

ENTRY PHASE

Prompt  Start Position

Input Prompt

- Direct text entry
- File upload (.txt, .md)
- Paste from clipboard

Select Analysis Mode

Which risk type?

Faithfulness

Faithfulness
Context adherence

Factuality

Factuality
World knowledge

Both

Both
Comprehensive

ANALYSIS PHASE

Run Analysis

Analyzer Agent
processes prompt

Processing

- Load 12-pillar guidelines

- Pattern matching
- PRD calculation

View Results

- Highlighted prompt
 - PRD gauges
- Violation cards

INTERPRETATION PHASE

Understanding Risks

Review PRD Scores

prompt_PRD | meta_PRD

20-50%

⚠ Moderate Risk
 $20\% \leq \text{PRD} < 50\%$
Some improvements
needed

> 50%

🔴 High Risk
 $\text{PRD} \geq 50\%$
Significant revision
needed

< 20%

Examine Highlights

● Medium | ● High | ●
Critical

Read Violation Details

Pillar + Rule + Mitigation

Review Initiator Questions

One question per broken rule

✓ Low Risk

PRD < 20%

Prompt well-structured

REFINEMENT PHASE
Iterative Improvement

Refine Prompt?

Yes

Open Chat Interface

Conversational Agent

Discuss Risks

- Ask clarification questions
- Receive suggestions
- Critical (non-yes-man) feedback

Apply Suggestions

Incorporate improvements

Yes



No

RE-ANALYSIS PHASE

Prepare Re-Analysis

Preparator Agent
synthesizes conversation

Preview Refined Prompt

- See changes summary
- Edit if needed

Edit

Confirm Preview?

Yes

Re-Analyze

Fresh analysis on
refined prompt

Compare PRD

Before vs After

No - Accept

No - Iterate

🎯 Measure improvement

PRD Acceptable?

Yes

EXIT PHASE

Export Results

- JSON (with metadata)
 - Markdown
 - PDF report

Deploy Prompt

Use in production
with reduced risk

🏁 End