, T, F, R (joiner/mover/leaver and access workflows)
- Optional: OT/Physical, Vendor/Procurement, Internal Audit/Risk packs

4.3 Adaptive Branching and Shadow Probes

Adaptive branching uses tiers to reduce fatigue while preserving detection of hidden maturity.

```
Tiers:
Tier 1: Strategic Gatekeepers (3-5 questions per domain/pack)
Tier 2: Formal vs Informal validation
Tier 3: Technical nuance / orchestration details
Tier 4: Evidence & verification prompts

Shadow Probes:
If Tier 1 is low but early signals show high execution/visibility/resilience,
run a minimal Tier 2 subset to detect Shadow Mature patterns.


Algorithm (high level):
for each (domain, pack):
    ask Tier1 questions
    if Tier1Score < StopThreshold:
        if ShadowSignal(pack, domain) == true:
            ask ShadowProbeSubset(Tier2)
        else:
            stop deepening and mark Foundational Gap
    else:
        proceed Tier2 -> Tier3 -> Tier4 as applicable
```

4.4 Evidence Capture Workflow

Evidence capture is structured and auditable. Each evidence item is tagged to the question and includes type, recency, scope, and sampling metadata.

- Evidence type: policy/procedure, system config, log extract, ticket export, case record, audit report, exercise result.
- Recency: evidence age in days; operational items must be recent.
- Scope: which population/systems are covered.
- Sampling: number of samples, selection method, and period covered.
- Access: who provided evidence and who approved its use.

5. Evidence Model

5.1 Evidence tiers and caps

Tier	Name	Examples	Default cap
A	Assertion	Respondent claim only	2
B	Artifact	Policy/process, configuration screenshot	3
C	Operational proof	Logs, tickets, case records, metrics	4
D	Independent verification	Audit sampling, exercise, external	4

		attestation	

5.2 Sampling adequacy rules

Sampling adequacy (SA) prevents cherry-picking for Tier C evidence.

```
Default SA rules for operational items:
n < 3 samples  -> SA = 0.40
n = 3 samples  -> SA = 0.70
n = 10 samples -> SA = 1.00
3 < n < 10     -> linear interpolation between 0.70 and 1.00
```

5.3 Negative controls (anti-gaming)

A small set of proof-based 'negative controls' is included to detect inflated claims. Failure caps affected scores and reduces CI.

- Show recent (redacted) insider-relevant triage records and outcomes.
- Show monitoring data retention + access auditability evidence.
- Show transparency notices or employee communications about monitoring.
- Show offboarding exceptions and approvals.
- Show evidence that case management access is role-restricted and audited.

6. Scoring Specification

This section defines calculations for question-level scoring, role aggregation, capability aggregation, axis scoring, and headline metrics.

6.1 Question-level adjusted score

```
Inputs per answer (q,r):
BS(q,r) ∈ {0,1,2,3,4}
ET(q,r) ∈ {A,B,C,D}
RD(q,r) in days
SC(q,r) ∈ [0,1]
SA(q,r) ∈ [0,1]
AC(q,r) ∈ [0,1] (default 1.0)


Evidence quality:
EQ_base(A)=0.40, EQ_base(B)=0.70, EQ_base(C)=0.90, EQ_base(D)=1.00
EQ(q,r) = EQ_base(ET) × SA(q,r)


Recency factor RF (default):
Operational: RD≤90 ->1.00; 91-180->0.85; 181-365->0.65; >365->0.40
Governance:   RD≤180->1.00; 181-365->0.85; 366-730->0.65; >730->0.45


Evidence caps:
Cap(A)=2; Cap(B)=3; Cap(C/D)=4
```

```
BS_cap(q,r)=min(BS(q,r), Cap(ET))


Adjusted score:
AS(q,r)=BS_cap(q,r) × min(EQ(q,r), RF(q,r), SC(q,r), AC(q,r))
NAS(q,r)=AS(q,r)/4
```

6.2 Role-weighted consensus and misalignment

Role weights RW(q,r) are configured per question type and sum to 1 across roles answering q. For example: legal defensibility questions weight Legal/DPO higher; culture questions weight HR higher; technical questions weight Security/IT higher.

```
Consensus:
CNS(q)=Σ RW(q,r)×NAS(q,r)

Misalignment (choose one global method):
Option A: MIS(q)=min(1, stdev({NAS(q,r)})/0.25)
Option B: MIS(q)=max({NAS(q,r)})-min({NAS(q,r)})

Evidence confidence:
EC(q)=Σ RW(q,r)×(EQ×RF×SC)
```

6.3 Capability aggregation (mean + harmonic + floor penalties)

```
For node X with questions QX and criticality weights CW(q):

WM(X)=Σ CW(q)×CNS(q) / Σ CW(q)

WHM(X)=Σ CW(q) / Σ [CW(q)/max(ε,CNS(q))], ε=0.05

Critical-control penalties:
For each critical q in CQX with threshold TH(q):
Pen(q)=0 if CNS(q)≥TH(q) else α(q)×(TH(q)-CNS(q))
FP(X)=Σ Pen(q)

Final score:
CS(X)=clamp(0.6×WHM(X)+0.4×WM(X)-FP(X),0,1)
```

6.4 Axis scoring

```
Axis mapping:
Each question maps to at most 2 axes with weights AXW(q,a) (sum to 1 over
mapped axes).
WA(q,a)=CW(q)×AXW(q,a)

AxisScore(a)=Σ WA(q,a)×CNS(q) / Σ WA(q,a)
AxisConf(a)= Σ WA(q,a)×EC(q)  / Σ WA(q,a)
AxisMIS(a)=  Σ WA(q,a)×MIS(q) / Σ WA(q,a)
```

6.5 Headline metrics (OR, SP, SI, CI)

```
Defaults:
SP=0.35×G+0.35×L+0.30×V
OR=0.35×E+0.35×T+0.30×R
(letters are AxisScore values)

Confidence attenuation (optional):
SP_adj=SP×min(1, AvgConf(G,L,V)+0.2)
OR_adj=OR×min(1, AvgConf(E,T,R)+0.2)

SI ratio:
SI=clamp(OR_adj/(SP_adj+ε),0,1), ε=0.05

CI:
CI=0.5×EvidenceCoverage+0.3×RoleCoverage+0.2×DepthCoverage
```

6.6 End-to-end scoring pipeline (implementation pseudocode)

```
Pipeline:
1) RoleDiscovery() -> present roles, outsourced roles, required packs
2) CollectResponsesByPack() -> BS, ET, RD, SC, SA, evidence metadata
3) For each (q,r): compute AS(q,r), NAS(q,r)
4) For each q: compute CNS(q), MIS(q), EC(q)
5) For each node X: compute CS(X), EC(X), MIS(X)
6) For each axis a: compute AxisScore(a), AxisConf(a), AxisMIS(a)
7) Compute SP, OR, SI, CI
8) Compute archetype probabilities and fracture list
9) Apply sector overlay multipliers (if enabled) and recompute affected
aggregates
10) Generate executive + operational reports and service triggers
```

7. Operationalizing F and W (Algorithms)

7.1 F: Friction indicators and normalization

F is derived from measurable operational indicators, then mapped into question responses or computed directly as an axis sub-score.

```
Example normalization approach (per indicator i):
- Compute raw metric value m_i (e.g., mean access approval time in days).
- Convert to percentile p_i relative to peer benchmark or historical baseline.
- Map to [0,1] risk score: fric_i = p_i (higher percentile = more friction).
- Combine indicators with weights: F_raw = Σ w_i × fric_i.
- Convert to AxisScore(F) via evidence-gated questions that validate
measurement, ownership, and remediation loop.
```

- Suggested friction indicators: exceptions/overrides per 100 users, average access lead time, break-glass frequency, number of unapproved SaaS apps, DLP block rate with override rate, helpdesk tickets tagged security-friction.

- Use evidence: ticket exports, IAM workflow metrics, CASB discovery, SIEM tags, service desk analytics.

7.2 W: Control Lag Index computation

W is operationalized through Control Lag: Change Rate multiplied by Adaptation Delay. Both components are normalized (0-1) and calibrated.

```
Step 1: ChangeRateIndex (CRI) in [0,1]
Inputs may include:
- number of major org events in last 12 months (M&A, layoffs, reorgs, cloud
migrations, outsourcing)
- percent workforce churn
- percent systems/applications changed or onboarded
Normalize to CRI using sector/size calibration bands.

Step 2: AdaptationDelayIndex (ADI) in [0,1]
Inputs may include:
- median days to update policies/standards after change
- median days to update JML / access models
- median days to onboard new systems into telemetry monitoring
- median days to update training/playbooks
Normalize to ADI using calibration bands.

ControlLagIndex = CRI × ADI
W axis score is the inverse of lag (or directly derived via questions):
W_risk = ControlLagIndex
W_score = 1 - W_risk
```

8. Sector Overlays (Weight Multipliers)

Sector overlays adjust criticality weights or axis weights via documented multipliers. Overlays must be disclosed in reports.

```
Overlay implementation:
For sector s, define multiplier M_s on either:
- question criticality: CW'(q)=CW(q)×M_s(q)  (preferred for specificity), or
- axis weights: WA'(q,a)=WA(q,a)×M_s(a)

Constraints:
- Multipliers bounded (e.g., 0.75 to 2.0)
- Total weight is re-normalized per aggregate so scores remain in [0,1]
- Overlay version and rationale are stored and reported
```

9. Fracture Analytics and Cascade Model

9.1 Fracture identification rules

- SI below threshold (e.g., <0.7) indicates structural lopsidedness; <0.5 is major fracture.

- High MIS at critical nodes indicates operating model fracture.
- High ControlLagIndex indicates transformation fracture.
- High friction indicators plus low H indicate bypass and resentment risk (Toxic Enforcer patterns).
- Critical-control failures (CNS<threshold) trigger floor penalties and immediate remediation flags.

9.2 Dependency graph and 'effective capability'

Cascade analysis is implemented as a transparent dependency graph rather than opaque prediction. It estimates how weaknesses in enabling axes reduce effective outcomes in dependent axes.

```
Example dependency edges (illustrative):
G -> T (investment governance)
L -> T (defensible monitoring)
V -> T (targeted instrumentation)
H -> E (adherence and reporting)
F -> E (bypass pressure)
W -> (E,T,R) (change destabilizes operations)

One implementation:
Let s be vector of axis scores in [0,1]. Let D be a matrix of influence weights
in [0,1].
Compute weakness vector w = 1 - s.
Propagate: w_eff = clamp(w + D·w, 0, 1)
Effective scores: s_eff = 1 - w_eff
Use s_eff for 'effective capability' reporting and roadmap prioritization (not
for headline SI unless explicitly chosen).
```

10. Cost of Inaction (Scenario-Based Ranges)

Cost-of-inaction is communicated as ranges with explicit assumptions. Business owners supply impact inputs; the model supplies likelihood bands based on V and W.

```
Scenario approach:
1) Select scenario family (data theft, sabotage, fraud, external collusion).
2) Determine crown jewel exposure band using V (low/medium/high exposure).
3) Determine volatility band using ControlLagIndex (W).
4) Use a likelihood band (e.g., 1-3%, 3-8%, 8-15%) calibrated per sector/size.
5) Use business-provided impact range (e.g., €0.5M-€3M).
6) ExpectedLossRange = LikelihoodBand × ImpactRange.
7) Proposed control roadmap estimates reduction in exposure/likelihood bands;
publish delta as 'value at risk reduced' (range).
```

11. Archetype Engine (Probabilistic)

Archetypes are computed as probabilities based on axis patterns, SI, and misalignment. Hybrids are allowed and marked tentative when CI is low.

```
Example archetype scoring (normalize to probabilities):
PaperDragon = 0.5×SP + 0.5×(1-OR) + 0.5×(1-SI)
ShadowMature = 0.5×(1-SP) + 0.5×OR + 0.5×(1-SI) + 0.3×MIS_overall
BlindTitan = 0.3×G + 0.3×T + 0.4×(1-V) + 0.3×(1-H)
ToxicEnforcer = 0.4×T + 0.4×L + 0.6×(1-H)
BalancedOperator = 0.4×OR + 0.2×SP + 0.4×SI + 0.2×H + 0.2×V

Confidence gating:
if CI < 0.6: archetype label = 'tentative'
```

12. Reporting and Service Triggers

12.1 Executive outputs

- 9-axis fingerprint with confidence shading.
- OR, SI, CI headlines (always publish all three).
- Top 5 fractures (misalignment, control lag, friction-bypass, critical failures).
- Archetype mix (probabilistic) and narrative.
- Roadmap triggers and sequencing recommendations.

12.2 Operational outputs

- Node-level scores: CS(X), EC(X), MIS(X) for domains/functions/sub-capabilities.
- Evidence backlog: where ET is below required for the current score.
- Control lag report: high ADI drivers and lag reducers.
- Friction report: top bypass drivers and ‘fast UX wins’.
- Role disagreement map: where Security/HR/Legal disagree.

12.3 Trigger rules (examples)

```
Example triggers:
If L < 0.55 OR AxisMIS(L) > 0.35 -> Legal & Privacy Defensibility Audit (+
transparency remediation)
If H < 0.55 -> Cultural Health & Behavioral Design Workshop (+ manager
intervention pathways)
If V < 0.55 -> Data Discovery & Crown Jewel + Critical Role Mapping
If F < 0.55 -> Security UX & Operational Optimization (reduce bypass/overrides)
If ControlLagIndex high -> Control Lag Reduction Sprint (policy, telemetry,
access, playbooks)
If R < 0.55 OR SI < 0.70 -> Insider Resilience Playbooks + Tabletop Exercise
If MIS_overall high -> Operating Model Alignment Workshop
```

13. Validation, Calibration, and Governance

- Pilot calibration on a representative sample of organizations (5-10) to tune multipliers, recency curves, penalties, and role weights.
- Inter-rater reliability checks: two assessors score the same evidence set; tune rubrics until acceptable agreement.
- Version control: all rubrics, thresholds, weights, and overlays are versioned and included in report metadata.
- Auditability: store a scoring trace that can reproduce every computed number from raw inputs.

14. Legal, Privacy, and Ethics Requirements

The methodology explicitly tests lawful basis, transparency, minimization, retention, access auditing, and proportionality. Protected reporting (whistleblowing) is treated as a protective mechanism, not automatically as a threat signal. Assess confidentiality controls, anti-retaliation enforcement, and escalation governance.

- Separation of duties for monitoring access (least privilege + audited access).
- Clear employee communications about monitoring scope and purpose.
- Documented DPIA/LIA where applicable and retention schedules.
- Controls to prevent misuse of monitoring data against whistleblowers.

Appendix A: Default Parameter Table

Parameter	Default	Notes
EQ_base tiers	A=0.40, B=0.70, C=0.90, D=1.00	Multiply by SA; tune during calibration.
Evidence caps	A cap=2, B cap=3, C/D cap=4	Prevents policy-only inflation.
Recency curves	Operational: 90/180/365; Governance: 180/365/730	RF decreases beyond windows.
Aggregation blend	CS=0.6×WHM + 0.4×WM - FP	Penalizes weak links; avoid pure averages.
Epsilon	ε=0.05	Avoids division by zero.
SI thresholds	SI<0.7 lopsided; SI<0.5 major fracture	Use alongside OR and SP.
CI weights	0.5 evidence, 0.3 role, 0.2 depth	Publish CI with all key scores.
MIS method	stdev/0.25 or range	Select one method globally for consistency.
Overlay bounds	0.75 to 2.0	Keep multipliers bounded and disclosed.

Appendix B: Worked Example (illustrative)

This example demonstrates how evidence gating and aggregation work. Numbers are illustrative.

```
Assume question q answered by Security and Legal:
Security: BS=4, ET=C, SA=0.7, RD=60, SC=0.8
Legal:    BS=3, ET=B, SA=1.0, RD=120, SC=1.0
Operational question -> RF(Sec)=1.0, RF(Legal)=0.85
Caps: Cap(C)=4, Cap(B)=3 -> BS_cap(Sec)=4, BS_cap(Legal)=3

EQ(Sec)=0.90*0.7=0.63
EQ(Legal)=0.70*1.0=0.70

AS(Sec)=4*min(0.63,1.0,0.8)=4*0.63=2.52 -> NAS=0.63
AS(Legal)=3*min(0.70,0.85,1.0)=3*0.70=2.10 -> NAS=0.525

Role weights RW: Sec=0.7, Legal=0.3
CNS=0.7*0.63 + 0.3*0.525 = 0.5985
MIS(range)=0.63-0.525=0.105
```