

Test case spec template

A reusable spec for single-turn and multi-turn tests.

QA

Designer

Developer

TEMPLATE

```
id:
title:
risk:
traces_to: []
severity: low | medium | high
tags: []
preconditions: {channel: , language: ,
persona: }
fixtures: {tools: {}, retrieval: []}
input: [{user: ""}] # or turns: with
expectations
expected: {must: [], must_not: [], should:
[]}
oracle: {assertions: [], judge: {rubric: ,
min_score: }}
pass_criteria: {runs: , required_pass_rate:
}
run_config: {model: , prompt_version: ,
temperature: }
```

CHOOSE THE ORACLE

Oracle	Use for
Assertion	Citations, forbidden phrases, length, JSON
Schema	Structured output, tool arguments
Tool-call check	Right tool, no write before confirmation
Rubric judge	Groundedness, tone, completeness
Human review	New or high-risk scenarios, judge calibration
Metamorphic	Paraphrase, language and order invariance

RULES OF THUMB

- Specify behavior, not exact wording
- Repeat runs; thresholds by severity
- Stub tools and retrieval; pin model and prompt
- Trace every test to a requirement, risk or incident
- Every incident becomes a permanent test