

AI-Enabled Quality Engineering

Executive discovery questionnaire - role-routed, multiple-selection version (v4)

Purpose: tailor the executive presentation, select relevant industry research, identify practical QA opportunities, and define a conservative phased pilot. Approximate ranges are sufficient; no customer data, source code, detailed rate cards, or formal financial commitments are required.

How to complete (allow 15-25 minutes; answer the sections you own)

Answer only the sections for your role and leave the others blank:

Executive sponsor, Finance, procurement: Sections 1, 5B and 6.

Engineering, delivery or QE leader: Sections 2, 3, 4, 5A and 6.

QA platform or DevOps leader: Sections 2, 3, 4, 5A and 6.

Square boxes allow multiple selections; respect the limit shown. Dropdown fields take one answer; in the grid question (Q17) tick one box per row. Click fields to enter text, then save the PDF normally.

Respondent information

Your role in this discussion (routes the questionnaire)

Name

Title / role

Function / business area

Date

1. Executive objectives and spending concerns

Executive sponsor and Finance route. Establish what the executive sponsor expects the conversation to solve and how success should be described.

1. What are the primary objectives for exploring AI in quality engineering?

Select up to three.

- | | |
|--|--|
| ☐ Reduce manual testing effort | ☐ Shorten regression and release cycles |
| ☐ Improve automation coverage and maintainability | ☐ Reduce escaped production defects |
| ☐ Improve audit and release evidence | ☐ Increase delivery capacity without increasing team size |
| ☐ Improve developer productivity | |

2. If a pilot proved one thing, which ONE outcome would justify continuing?

Select one. This defines what "success" will mean for the executive sponsor.

3. Which outcomes must NOT deteriorate for a pilot to count as a success?

Select all that apply. These become the quality and delivery floors.

- | | |
|---|---|
| ☐ Escaped defects or production incidents | ☐ Release stability or change-failure rate |
| ☐ Payment correctness and critical defect resolution | ☐ Privacy and sensitive-data handling |
| ☐ Audit and release evidence | ☐ Engineer adoption and morale |

1. Executive objectives and spending concerns

Continued

4. What are the greatest concerns with current or previous AI initiatives?

Select up to three.

- | | |
|---|--|
| ☐ Unclear or unproven business value | ☐ Excessive consulting or implementation cost |
| ☐ High licensing, token, or model-consumption cost | ☐ Too many overlapping tools or platforms |
| ☐ Low user adoption or weak workflow fit | ☐ Security, privacy, or data-residency risk |
| ☐ Difficulty moving beyond pilots | ☐ Weak governance, controls, or accountability |
| ☐ Productivity gains that do not become budget savings | ☐ Quality regressions from AI-generated artifacts |
| ☐ Regulatory or audit scrutiny of AI use | ☐ Previous AI pilots that did not deliver |

2. Delivery scope and QA operating model

Engineering, delivery and QE route. Approximate ranges are sufficient; detailed rate cards and organization charts are not required.

5. Approximately how many delivery teams or applications are in the potential scope?

Select one range.

Target business area, portfolio, or application type (optional):

6. Roughly what share of in-scope applications are modern cloud or API-based systems with automated build and deployment pipelines?

Select one range.

7. For a typical delivery squad, provide the approximate number of resources by role.

Enter approximate numbers for a typical or median squad; ranges are acceptable (for example 6 / 2 / 1 / 1 / shared).

Role / capability	Approx. number
Developers	
Manual QA analysts	
Automation engineers / SDETs / QEs	
Business analysts and product owners	
Test-data, environment and platform engineers	

8. Approximately what percentage of QA headcount is external, offshore, or delivered through a managed service?

Select one range (by headcount, not spend).

2. Delivery scope and QA operating model

Continued

9. Which descriptions apply to the current quality-engineering operating model?

Select all that apply.

- | | |
|--|---|
| ☐ Centralized QA organization | ☐ QA / QE embedded in delivery squads |
| ☐ Managed-service model | ☐ Hybrid centralized and embedded model |
| ☐ Dedicated automation / SDET capability | ☐ Business-led UAT outside the QA organization |
| ☐ Model varies significantly by business unit | ☐ Operating model is being redesigned |

10. Where does the QA organization spend the most human effort?

Select your top THREE.

- | | |
|---|---|
| ☐ Requirements review and acceptance-criteria refinement | ☐ Test planning, scenario design, and test-case creation |
| ☐ Building automated tests | ☐ Maintaining or repairing automated tests |
| ☐ Preparing test data | ☐ Preparing or troubleshooting test environments |
| ☐ Manual functional test execution | ☐ Regression execution |
| ☐ Failed-test triage and root-cause investigation | ☐ Defect creation, routing, retesting, and closure |
| ☐ UAT coordination and support | ☐ Quality reporting, traceability, and release evidence |

3. Testing maturity and workflow

Engineering, delivery and QE route. Identify where human effort, wait time, rework, and quality risk are concentrated today.

11. Roughly what share of regression test cases execute automatically without human intervention?

Select one range.

12. Which test types are consistently automated within CI/CD?

Select all that apply.

- | | | |
|---------------------------------------|--------------------------------------|--|
| ☐ Unit | ☐ Component | ☐ API |
| ☐ Contract | ☐ Integration | ☐ UI / end-to-end |
| ☐ Regression | ☐ Performance | ☐ Accessibility |
| ☐ Data quality | ☐ Few or none | ☐ Varies significantly by application |

13. What is the typical elapsed regression-testing duration for a major release?

Select one range, measured in calendar days.

14. Approximately how many active human hours does a full regression cycle require?

Select one range. Count hands-on time across all people, not elapsed time.

3. Testing maturity and workflow

Continued

15. Which issues most frequently delay testing or releases?

Select up to three.

- | | |
|---|--|
| ☐ Incomplete or changing requirements | ☐ Insufficient unit or component testing |
| ☐ Manual test-case creation | ☐ Manual regression execution |
| ☐ Unstable automation or flaky tests | ☐ Test-data availability |
| ☐ Test-environment availability or instability | ☐ Downstream-system dependencies |
| ☐ Legacy or mainframe integration | ☐ Defect triage and ownership |
| ☐ Long waits for defect fixes and retesting | ☐ UAT, release evidence, or approval requirements |

16. Which quality checks are mandatory release gates?

Select all that apply.

- | | |
|---|--|
| ☐ Unit-test pass rate and/or code coverage | ☐ API, integration, or regression pass rate |
| ☐ Critical and high-priority defect thresholds | ☐ Performance or resilience thresholds |
| ☐ Accessibility compliance | ☐ UAT or business approval |
| ☐ Production-readiness / change approval | ☐ Traceability, audit, or control evidence |

4. Metrics, toolchain, and AI readiness

Engineering, QE and platform routes. These responses determine whether a leadership savings hypothesis can be evidence-based.

17. For each baseline measure, indicate whether it is tracked and trusted, tracked but unreliable, or not tracked.

Select one per row. Leave a row blank if you do not know.

Measure	Tracked and trusted	Tracked but unreliable	Not tracked
QA effort hours by activity, release, or application	☐	☐	☐
Regression duration and execution volume	☐	☐	☐
Automation coverage and maintenance effort	☐	☐	☐
Flaky-test or rerun rate	☐	☐	☐
Defect volume, severity, reopen rate, and escaped defects	☐	☐	☐
Change-failure, rollback, or production-incident rate	☐	☐	☐
Mean time to triage and remediate defects	☐	☐	☐
Release frequency and lead time for change	☐	☐	☐
QA labour, contractor, managed-service, or tool spend	☐	☐	☐
Test-environment availability and test-data lead time	☐	☐	☐

18. Which enterprise AI platform is approved for engineering pilots, and which coding or testing assistants are approved for engineers today?

Select one platform; list assistants if known.

Approved enterprise AI platform

Approved coding or testing assistants (for example GitHub Copilot, test-tool AI features):

4. Metrics, toolchain, and AI readiness

Continued

19. Identify the primary tools currently used.

Optional. Product names are helpful but not required.

Capability	Current platform or tool
Requirements / work management	
Source control and CI/CD	
Test management and automation	
Observability / logs / traces	
Test data, environments and device/browser coverage	

20. Which AI-assisted QA capabilities are already in use or being evaluated?

Tick "In use" and/or "Evaluating" for each row that applies.

	In use	Evaluating
Requirements, acceptance criteria, or test-scenario generation	☐	☐
Unit, API, or UI test generation	☐	☐
Test automation maintenance or self-healing	☐	☐
Regression-test selection or prioritization	☐	☐
Synthetic test-data generation	☐	☐
Failed-test triage or root-cause analysis	☐	☐
Defect creation, classification, or routing	☐	☐
Release-quality summaries or evidence preparation	☐	☐
None / only informal experimentation	☐	☐

5A. Shared QA platform and AI controls

Engineering and platform routes. Identify reusable QA services, ownership and the action limits of a first pilot.

21. Which shared QA platform gaps most limit reuse across teams?

Select up to three.

- | | |
|---|---|
| ☐ Requirements and test-context connectors | ☐ Reusable automation frameworks and templates |
| ☐ Synthetic test data and fixture services | ☐ On-demand test environments and dependency stubs |
| ☐ Web, mobile and API runner orchestration | ☐ Shared test results, evidence and traceability |
| ☐ Failure-triage and test-maintenance workflow | ☐ Platform ownership, support and adoption |

22. What is the HIGHEST level of AI action acceptable during an initial pilot?

Select one. Each level includes the levels above it.

5A. Shared QA platform and AI controls

Continued

23. Which AI controls are NOT yet in place for AI tooling and would need to be addressed before a pilot?

Select all that apply. Leave blank if all are in place.

- | | |
|--|--|
| ☐ Source-code and intellectual-property confidentiality | ☐ Customer-data privacy and sensitive-data handling |
| ☐ Data residency and retention | ☐ Approved-model and vendor enforcement |
| ☐ Role-based access and segregation of duties for AI identities | ☐ Audit logging and traceability of prompts, outputs, and approvals |
| ☐ Human approval for production-impacting actions | ☐ Explainability and supporting evidence |
| ☐ Model, prompt, and generated-artifact versioning | ☐ License, token, and model-consumption cost controls |
| ☐ Vendor portability and ability to disable AI safely | ☐ Model inventory and risk rating (OSFI E-23 readiness) |

5B. Economics and financial capture

Executive sponsor and Finance route. A productivity gain becomes a saving only when Finance can name the budget line that changes.

24. Which forms of spending or capacity could realistically be reduced, avoided, or redeployed within the next 12 months?

Select all that apply.

- | | |
|--|--|
| ☐ Contractor renewals falling due in the next 12 months | ☐ Managed-service scope or unit volumes |
| ☐ Planned QA hiring | ☐ Overtime or surge capacity |
| ☐ Testing-tool licenses | ☐ Test infrastructure or execution cost |
| ☐ Defect rework and production-incident effort | ☐ Capacity redeployed to additional delivery work |
| ☐ No capture mechanism has yet been identified | ☐ Not my decision |

25. Approximately what is the annual addressable QA and testing spend in scope (internal labour plus external services and tooling)?

Select one range. Prefer not to say is acceptable.

26. Roughly what share of that spend is variable (contractors, offshore, managed services) rather than employees?

Select one range.

27. If QA capacity were released, what would most likely happen to it?

Select one.

28. What is the minimum evidence Finance would accept to recognize a saving?

Select one.

6. Pilot direction and executive input

All routes. A conservative pilot should be limited, measurable, and subject to an explicit go/no-go decision.

29. Which application types could be suitable for an initial pilot?

Select all that apply.

- | | |
|---|--|
| ☐ Modern API or microservices application | ☐ Internal web application |
| ☐ Customer-facing web application | ☐ Mobile application |
| ☐ Legacy or mainframe-integrated application | ☐ Data or analytics application |
| ☐ Packaged platform | ☐ More than one application type for comparison |
| ☐ Candidate not yet selected | |

30. Which use cases could be suitable for a limited pilot?

Select up to three.

- | | |
|--|---|
| ☐ Requirement analysis and test-scenario generation | ☐ Executable API or component-test generation |
| ☐ UI automation generation or maintenance | ☐ Change-impact analysis and regression-test selection |
| ☐ Failed-test triage and root-cause analysis | ☐ Flaky-test detection and automation repair |
| ☐ Synthetic test-data generation | ☐ Release evidence and readiness summaries |

31. What minimum net reduction in targeted human effort would justify a second phase?

Select one. "Net" deducts review, correction, and control effort.

32. Beyond the effort threshold, which outcomes would justify further investment?

Select up to two.

- | | |
|--|---|
| ☐ Material reduction in regression duration | ☐ Material reduction in failure-triage time |
| ☐ Material reduction in automation-maintenance effort | ☐ No deterioration in escaped defects or change-failure rate |
| ☐ Shorter test-data and environment wait times | ☐ Demonstrated contractor reduction or hiring avoidance |
| ☐ Credible payback within 12-18 months | |

33. What conditions should stop or prevent expansion of the pilot?

Select all that apply.

- | | |
|---|---|
| ☐ Insufficient measurable savings | ☐ High implementation or integration cost |
| ☐ High ongoing license or model-consumption cost | ☐ Poor accuracy or excessive human rework |
| ☐ Security, privacy, residency, or audit concerns | ☐ Inability to integrate with the existing toolchain |
| ☐ Quality or release-stability deterioration | ☐ Low adoption or workflow abandonment |
| ☐ No credible mechanism to capture the value financially | |

34. What is the single most expensive or frustrating part of the current QA workflow?

Optional. Up to 1,000 characters; scroll within the field to read longer answers.

6. Pilot direction and executive input

Continued

35. What question would the executive sponsor most want the presentation to answer?

Optional. Up to 1,000 characters; scroll within the field to read longer answers.

36. Is there a representative application, team, or recent release that should be used as an illustrative case, and who is the delivery lead we may contact?

Optional. Up to 1,000 characters; scroll within the field to read longer answers.

Thank you. Responses will be used to prepare a focused executive presentation, an initial savings hypothesis with explicit assumptions, and a conservative phased pilot with a defined cost ceiling and go/no-go decision.