Ethical AI Evaluation Report

Project: test use case

Category: N/A

Generated: December 29, 2025

Questionnaire: general-v1

Quick Navigation

Dashboard   Risks   Tensions   Recommendations   Principle Analysis   Methodology

Appendix

Overall Risk Summary

0.25/4.0

Low Risk

Based on 3 evaluator submission(s)

Team Completion

0/3

Experts Submitted

3 started, 3 assigned

Ethical Tensions

2

Total Tensions Identified

0 Accepted, 0 Under Review, 0 Disputed

Evidence Coverage

0.0%

Tensions with Evidence

0 of 2 tensions
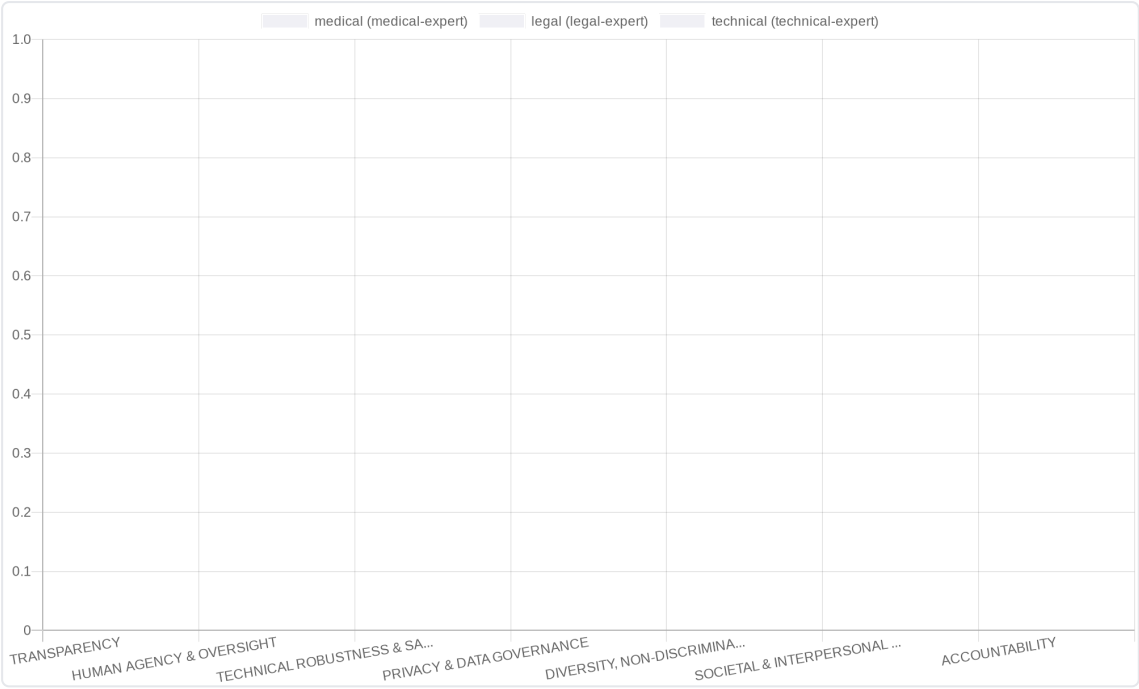
Ethical Principles Score Overview

Scale (0-4): Higher Score = Higher Risk

- **0.0-1.0:** Low risk / Good
- **1.0-2.0:** Moderate risk
- **2.0-3.0:** High risk
- **3.0-4.0:** Critical risk

Note: Scores are canonical from scores collection; Gemini does not compute scores.

Principle × Evaluator Score Matrix

medical (medical-expert) legal (legal-expert) technical (technical-expert)

```
1.0

0.9

0.8

0.7

0.6

0.5

0.4

0.3

0.2

0.1

0
    TRANSPARENCY        TECHNICAL ROBUSTNESS & SA...   DIVERSITY, NON-DISCRIMINA...        ACCOUNTABILITY
         HUMAN AGENCY & OVERSIGHT    PRIVACY & DATA GOVERNANCE    SOCIETAL & INTERPERSONAL ...
```

Heatmap Legend

Cells show evaluator's average risk score per principle (0-4 scale).

N/A = evaluator did not submit responses for this principle.

Only evaluators with submitted responses (status="submitted") are shown.

Note: Evaluators are derived from scores collection - no hardcoded "Expert 1/2" labels.

Evidence Coverage

With Evidence **No Evidence**

0

100

Coverage Explanation

Percentage of tensions that have at least one evidence item attached.

With Evidence: Tensions with evidence count > 0

No Evidence: Tensions with no evidence attached

Executive Summary

- Overall ethical risk level: 0.25/4.0

- 0 of 3 assigned experts completed evaluations

- 2 ethical tensions identified

Top Risk Drivers

Note: AvgRisk is derived from scores collection and mapped to principle tags. Only questions with submitted responses are included.

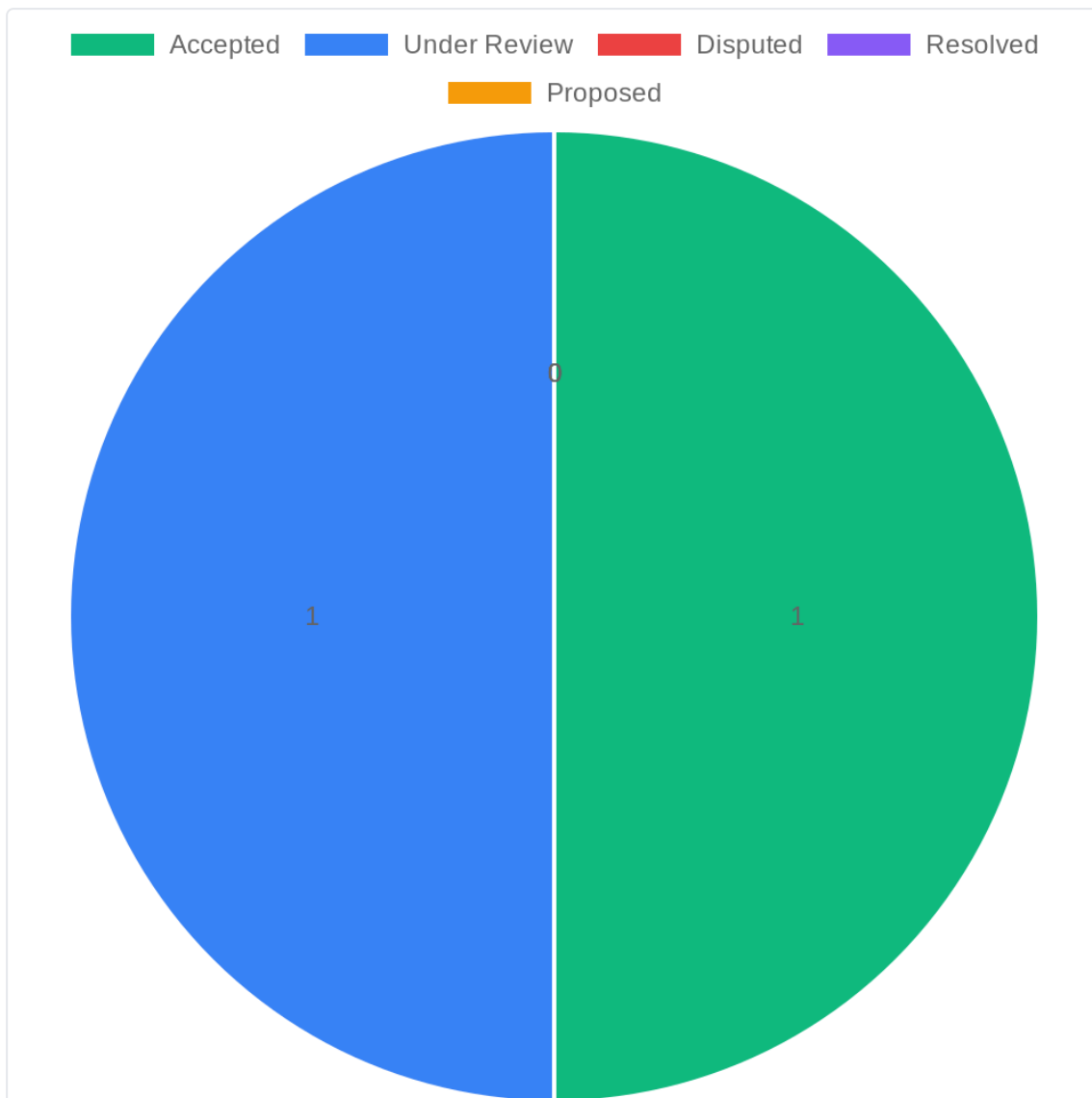| Rank	Principle	Question	Avg Risk Score	Roles	Answer Excerpt
1	SOCIETAL & INTERPERSONAL WELL-BEING	693c6b9566f16d2d496635a5	**3.00**	medical-expert, legal-expert, technical-expert	No text answer provided
2	TRANSPARENCY	693c6b9366f16d2d4966356e	**0.00**	medical-expert, legal-expert, technical-expert	No text answer provided
3	TRANSPARENCY	693c6b9366f16d2d49663576	**0.00**	medical-expert, legal-expert, technical-expert	No text answer provided
4	HUMAN AGENCY & OVERSIGHT	693c6b9366f16d2d4966357c	**0.00**	medical-expert, legal-expert, technical-expert	No text answer provided
5	HUMAN AGENCY & OVERSIGHT	693c6b9466f16d2d49663583	**0.00**	medical-expert, legal-expert, technical-expert	No text answer provided
6	TECHNICAL ROBUSTNESS & SAFETY	693c6b9466f16d2d4966358a	**0.00**	medical-expert, legal-expert, technical-expert	No text answer provided

Rank	Principle	Question	Avg Risk Score	Roles	Answer Excerpt
7	PRIVACY & DATA GOVERNANCE	693c6b9466f16d2d49663591	**0.00**	medical-expert, legal-expert, technical-expert	No text answer provided
8	PRIVACY & DATA GOVERNANCE	693c6b9566f16d2d49663598	**0.00**	medical-expert, legal-expert, technical-expert	No text answer provided
9	DIVERSITY, NON-DISCRIMINATION & FAIRNESS	693c6b9566f16d2d4966359e	**0.00**	medical-expert, legal-expert, technical-expert	No text answer provided
10	SOCIETAL & INTERPERSONAL WELL-BEING	693c6b9666f16d2d496635a8	**0.00**	medical-expert, legal-expert, technical-expert	No text answer provided

Analysis

← Back to Dashboard

Ethical Tensions

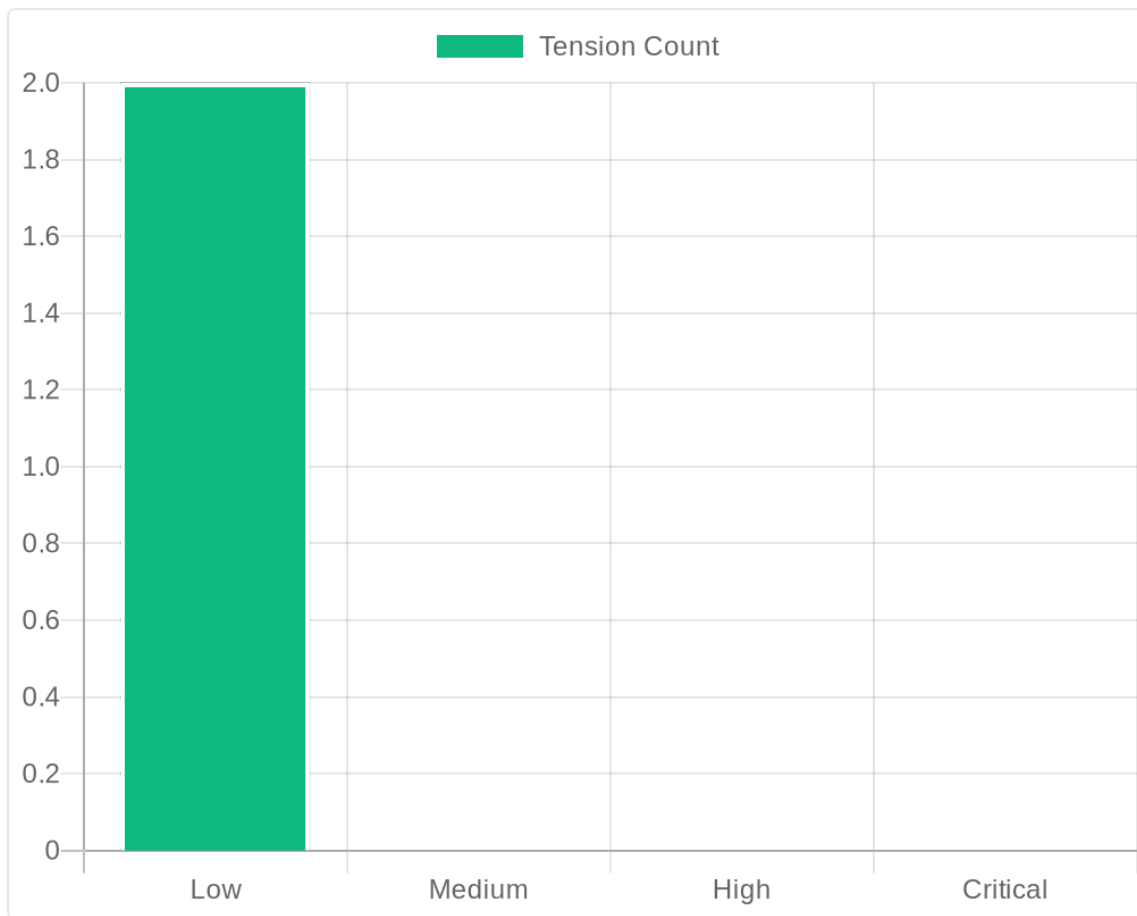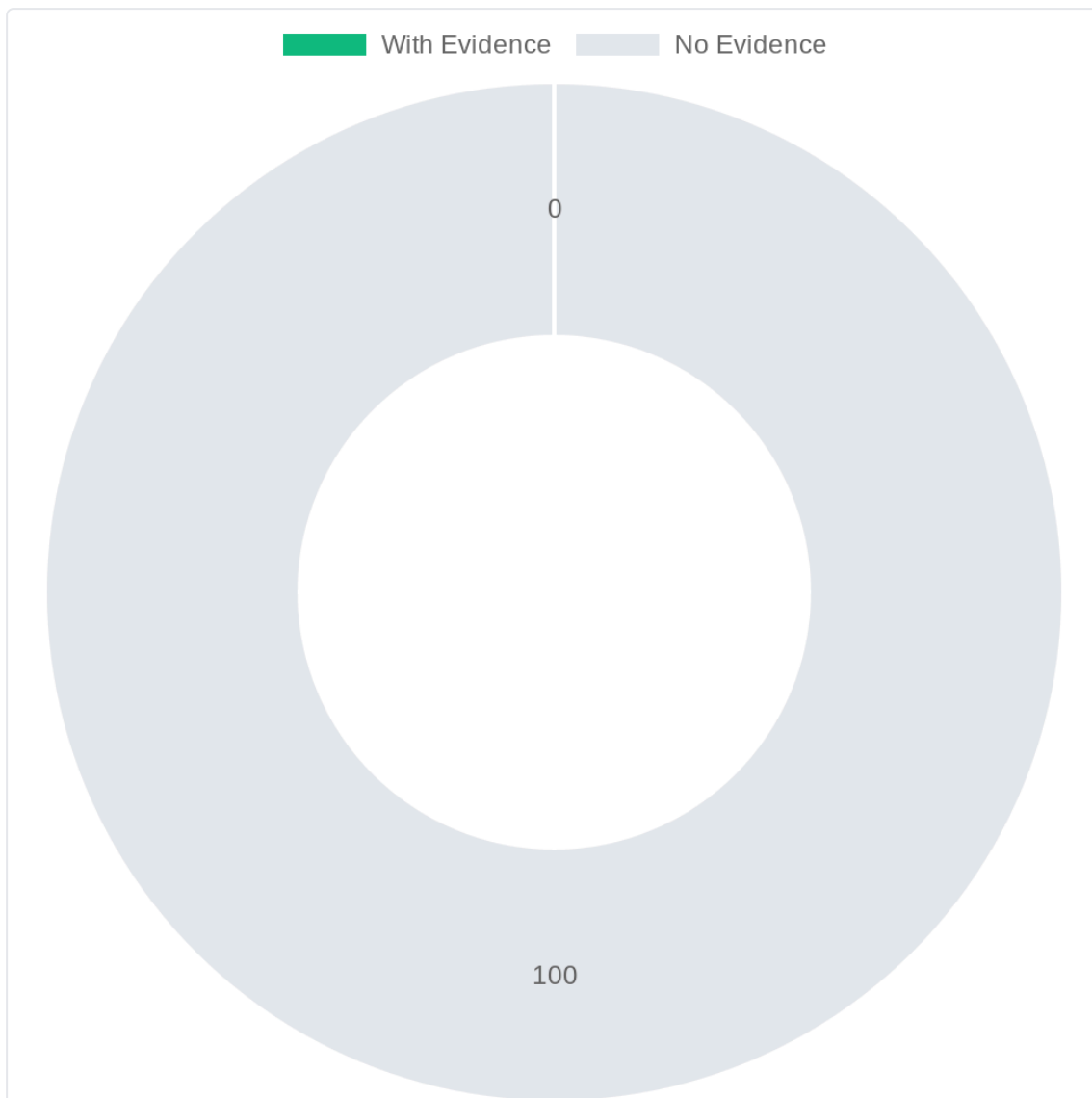# Tension Review State Distribution

Review State Explanation

Review state is computed from votes/consensus rules:

- **Proposed:** Newly identified, awaiting review
- **Under Review:** Active discussion in progress
- **Accepted:** Consensus reached, mitigation approved
- **Disputed:** Conflicting opinions, requires resolution
- **Resolved:** Tension addressed and closed

Tension Severity Distribution

Evidence Coverage

With Evidence   No Evidence

0

100

Coverage Explanation

Percentage of tensions that have at least one evidence item attached.

With Evidence: Tensions with evidence count > 0

No Evidence: Tensions with no evidence attached

Conflict	Severity	Review State	Evidence	Consensus	Comments	Created By
Transparency ↔ Safety	high	Under review	**No evidence attached**	1 agree / 1 disagree (50.0% agree)	0	unknown
Privacy ↔ Accountability	high	Accepted	**No evidence attached**	2 agree / 0 disagree (100.0% agree)	0	unknown

Detailed Tension Analysis

← Back to Dashboard

Principle-by-Principle Analysis

← Back to Dashboard

Methodology & Data Sources

This report is generated using the Z-Inspection methodology for ethical AI evaluation.

Data Sources

- **responses collection:** All expert answers and qualitative context
- **scores collection:** Canonical computed metrics (ONLY source of quantitative scores)
- **tensions collection:** Ethical tensions, claims, evidence, mitigations, and consensus
- **projectassignments collection:** Expert assignments and participation tracking

> **IMPORTANT:** Quantitative scores come from the scores collection and are NOT computed by Gemini AI. All numeric metrics are deterministic and traceable to MongoDB data.

Risk Score Mapping

- **Score 4:** Best/Low risk
- **Score 3:** Good/Acceptable risk
- **Score 2:** Moderate risk
- **Score 1:** High risk
- **Score 0:** Critical/Worst risk

Risk Percentage Formula: Percentage of responses with score < 3.0 (indicating risk)

Appendix

Evaluators

Only evaluators with submitted responses are listed below:

Name	Role	Email	Status
medical	medical-expert	medical@zinspection.com	Submitted
legal	legal-expert	legal@zinspection.com	Submitted
technical	technical-expert	technical@zinspection.com	Submitted

Limitations

No specific limitations identified.

Glossary

- **Risk Score:** 0-4 scale where 4 = best/low risk, 0 = worst/high risk
- **Risk %:** Percentage of responses with score < 3.0
- **Safe %:** Percentage of responses with score >= 3.0
- **Severity Levels:** Critical, High, Medium, Low (based on avgRiskScore)
- **Evidence:** Policy documents, test results, user feedback, logs, incidents, or other supporting materials
- **Review State:** Proposed, Under Review, Accepted, Disputed, or Resolved

test use case