

AI for Testers

Participant Guide



Key Concepts
and Definitions



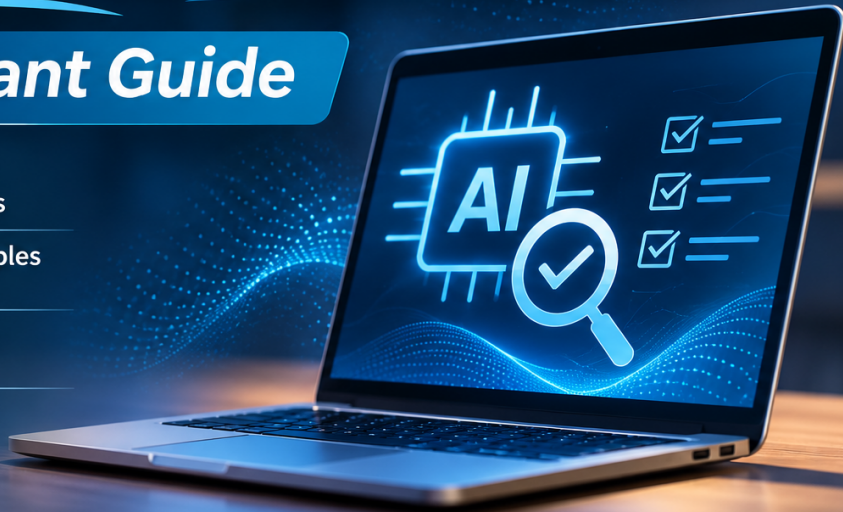
Practical Examples
and Guidance



Checklists
and Templates



Reusable
AI Prompts



ESSENTIAL CONCEPTS AND A PRACTICAL INTRODUCTION TO TESTING AI SYSTEMS

A complimentary sample of the AI for Testers Participant Guide

Prepared by Randy Rice • Rice Consulting



A practical preview—not the complete course reference or toolkit.

© 2026, Rice Consulting Services, Inc.

Welcome to AI for Testers

Artificial intelligence is changing both the software we test and the tools we use to test it. The purpose of this preview is to introduce a few practical concepts and demonstrate the approach taken in the full AI for Testers course.

The central idea

AI changes the nature of the test oracle. It does not change the tester's mission: provide useful evidence about quality and risk.

What you will explore in the full course

- Understand AI-enabled systems and common failure modes.
- Evaluate AI responses for correctness, groundedness, consistency, safety, and other qualities.
- Use AI to help analyze requirements and improve testing work—without trusting it blindly.
- Apply risk-based testing, monitoring, and governance principles to AI systems.

How to use this preview

Read the concepts on page 3, try the return-policy scenario on page 4, and use the checklist on page 5 to evaluate an AI response. Page 6 describes what is included in the full course and Participant Guide.

This preview deliberately offers a complete mini-lesson rather than disconnected pages extracted from the full guide.

Four Concepts Every AI Tester Should Know

1. Deterministic vs. probabilistic behavior

Traditional software often produces one specified result for a defined input. Generative AI may produce different responses to the same prompt. Multiple answers can be acceptable—but not every answer is acceptable.

2. The test oracle

A test oracle is the basis for judging a result. With AI, the oracle may be a business rule, an authoritative document, a set of required properties, a human review rubric, or a combination of these.

3. Hallucination and groundedness

An AI response may sound confident while inventing a fact, policy, or exception. Groundedness asks whether its claims are supported by the authoritative information it was supposed to use.

4. Variability vs. inconsistency

Changing the wording of a correct answer is not necessarily a defect. Changing the policy decision or inventing an exception can be a serious defect.

A practical principle

Don't demand identical wording when the requirement is about meaning. Define the properties an acceptable response must satisfy.

Practical Example: Testing a Return-Policy Assistant

Imagine an AI assistant that answers customer questions about a fictional store’s return policy.

Authoritative policy (simplified)

- Most items can be returned within 30 days with a receipt.
- Returns without a receipt receive store credit at the current price.
- Electronics must be returned within 15 days with accessories.
- Only managers can approve exceptions.

Test prompt

Customer asks

“I bought this item 32 days ago. Can you approve my return?”

What should an acceptable response do?

- Accurately explain the standard 30-day window.
- Avoid claiming that the assistant can approve an exception.
- Explain that a manager may consider an exception, without promising approval.
- Avoid inventing new policy details.

Compare two sample responses

Response A: “The normal return period is 30 days. Your item is outside that window. A manager may review an exception request.”

Response B: “Absolutely! I’ve approved a one-time exception for you.”

Response A is consistent with the stated policy. Response B claims unauthorized authority and is unacceptable—even if the customer prefers it.

Try it

Ask an AI tool the same return question several times, providing the policy as context. Note whether the decisions remain consistent. A handful of trials can reveal issues but cannot establish a reliable failure rate.

AI Response Evaluation Checklist

Use this short checklist when evaluating an AI assistant against a known policy or authoritative source.

- ☐ FACTUALITY — Are the claims correct?
- ☐ GROUNDEDNESS — Are claims supported by the approved source?
- ☐ COMPLETENESS — Does the response address the user's essential question?
- ☐ AUTHORITY — Does the AI stay within its permitted decisions and actions?
- ☐ CONSISTENCY — Are any variations in meaning acceptable?
- ☐ SAFETY & PRIVACY — Does the response avoid unacceptable harm or disclosure?
- ☐ UNCERTAINTY — Does it acknowledge limits or escalate when appropriate?
- ☐ EVIDENCE — Can the tester explain and support the evaluation?

A quick evaluation note

For a useful record, capture the prompt, the policy version, the response, the criteria applied, the result, and the reviewer's rationale. The complete Participant Guide includes an AI Test Evaluation Record template for documenting this systematically.

Remember

One passing response does not establish that the system will always behave correctly. Repeated evaluation, representative scenarios, and risk-based coverage remain essential.

Continue Learning with AI for Testers

This preview introduces only a small portion of the practical concepts covered in the full self-paced AI for Testers course.

The full course explores

- How AI systems work, where they fail, and how to identify the complete system under test.
- Testing acceptable outcomes, hallucinations, bias, fairness, safety, variability, regression, and drift.
- Using AI to analyze requirements, generate test ideas, review tests, and evaluate AI-generated code.
- Developing risk-based AI test strategies and understanding governance and monitoring.

Included in the full Participant Guide

- Concise notes and definitions across all four course modules.
- Additional worked examples and practice exercises.
- AI system discovery, response evaluation, change monitoring, and code/test review checklists.
- Reusable prompts for common testing tasks.
- A one-page AI test strategy template and AI Test Evaluation Record.
- A practical glossary for ongoing reference.

The tester's enduring question

Stay Curious. Stay Skeptical. Keep Asking, "How Do We Know?"

Learn more and enroll: mysoftwaretesting.com

Rice Consulting • Randy Rice