



AI Testing & Validation Strategy

Table of contents

Table of contents.....2

Introduction..... 2

Multi-Dimensional Validation Framework..... 2

 1. Effectiveness..... 2

 2. Reliability..... 3

 3. Instruction Adherence..... 3

 4. Precision..... 3

 5. Efficiency..... 3

Complexity-Aware Testing Approach..... 3

Continuous Monitoring & Observability..... 4

Quality Gates for AI..... 4

Customer Assurance..... 4

Introduction

At Matillion, we are AI-first in everything we do. Our agentic data engineer, Maia, is designed to accelerate data productivity by understanding natural language instructions and delivering high-quality pipeline solutions. With AI embedded throughout our platform, we apply the same world-class quality standards to AI-generated outputs as we do to traditional software development. This document outlines how we ensure our AI services are reliable, precise, and trustworthy.

Multi-Dimensional Validation Framework

We validate AI functionality across five critical dimensions, ensuring comprehensive quality assurance:

1. Effectiveness

Primary Focus: Did the AI achieve the intended goal?

Validation: We assess overall goal achievement with clear success thresholds that adapt based on task complexity.

Assurance: Every AI interaction is measured against predefined success criteria.

2. Reliability

Primary Focus: Can customers trust the AI consistently?

Validation: We monitor success rates across all operations, targeting trusted reliability.

Assurance: AI responses are consistent, repeatable, and dependable across varied scenarios.

3. Instruction Adherence

Primary Focus: Did the AI correctly understand and follow requirements?

Validation: Structured evaluation ensures the AI interprets customer intent accurately.

Assurance: AI outputs align precisely with user instructions and business objectives.

4. Precision

Primary Focus: Is the output technically sound and following best practices?

Validation: Automated validation checks ensure outputs meet technical standards and quality benchmarks.

Assurance: AI-generated solutions adhere to industry best practices and platform guidelines.

5. Efficiency

Primary Focus: Was the goal achieved economically and quickly?

Validation: We monitor resource usage, response times, and optimisation opportunities.

Assurance: AI delivers value without unnecessary consumption of time or resources.

Complexity-Aware Testing Approach

Not all AI tasks are equal. We categorise interactions by complexity and apply appropriate quality thresholds:

- **Explain Tasks:** Read-only operations with 95%+ expected success
- **Simple Tasks:** Single-component solutions with 85%+ expected success
- **Moderate Tasks:** Multi-component workflows with 75%+ expected success
- **Complex Tasks:** Advanced multi-file solutions with 65%+ expected success

This graduated approach ensures realistic expectations while maintaining rigorous quality standards.

Continuous Monitoring & Observability

AI quality is not a one-time check. We implement continuous validation:

- **Real-Time Telemetry:** Every AI interaction is monitored and logged.
- **Automated Evaluation:** Post-deployment testing validates AI behavior in live environments.
- **Trend Analysis:** We track performance over time to identify and address degradation proactively.
- **Customer Feedback Integration:** User satisfaction metrics inform continuous improvement.

Quality Gates for AI

AI services pass through the same rigorous quality gates as all Matillion services:

- **Automated Pre-Deployment Testing:** Comprehensive validation before any release.
- **Multi-Environment Promotion:** Changes progress through development, pre-production, and production environments.
- **Feature Flagging:** Controlled rollout ensures safe adoption.
- **Synthetic Monitoring:** Continuous health checks in production.
- **Zero-Bug Policy:** All identified issues are resolved before customer exposure.

Customer Assurance

Our commitment to quality means:

- **Proactive Testing:** We identify and prevent issues before they reach customers.
- **Transparent Standards:** Clear success criteria and quality benchmarks guide development.
- **Continuous Validation:** AI outputs are validated continuously, not just at release.
- **Shift-Left Culture:** Quality is embedded from the earliest stages of development.
- **Fastidious Verification:** Every AI output undergoes rigorous validation to ensure accuracy and reliability..

For more information about Matillion's quality standards, please contact your account representative.

Date	Author	Description
December 2025	Steve Warburton	First publication <ul style="list-style-type: none">• Initial release of the document.• Formatting and layout set for readability.