Sample Exam – Questions

Sample Exam set A
Version 1.1

ISTQB® Testing with Generative AI Syllabus Specialist Level

Compatible with Syllabus version 1.1

International Software Testing Qualifications Board

Version 1.1 27/04/2026

Copyright Notice

Document Responsibility

The ISTQB® Examination Working Group is responsible for this document.

This document is maintained by a core team from ISTQB® consisting of the Syllabus Working Group and Exam Working Group.

Acknowledgements

This document was produced by a core team from ISTQB®: Abbas Ahmad, Alessandro Collino, Bruno Legeard.

The core team thanks the Exam Working Group review team, the Syllabus Working Group and Member Boards for their suggestions and input.

Revision History

Sample Exam – Questions Layout Template used:	Version 2.12	Date: November 27, 2024

Version	Date	Remarks
V1.0	2024/12/30	CT-GenAI Sample Exam set A 1.0 Alpha Release
V1.0	2025/04/17	CT-GenAI Sample Exam set A 1.0 Beta Release
V1.0	2025/06/10	CT-GenAI Sample Exam set A 1.0 Release
V1.1	2026/04/27	CT-GenAI Sample Exam set A 1.1 Minor Update

Table of Contents

Introduction

Purpose of this document

The example questions and answers and associated justifications in this sample exam have been created by a team of subject matter experts and experienced question writers with the aim of:

- Assisting ISTQB® Member Boards and Exam Boards in their question writing activities
- Providing training providers and exam candidates with examples of exam questions

These questions cannot be used as-is in any official examination.

Note, that real exams may include a wide variety of questions, and this sample exam ***is not*** intended to include examples of all possible question types, styles or lengths, also this sample exam may both be more difficult or less difficult than any official exam.

Instructions

In this document you may find:

- Questions[1], including for each question:
  - Any scenario needed by the question stem
  - Point value
  - Response (answer) option set
- Additional questions, including for each question [does not apply to all sample exams]:
  - Any scenario needed by the question stem
  - Point value
  - Response (answer) option set

- *Answers, including justification are contained in a separate document*

[1] In this sample exam the questions are sorted by the LO they target; this cannot be expected of a live exam.

Questions

Question #1 (1 Point)

Match each type of AI technology (1-4) with its CORRECT description (A-D):

1. Symbolic AI
2. Classical machine learning
3. Deep learning
4. Generative AI

A. Uses neural networks to automatically learn features from data.
B. Uses rule-based systems to mimic human decision-making.
C. Uses deep learning to create new data by learning from its training data.
D. Uses a data-driven approach that requires feature selection.

a) 1D, 2B, 3A, 4C
b) 1D, 2C, 3B, 4A
c) 1C, 2B, 3D, 4A
d) 1B, 2D, 3A, 4C

Select ONE option.

Question #2 (1 Point)

Consider the realm of Large Language Models (LLMs). Which of the following options BEST explains why context window limitations affect LLM's text processing capabilities?

a) Because context windows restrict temporal processing sequences, preventing LLMs from maintaining chronological consistency across extended text analysis.
b) Because context windows prevent cross-referencing capabilities, limiting LLMs' ability to connect information across different document sources simultaneously.
c) Because context windows force LLMs to discard earlier information, which may contain relevant details needed for understanding later content.
d) Because context windows constrain parsing granularity levels, restricting LLMs from adjusting between character-level and document-level analysis approaches.

Select ONE option.

Question #3 (1 Point)

Which of the following statements BEST describes tokenization in processing text for LLMs?

a) Tokenization converts tokens into high-dimensional vectors to capture their meaning.
b) Tokenization creates the building blocks used to understand and generate text.
c) Tokenization generates contextually appropriate responses using neural networks.
d) Tokenization predicts the next token in a sequence based on learned relationships.

Select ONE option.

Question #4 (1 Point)

In the context of software testing, which of the following statements (i-v) about foundation, instruction-tuned, and reasoning LLMs are CORRECT?

i. Foundation LLMs excel at generating test cases from high-level requirements without structured input.
ii. Reasoning LLMs excel at creating test scripts that strictly follow predefined organizational templates.
iii. Instruction-tuned LLMs excel at autonomously prioritizing test execution based on real-time user feedback.
iv. Reasoning LLMs excel at synthesizing data from defect reports to detect trends and prioritize test efforts.
v. Instruction-tuned LLMs excel at generating test cases that adhere to Gherkin language syntax.

a) i, ii, and iii
b) ii, iii, and iv
c) i, ii, and v
d) iv, and v

Select ONE option.

Question #5 (1 Point)

Which of the following statements BEST describes the relation between multimodal LLMs and vision-language models?

a) Multimodal LLMs are a subset of vision-language models designed to handle diverse inputs.
b) Vision-language models are a subset of multimodal LLMs focusing on visual and textual data.
c) Vision-language models are unrelated to multimodal LLMs and focus only on the user interface.
d) Multimodal LLMs and vision-language models are interchangeable terms.

Select ONE option.

Question #6 (1 Point)

Which TWO of the following options represent key capabilities of LLMs in test tasks?

a) Identifying ambiguities and inconsistencies in requirements.
b) Generating complete application code for deployment.
c) Automating the execution of all test scripts without human intervention.
d) Performing exploratory testing on software applications.
e) Creating diverse test data with various combinations and boundary values.

Select TWO options.

Question #7 (1 Point)

Which of the following statements BEST explains the difference between AI chatbots and LLM-powered testing applications in the context of software testing?

a) AI chatbots are more suited for specific test tasks, while LLM-powered testing applications focus on ad hoc interactions.
b) Both AI chatbots and LLM-powered testing applications are designed to perform identical tasks without any configuration differences.
c) LLM-powered testing applications rely on conversational prompts, while AI chatbots require integration into test tools and test processes.
d) AI chatbots offer conversational interfaces for ad hoc test tasks, while LLM-powered testing applications provide customized solutions for specific test tasks.

Select ONE option.

Question #8 (1 Point)

A tester is examining a structured prompt used to obtain LLM assistance for performance test analysis. One of the components of this prompt reads: “Test reports from performance testing tools, system monitoring logs during peak usage periods, and application performance benchmarks from previous releases”.

In which component of the six-part prompt structure would this description MOST LIKELY appear?

a) Context
b) Input data
c) Constraints
d) Output format

Select ONE option.

Question #9 (1 Point)

A tester wants an LLM to analyze a requirements specification for potential defects. In the structured prompt the tester is using, one line reads: “The potential defects must be provided in a markdown table with the following columns: ID, requirement reference, defect type, description, severity”.

In which component of the six-part prompt structure would this line MOST LIKELY appear?

a) Instructions
b) Constraints
c) Output format
d) Context

Select ONE option.

Question #10 (1 Point)

Which of the following BEST differentiates prompt chaining, few-shot prompting, and meta prompting techniques?

a) Prompt chaining focuses on providing examples, few-shot prompting breaks tasks into subtasks, and meta prompting refines prompts manually.
b) Few-shot prompting provides guidance with examples, prompt chaining breaks tasks into multiple prompts, and meta prompting allows the model to iteratively refine its own prompts.
c) Meta prompting emphasizes breaking down tasks into steps, prompt chaining uses examples, and few-shot prompting focuses on manual optimization of prompts.
d) Prompt chaining provides guidance without examples, few-shot provides guidance with examples, and meta prompting relies on tester-defined prompts.

Select ONE option.

Question #11 (1 Point)

What is the primary function of a system prompt in interactions with LLMs?

a) To provide a framework for the LLM behavior for the entire conversation.
b) To provide specific questions or instructions from the user to the LLM.
c) To adjust dynamically with each user interaction and set the conversation's context.
d) To include visible input from the user and set rules for the conversation.

Select ONE option.

Question #12 (2 Points)

You are tasked with applying the following test approach to a set of stable requirements for a new project: generate test conditions, prioritize them based on risk level, and identify potential coverage gaps. The requirements have already been thoroughly reviewed for defects.

Which of the following sequences of steps (i-v) should you follow to effectively apply Generative AI to implement this test approach using a prompt chaining technique to implement this test approach?

i. Submit the requirements to the LLM and prompt it to produce test conditions based on those requirements.
ii. Provide the test conditions to the LLM, ensuring it understands the context for prioritization, and prompt it to prioritize those test conditions accordingly.
iii. Provide the prioritized test conditions to the LLM and prompt it to analyze them to determine whether all aspects of the requirements are addressed in the test conditions.
iv. Submit the requirements to the LLM and prompt it to produce prioritized test conditions that address all aspects of the requirements.
v. Submit the requirements to the LLM and prompt it to detect inconsistencies and ambiguities in those requirements.

a) i, ii, and iii
b) iv, and ii
c) i, iii, and v
d) v, and iv

Select ONE option.

Question #13 (2 Points)

Consider applying the few-shot structured prompting technique to generate Gherkin-style test cases (i.e., scenario-based) for the following user story and acceptance criterion:

- User story: “As a user, I want to reset my password so that I can regain access to my account if I forget it.”
- Acceptance criterion: “When a user submits a registered email address then they receive a password reset email.”

You can rely on predefined examples that include user stories, acceptance criteria, and Gherkin-style test cases. Your task is to create a prompt to guide the LLM in generating accurate test cases aligned with the acceptance criterion for the user story above.

Which one of the following prompts is BEST suited to this task?

a) Prompt A

Role: Act as a test analyst.
Context: You are testing a password reset functionality.
Instruction: Generate Gherkin-style test cases for the user story and acceptance criterion. using the following predefined examples as a guide: << predefined examples >>.
Input Data: <<< user story >>> and <<< acceptance criterion >>>.
Constraints: Rely on best practices to create test cases.
Output Format: Generate test cases with expected results.

b) Prompt B

Role: Act as a test analyst specializing in Gherkin-style test cases.
Context: You are testing a password reset functionality.
Instruction: Generate Gherkin-style test cases for the user story and acceptance criterion, using the following predefined examples as a guide: << predefined examples >>.
Input Data: <<< user story >>> and <<< acceptance criterion >>>.
Constraints: Use "Given-When-Then" syntax and ensure alignment with the acceptance criterion.
Output Format: Respect the given Gherkin-style test case format.

c) Prompt C

Role: Act as a test analyst.
Context: You are testing a password reset functionality.
Instruction: Generate Gherkin-style test cases for the user story and acceptance criterion. Rely on best practices to create test cases.
Input Data: <<< user story >>> and <<< acceptance criterion >>>.
Constraints: Use "Given-When-Then" syntax and ensure alignment with the acceptance criterion.
Output Format: Respect the given Gherkin-style test case format.

d) Prompt D

Role: Act as a test analyst.
Context: You are testing a password reset functionality.
Instruction: Generate at least two Gherkin-style test cases for the user story and acceptance criterion. Focus on edge cases.
Input Data: <<< user story >>> and <<< acceptance criterion >>>.
Constraints: Ensure all test cases follow "Given-When-Then" syntax.
Output Format: Respect the given Gherkin-style test case format.

Select ONE option.

Question #14 (2 Points)

You are tasked with applying structured prompting to analyze regression test results. Here is an initial draft of the prompt:

> Role: Act as a test analyst.
> Context: Analyze raw regression test results from a recent test execution cycle.
> Instruction: Identify discrepancies in the test results.
> Input Data: Use the attached file containing raw test results.
> Constraints: Use the known anomalies list for cross-checking.
> Output Format: Provide a list of discrepancies using a table format.

You are asked to improve this prompt. Which of the following improvements would BEST align the prompt with structured prompt engineering best practices for comprehensive regression test report analysis?

a) Add a step to cluster similar issues and cross-check findings against the known anomalies list.
b) Specify that the role is a regression test analyst specializing in actionable insights.
c) Expand the instruction to include separating expected results and actual results, clustering issues, and highlighting discrepancies.
d) Include references to regression testing principles such as "Given-When-Then" in the constraints.

Select ONE option.

Question #15 (2 Points)

You are using an LLM to assist in preparing actionable test metrics from raw data. The metrics include test progress, defect trends, and coverage, which are graphically displayed and explained with text. Your goal is to improve the test process to ensure the generated metrics are accurate, actionable, and easily interpretable by stakeholders.

Here is an initial draft of a prompt used to instruct the AI:

> Role: Act as a test manager.
> Context: You are provided with raw data from test tools.
> Instruction: Generate test progress metrics, defect trend metrics, and coverage metrics from the raw data.
> Input Data: Use the attached file containing raw test results.
> Constraints: Ensure that the output is concise and understandable.
> Output Format: Display metrics on a dashboard.

You are asked to improve this prompt. Which of the following improvements would BEST enhance the LLM's ability to produce accurate and actionable metrics?

a) Specify that the role is a test manager focusing on actionable insights and decision support, ensuring comprehensive analysis of test data.
b) Add an instruction to include potential risks identified from the trends in the generated metrics, along with their impact assessment and priority levels.
c) Expand the output format to include a plain-language summary that interprets the metrics and outlines next steps for stakeholders.
d) Emphasize constraints that the output is also easily interpretable by stakeholders, using clear language and avoiding technical jargon throughout the response.

Select ONE option.

Question #16 (2 Points)

Your goal is to create test cases for an AI-based system that suffers from the test oracle problem, preventing you from determining expected results. You can only count on a few existing test cases with known expected results. Through appropriate analysis, you have identified a set of well-defined transformation rules that specify how changes to inputs affect expected results. These rules can be applied to all existing test cases. You have decided to rely on Generative AI, providing a given LLM with the following information: the existing test cases with their inputs and expected results, a clear description of the transformation rules, and guidelines for generating additional test cases by precisely applying these rules to the relevant existing test cases. With the specified information, the chosen LLM can directly generate additional test cases in line with your expectations.

Which of the following prompting techniques is BEST suited to achieve your goal in this scenario?

a) Few-shot prompting
b) Prompt chaining
c) Meta prompting
d) Zero-shot prompting

Select ONE option.

Question #17 (1 Point)

You are leveraging Generative AI to assist in testing an entertainment software application. The Generative AI model generates test cases for user interaction scenarios, test scripts for API interactions, and synthetic test data to address edge cases.

To effectively evaluate the Generative AI model's performance and to refine prompts, which combination of metrics and actions BEST ensures comprehensive assessment and improvement?

a) Evaluate the diversity of test cases to ensure varied input scenarios and use test execution success rate to validate the functionality of generated API test scripts.
b) Apply accuracy and completeness metrics to validate test cases against entertainment software requirements and rely on time efficiency to compare AI-generated test scripts with manual test efforts.
c) Focus on precision to ensure generated test data meets entertainment software compliance standards, while contextual fit and test execution success rate assesses the alignment and usability of test scripts.
d) Prioritize relevance and contextual fit for all outputs to maintain consistency with entertainment software requirements and include diversity metrics to expand edge case coverage.

Select ONE option.

Question #18 (1 Point)

Which of the following techniques for evaluating and iteratively refining prompts is BEST suited for determining why an LLM consistently generates test cases with wrong expected results that contradict the input requirements, thereby providing insights to optimize the prompt and prevent similar errors?

a) Output analysis
b) A/B testing of prompts
c) Adjusting prompt length and specificity
d) Integrating user feedback

Select ONE option.

Question #19 (1 Point)
What is a hallucination in the context of LLM outputs?

a) A logical error where the LLM fails to follow a multi-step reasoning process accurately.
b) A bias in the LLM output caused by the training data favoring certain perspectives.
c) A generation of irrelevant or factually incorrect output by the LLM for a given task.
d) A limitation of the LLM to understand non-English perspectives in test generation tasks.

Select ONE option.

Question #20 (2 Points)
You are using Generative AI to create test cases for an e-commerce (e-shop) application. The following features have been explicitly mentioned in the project briefing:

- cart management
- discount code application
- order confirmation email generation

Based on these details, which of the following AI-generated test cases MOST LIKELY represents a hallucination?

a) Verify that a user can add multiple items to their cart and proceed to checkout.
b) Verify that a user cannot apply an expired discount code during checkout.
c) Verify that a user receives a confirmation email after successfully placing an order.
d) Verify that a user can create a wishlist to save favorite items for later.

Select ONE option.

Question #21 (1 Point)
Which of the following options refers to a benefit that is MOST directly associated with using clear and structured input data formats when working with LLMs for test tasks?

a) Helps reduce the effort to fine-tune the LLMs for test tasks.
b) Helps LLMs generate less ambiguous outputs for test tasks.
c) Helps LLMs generate more context-relevant outputs for test tasks.
d) Helps LLMs generate more creative outputs for test tasks.

Select ONE option.

Question #22 (1 Point)
Which strategy can help reduce variability in LLM outputs by narrowing the probability distribution during inference?

a) Increasing the learning rate.
b) Lowering the temperature setting.
c) Increasing the random seed.
d) Lowering the random seed.

Select ONE option.

Question #23 (1 Point)

Which of the following statements about data privacy concerns related to using Generative AI for software testing is INCORRECT?

a) Generative AI can unintentionally expose sensitive data through its outputs.
b) Generative AI tools may store and process sensitive data without explicit user consent, leading to misuse.
c) Using Generative AI tools without adhering to data protection regulations, such as General Data Protection Regulation (GDPR), can lead to legal disputes.
d) An LLM is likely to expose real sensitive data if it hallucinates while generating synthetic test data, regardless of the data it was trained on.

Select ONE option.

Question #24 (1 Point)

An attacker injects falsified test results into the training dataset of an LLM intended to recommend optimal test coverage strategies. What type of attack vector does this description BEST refer to?

a) Malicious code generation
b) Context Manipulation
c) Request manipulation
d) Data poisoning

Select ONE option.

Question #25 (1 Point)

Match each type of attack vector against an LLM (1-4) with the corresponding example (A-D):

1. Context Manipulation
2. Request manipulation
3. Data poisoning
4. Malicious code generation

A. An attacker maliciously modifies the data associated with traceability links between requirements and test cases into the dataset used for fine-tuning an LLM, compromising its accuracy in generating test cases from requirements.
B. An attacker maliciously crafts and provides deceptive prompts that induce an LLM, fine-tuned to assist testers in automated test script generation, to produce vulnerable test scripts with hidden security flaws.
C. An attacker maliciously provides large specially crafted prompts that induce an LLM, fine-tuned to assist testers in generating test cases, to accidentally reveal confidential API keys inherited from past test projects.
D. An attacker maliciously submits carefully modified reference screenshots into a visual testing framework that uses an LLM for comparative visual analysis, to trick the LLM into systematically ignoring genuine UI issues during regression testing.

a) 1C, 2D, 3A, 4B
b) 1B, 2D, 3A, 4C
c) 1D, 2C, 3B, 4A
d) 1C, 2B, 3D, 4A

Select ONE option.

Question #26 (1 Point)

Which of the following strategies BEST addresses data privacy risks in the context of Generative AI-powered software testing?

a) Using multiple LLMs to evaluate and compare test results for improved accuracy.
b) Replacing sensitive test data with an anonymized version of the same.
c) Allowing unrestricted access to sensitive test data to improve Generative AI model training.
d) Disabling encryption of sensitive test data to streamline data storage and transmission processes.

Select ONE option.

Question #27 (1 Point)

Which of the following options about the impact of LLM usage on energy consumption and CO_2 emission is CORRECT?

a) Image generation tasks consume substantially more energy than text generation tasks, but produce fewer CO_2 emissions.
b) Generative AI-powered searches consume significantly less energy than traditional web searches due to their optimized algorithms.
c) Image generation tasks consume substantially more energy than text generation tasks due to their higher computational complexity.
d) Text generation tasks consume very little energy, allowing them to be performed by millions of users without significant energy consumption.

Select ONE option.

Question #28 (1 Point)

Which TWO of the following standards, or parts of them, are MOST relevant to the use of Generative AI in software testing?

a) ISO/IEC 25010:2023
b) ISO/IEC 23053:2022
c) ISO/IEC/IEEE 29119-2:2021
d) ISO/IEC 42001:2023
e) ISO/IEC/IEEE 29119-3:2021

Select TWO options.

Question #29 (1 Point)

Which of the following components of an LLM-powered testing application is responsible for combining user input with structured and semantically similar data to prepare a prompt for the LLM?

a) Back-end
b) Front-end
c) Authentication component
d) Post-processing component

Select ONE option.

Question #30 (1 Point)

You are a tester working on a banking application that includes features such as user login, account management, and secure transactions. The system documentation, including API specifications and security requirements, is stored in a vector database, while historical test cases are stored in a relational database. Your task is to generate test cases using a Retrieval-Augmented Generation (RAG) framework to ensure alignment with the latest specifications and requirements.

Which of the following options represents the MOST appropriate use of the RAG framework in this scenario?

a) Submit a query specifying one function to be tested. The RAG framework will retrieve relevant specifications and requirements from the vector database, combine them with historical test cases, and automatically generate accurate and context-aware test cases through the LLM.
b) Submit a query specifying all functions to be tested. The RAG framework will retrieve relevant specifications and requirements from the vector database, combine them with historical test cases, and automatically generate accurate and context-aware test cases through the LLM.
c) Use the RAG framework to retrieve historical test cases from the relational database and security requirements from the vector database. Manually review the retrieved information before refining the query for the LLM to generate targeted test cases.
d) Rely on the LLM's internal training data to generate test cases while using the RAG framework or reference retrieval without directly integrating retrieved information into the generation process.

Select ONE option.

Question #31 (1 Point)

Which of the following options BEST describes the enhancements that autonomous and semi-autonomous LLM-powered agents bring to automating test processes?

a) They can enhance both efficiency and quality in automating test processes through a balanced use of single-agent and multi-agent systems.
b) They can enhance quality in automating test processes by adding complex verification checks although at the expense of efficiency.
c) They can enhance both efficiency and quality in automating test processes by leveraging their ability to operate with varying levels of human interaction.
d) They can enhance both efficiency and quality in automating test processes while eliminating the need for verification within these test processes.

Select ONE option.

Question #32 (1 Point)

Which of the following statements about fine-tuning language models for specific test tasks is INCORRECT?

a) Fine-tuning involves training a pre-trained model on task-specific data to enhance its performance and domain knowledge.
b) Fine-tuning equips a language model with new capabilities by replacing its general knowledge with task-specific reasoning, ensuring the absence of overfitting.
c) Fine-tuning modifies a pre-trained model's parameters using a targeted dataset to adapt it for a specific domain or task.
d) Fine-tuning requires high-quality, task-specific datasets to avoid biased or inaccurate results.

Select ONE option.

Question #33 (1 Point)

Which of the following BEST describes the primary focus of Large Language Model Operations (LLMOps) when deploying and managing LLMs for test tasks?

a) Preventing reliance on Generative AI in test processes.
b) Managing LLMs effectively across their lifecycle, including privacy, security, and cost considerations.
c) Limiting LLM usage to chatbot-based testing solutions to reduce complexity.
d) Fully automating all test tasks without requiring human oversight.

Select ONE option.

Question #34 (1 Point)

Which of the following statements about shadow AI is CORRECT?

a) Shadow AI enforces compliance with organizational data policies and general AI regulations.
b) Shadow AI eliminates the need for clear licensing agreements in AI tools.
c) Shadow AI reduces the risk of intellectual property disputes.
d) Shadow AI may lead to unauthorized accesses to sensitive information.

Select ONE option.

Question #35 (1 Point)

What is a key aspect to consider when defining a Generative AI strategy for software testing?

a) Prepare training programs designed to ensure that team members obtain certifications specific to each LLM they use.
b) Select LLMs that can be integrated appropriately with existing test environments and test tools.
c) Ensure that as much input data as possible is available to increase the likelihood of obtaining effective LLM outputs.
d) Collect standard supervised machine learning metrics to evaluate the effectiveness of LLM outputs.

Select ONE option.

Question #36 (1 Point)

Which of the following statements BEST describes one of the key criteria to select an appropriate LLM for specific test tasks within a test organization?

a) A key selection criterion involves evaluating the LLM's performance for the test tasks against the publicly available benchmarks for LLMs in code generation.
b) A key selection criterion involves evaluating recurring costs such as those associated with the computational resources required to run the LLM.
c) A key selection criterion involves evaluating the LLM against the publicly available community benchmarks to ensure full compatibility with them.
d) A key selection criterion involves evaluating recurring costs such as those associated with a proof of concept aimed at demonstrating the suitability of an LLM for the test tasks.

Select ONE option.

Question #37 (1 Point)

What are the key phases in the adoption of Generative AI in a test organization?

a) Discovery, initiation and usage definition, utilization and iteration
b) Awareness, usage prioritization, performance monitoring
c) Planning, experimentation, evaluation and refinement
d) Training, testing, implementation, and scaling

Select ONE option.

Question #38 (1 Point)

Which of the following options BEST refers to an example of knowledge and/or skills required for testers to work effectively with LLMs in test processes?

a) Mastering techniques specifically aimed at preventing LLMs from hallucinating and incurring reasoning errors when performing specific test tasks.
b) Selecting and implementing suitable test automation approaches, such as keyword-driven test automation, for automating test processes.
c) Choosing the most appropriate LLM based on criteria such as its capability to be adapted or customized to perform specific test tasks.
d) Ensuring the validation and test data used in the development of the LLMS are of the highest quality.

Select ONE option.

Question #39 (1 Point)

What is the BEST approach for cultivating skills within test teams to specifically support the adoption of Generative AI?

a) Rely mainly on external expert courses with hands-on practice, aiming to integrate AI into all daily test tasks at once.
b) Encourage independent experimentation with various LLMs without following a structured process.
c) Adopt a hands-on, gradual learning process supported by guided exercises, peer learning, and knowledge-sharing communities.
d) Rely mainly on theoretical courses from external experts, aiming to gradually integrate AI into daily test tasks in line with actual learning.

Select ONE option.

Question #40 (1 Point)

Which of the following answers BEST describes how the roles and responsibilities of testers and test managers within a test organization are impacted by the adoption of Generative AI for software testing?

a) Testers shift their focus from manually designing test cases to guiding and verifying AI-generated testware.
b) Test managers shift their focus from managing test projects to understanding the inner working of Generative AI technologies.
c) Testers shift their focus from manually designing test cases to overseeing AI-based test processes.
d) Test managers shift their focus from relying on people to relying solely on Generative AI to boost productivity in test tasks.

Select ONE option.

Appendix: Additional Questions

Question #A1 (1 Point)

Consider the following steps (on the left) involved in implementing RAG (Retrieval-Augmented Generation):

Steps	Answer	
Clean and process document chunks	**First**	
Generate a response using retrieved chunks and the LLM		
Break large documents into smaller chunks		
Store embeddings in a vector database		
Retrieve relevant chunks based on semantic similarity	**Last**	

Order these steps in sequence from the first to the last.

Question #A2 (1 Point)

Match each prompting technique (on the left) to an appropriate example of its application in software testing (on the right):

Prompt Chaining	Breaking down test analysis tasks into smaller steps requiring iterative LLM interactions and human verification
Few-Shot Prompting	Generating Gherkin-style test cases from user stories using the LLM with pre-defined examples
Meta Prompting	Prioritizing test cases based on risk by leveraging the LLM's inherent reasoning without examples
Zero-Shot Prompting	Interacting with the LLM to iteratively refine prompts for creating test oracles from ambiguous requirements

Assign each individual prompting technique with an appropriate example of its application in software testing. No item may be left unpaired or paired with more than one other item.

Question #A3 (1 Point)

Consider the following activities (on the left) and key phases (on the right) involved in the adoption of Generative AI within a test organization:

Activities	Key phases
Identifying and prioritizing practical use cases	Discovery
Training testers on Generative AI	Initiation and Usage Definition
Providing testers access to LLMs	Utilization and Iteration
Experimenting with initial use cases to familiarize testers with Generative AI	
Managing the evolution of test processes following the full integration of Generative AI	

Assign each activity (item) to a key phase (group). No group can be left empty and no item can be assigned to more than one group.