DUMPSARENA

Certified Tester AI Testing Exam

ISTQB CT-AI

Version Demo

Total Demo Questions: 14

Total Premium Questions: 142

Buy Premium PDF

https://dumpsarena.co

sales@dumpsarena.co

sales@dumpsarena.co
dumpsarena.co

Topic Break Down

Topic	No. of Questions
Topic 1, Introduction to Artificial Intelligence (AI)	30
Topic 2, Testing AI-Based Systems Overview	8
Topic 3, Testing AI-Based Systems	11
Topic 4, Testing AI-Based Systems Quality Characteristics	40
Topic 5, Methods and Techniques for Testing AI-Based Systems	28
Topic 6, Testing AI-Based Systems in Production	6
Topic 7, Supporting the Testing of AI-Based Systems	19
Total	142

QUESTION NO: 1

Which statement regarding flexibility and adaptability of AI-based systems is correct?

Choose ONE option (1 out of 4)

A. Adaptability and flexibility are important when the system needs to change its behavior and determine the change on its own.

B. Adaptability is considered to be the ability of the system to be used in unspecified situations.

C. Self-learning AI-based systems are classified according to whether they are adaptable only or flexible only.

D. Flexibility is considered to be the ease with which the system can be reprogrammed to a changed operating condition.

ANSWER: A

Explanation:

“Adaptability and flexibility are important when the system needs to change its behavior and determine the change on its own.” is correct because AI-based systems often operate in environments where inputs, conditions, data patterns, or operational needs evolve after deployment. Flexibility supports the system’s ability to accommodate situations that were not fully specified during development. Adaptability concerns the system’s ability to alter its operation appropriately when conditions change. These characteristics are especially significant for self-learning or otherwise autonomous AI-based systems, because the system may need to identify that its current behavior is no longer suitable and adjust its behavior without a developer manually defining every individual response in advance.

From a testing perspective, this means testers need to consider how the system detects relevant changes, what limits govern its behavioral changes, how those changes are monitored, and whether the resulting behavior remains safe, effective, and aligned with intended use. The ISTQB AI Testing certification addresses the distinctive characteristics and risks of AI-based systems, including the need for testing approaches suited to systems whose behavior can evolve. See the ISTQB Certified Tester AI Testing certification page and the ISTQB glossary entry search for adaptability.

QUESTION NO: 2

Which two test procedures are BEST suited for CleverPropose system testing?

Choose TWO options (2 out of 5)

A. Back-to-back testing

B. Adversarial testing

C. Metamorphic testing

D. Exploratory data analysis

E. Pairwise testing

ANSWER: A C

Explanation:

Back-to-back testing is well suited where CleverPropose succeeds or replaces an established conventional advisory system. It runs the new AI-based system and a reference implementation on equivalent inputs, then compares their outputs. The existing system therefore supplies a practical test oracle or comparison baseline, which is especially valuable for decision-support functionality where deriving every expected result manually would be difficult. Differences can be investigated to determine whether they reflect intended new behaviour, model limitations, or defects.

Metamorphic testing is also particularly useful for AI systems because precise expected outputs are often unavailable or expensive to establish. Instead of asserting one exact recommendation for each input, testers specify relations that outputs should maintain when inputs are transformed in a meaningful way. For example, a controlled change to a relevant financial

characteristic can be expected to produce a consistent or bounded change in a proposal or risk-related result. These relations permit systematic testing despite an oracle problem and support coverage beyond a limited regression suite.

Both procedures are covered within the AI-testing techniques and oracle-related considerations of the ISTQB Certified Tester AI Testing syllabus. See the ISTQB CT-AI certification page and the official CT-AI documents page.

QUESTION NO: 3

An ML model was trained to predict energy consumption from temperature readings expressed in degrees Celsius. After deployment, a revised data pipeline supplies temperature readings in degrees Fahrenheit without converting them. The model produces implausible predictions although its implementation and trained weights have not changed.

Which ONE of the following is the MOST likely cause?

A. The production input data is not being preprocessed consistently with the training data

B. The model has become non-deterministic because it is operating in production

C. The model must use unsupervised learning to process temperature values

D. The model weights have changed because the input values have a different unit

ANSWER: A

Explanation:

The production input data is not being prepared consistently with the training data. The model learned relationships using values expressed in degrees Celsius. Supplying Fahrenheit values changes the numerical meaning and distribution of the input feature, so the learned relationship is applied incorrectly.

This is a data-pipeline and data-preparation defect rather than evidence that the trained weights have changed or that the model necessarily requires a different learning algorithm. The appropriate corrective action is to ensure that production preprocessing matches the preprocessing used for training and testing.

QUESTION NO: 4

Which of the following statements regarding experience-based testing for AI-based systems is correct?

Choose ONE option (1 out of 4)

A. Intuitive test case design for AI-based systems involves interactive, hypothesis-driven examination of data for correlations or developmental trends.

B. In checklist-based testing of AI-based systems, the existing test cases are dynamically adapted, for example based on metamorphic testing.

C. Exploratory testing is often used for AI-based systems because there are often insufficient specifications or problems with the test oracle for AI-based systems.

D. Tour refers to intuitive test case design for AI-based systems based on multiple, sequential test cases using systematically biased training data.

ANSWER: C

Explanation:

Exploratory testing is often used for AI-based systems because there are often insufficient specifications or problems with the test oracle for AI-based systems. This is correct because many AI systems, especially machine-learning-based systems, do not have fully specified expected results for every possible input. Their outputs can be probabilistic, context-sensitive, or difficult to determine independently, creating a test-oracle problem. In addition, requirements and expected behaviour may be incomplete, ambiguous, or continuously refined as data and models evolve. Exploratory testing addresses these

conditions by combining test design, execution, learning, and assessment in an interactive process. A tester can investigate model behaviour, form and revise hypotheses, select meaningful data variations, and use domain knowledge to identify anomalous, unsafe, biased, or otherwise unexpected outcomes. This makes it particularly valuable alongside more structured AI testing techniques, including data validation, metamorphic testing, and robustness testing. The ISTQB AI Testing syllabus identifies the limited availability of specifications and the oracle problem as important characteristics affecting testing of AI-based systems, and includes exploratory testing among the applicable experience-based techniques. See the ISTQB Certified Tester AI Testing certification page and the CT-AI syllabus resources.

QUESTION NO: 5

A team uses an AI-based tool to select a reduced regression test suite from historical defect data, previous test results, and recent code changes. The product includes safety-related functions that must be tested on every release.

Which TWO actions are MOST appropriate when using the tool's recommendations?

Choose TWO options (2 out of 5)

A. Retain mandatory safety-related tests regardless of the tool's ranking

B. Monitor selections against later defects and actual test outcomes

C. Remove all tests that the tool does not select without further review

D. Prioritize tests solely according to their shortest execution time

E. Keep the selection model unchanged after major product changes

ANSWER: A B

Explanation:

Mandatory safety-related tests must remain in the release suite even if the AI-based tool assigns them a low predicted defect-detection value. The selection must also be monitored against actual escaped defects and test outcomes so that stale patterns, poor predictions, and changes in the system can be identified and the model can be reassessed.

The tool's recommendation is decision support rather than proof that omitted tests have no value. Selecting tests solely because they execute quickly optimizes duration rather than risk coverage. Disabling human review and retaining a fixed model indefinitely would increase the risk that unsuitable historical patterns drive testing decisions.

QUESTION NO: 6

Which option describes a reasonable application of AIB testing for a self-learning system after it has changed its behavior due to user input?

Choose ONE option (1 out of 4)

A. Generating test cases for the system before and after the change, since neither has a test oracle

B. Comparing outputs before and after the change using different inputs

C. Comparing outputs before and after the change using identical inputs

D. Comparing outputs of a non-self-learning system with those of the changed self-learning system

ANSWER: C

Explanation:

Comparing outputs before and after the change using identical inputs is a reasonable application of AI behaviour (AIB) testing. A self-learning system can alter its learned model or decision behaviour after receiving user input or feedback. To assess the effect of that learning, testers run a defined, representative set of inputs against the system before the

behavioural change and run that same set again afterwards. They then compare the corresponding outputs. Holding the inputs constant makes the comparison meaningful: observed output differences indicate a change in the system's behaviour for those cases, rather than a difference caused by presenting different data.

This approach supports the evaluation of behavioural evolution where a conventional expected-result oracle may be difficult to establish for every individual AI output. It can reveal changed classifications, recommendations, predictions, rankings, or confidence values and provides evidence for deciding whether the learning has produced an acceptable behavioural change. The test set should be controlled and retained, with the system version and relevant learning conditions recorded, so that results are repeatable and differences can be investigated. This aligns with the CT-AI syllabus's treatment of AI behaviour testing for self-learning systems; see the ISTQB Certified Tester AI Testing certification page and the CT-AI syllabus.

QUESTION NO: 7

Which AI-specific test objective and acceptance criterion should be selected MOST LIKELY for testing GPT_Legal?

Choose ONE option (1 out of 4)

A. Test objective: Evidence of functional safety Acceptance criterion: The system recognizes failures in the transmission of information and data with the DPMA system and the evaluation system by means of self-tests.

B. Test objective: Evidence of evolution Acceptance criterion: The quality of the research results does not deteriorate with further training.

C. Test objective: Evidence of compatibility Acceptance criterion: The system can exchange information with the DPMA system and the evaluation system.

D. Test objective: Evidence that the data is free from inappropriate bias Acceptance criterion: The DPMA's analysis data is statistically compared to data from other sources.

ANSWER: B

Explanation:

Test objective: Evidence of evolution with the acceptance criterion **"The quality of the research results does not deteriorate with further training"** is the most appropriate selection for GPT_Legal. In the CT-AI syllabus, evolution is an AI-specific quality characteristic concerned with an AI system's ability to maintain or improve its performance as it is trained further and exposed to additional data. This is particularly relevant to a self-learning legal-research system, because its training is expected to continue after initial deployment and may incorporate new legal documents, terminology, research patterns, and operational contexts.

The stated criterion gives a clear, measurable regression-oriented objective: compare the quality of research results before and after further training and verify that the updated model has not degraded. Such testing can use a fixed, representative benchmark set with agreed quality measures, for example correctness, relevance, completeness, and false-positive rate. It directly supports safe model evolution by detecting harmful changes introduced by retraining while allowing validated improvements to be released.

This use of evolution as a test objective is consistent with the AI-specific test objectives described in the ISTQB Certified Tester AI Testing (CT-AI) certification and the associated CT-AI syllabus documents.

QUESTION NO: 8

A medical workflow routes images through an AI-based image-quality model and then through an ML model that identifies cases requiring specialist review. The team is preparing a test strategy for the complete workflow.

Choose TWO options (2 out of 5)

Which TWO activities are MOST appropriate?

A. Test each model independently against its relevant performance criteria

B. Test representative end-to-end cases through the integrated workflow

C. Test only the final workflow because component testing is unnecessary

D. Test only each individual model because integration cannot affect model behaviour

E. Use only perfect synthetic images to avoid variation in the input data

ANSWER: A B

Explanation:

Each model should be tested independently against its own relevant quality criteria, because a weakness in either model can affect the clinical workflow. The integrated workflow should also be tested with representative end-to-end cases, including cases in which the image-quality model rejects or accepts an image, because its output determines what reaches the downstream model.

Testing only the final workflow can conceal the source of a failure. Testing only individual models does not reveal interface and sequencing effects. Replacing representative clinical cases with synthetic perfect images would reduce evidence for operational performance.

QUESTION NO: 9

You are developing a “flower” ML model… Which of the following describes an objection that you can NEGLECT in your risk assessment?

Choose ONE option (1 out of 4)

A. The possible inputs for the ‘leaf’ and ‘flower’ ML models are so different that reuse has few advantages over new development.

B. The probability of misclassification of the ML model "flower" is higher when it is reused than when it is developed from scratch.

C. The classification behavior of the "flower" ML model is more difficult to understand when it is reused compared to when it is developed from scratch.

D. The possible outputs of the "leaf" and "flower" ML models are so different that reuse has few advantages over new development.

ANSWER: D

Explanation:

The possible outputs of the "leaf" and "flower" ML models are so different that reuse has few advantages over new development. is the objection that can be neglected. In this transfer-learning scenario, the existing leaf-based model and the intended flower-based model are used to identify the same plant species. Although the observable features supplied to each model differ—leaves versus flowers—the required classification result remains the plant-species label. Therefore, there is no relevant output-label mismatch between the source and target tasks.

Transfer learning reuses knowledge learned in one task to support another task. Its suitability depends substantially on the relationship between the original and intended tasks, including the nature of their input data and whether learned patterns can transfer effectively. When the target output classes are unchanged, differing output spaces do not create an additional reuse concern. The shared output objective means that the model can still be adapted to classify the same species from a new visual representation. This makes the stated output-difference objection inapplicable to the described risk assessment.

The ISTQB Certified Tester AI Testing syllabus discusses pre-trained models and transfer learning, including the importance of similarity between source and target tasks: ISTQB Certified Tester AI Testing.

QUESTION NO: 10

A company suspects that an attacker has inserted manipulated records into data collected for retraining an ML model. The manipulated records are intended to cause a particular category of transactions to be misclassified.

Choose TWO options (2 out of 5)

Which TWO activities are MOST appropriate for investigating this risk?

A. Use exploratory data analysis to investigate outliers and unexpected data distributions

B. Check the provenance of records and validate their approved data sources

C. Increase the number of layers in the ML model

D. Remove all rare records before investigating them

E. Repeat the same functional test until identical output is obtained

ANSWER: A B

Explanation:

Exploratory data analysis can reveal unexpected distributions, anomalous clusters, and outliers that may indicate manipulated data. Checking data provenance and validating that records originate from approved, controlled sources also addresses the possibility that untrusted or altered records have entered the retraining data.

Increasing the number of model layers does not investigate the integrity of the data. Removing all rare records would risk discarding valid but important cases. Executing the same test repeatedly may address probabilistic behaviour but does not establish whether the training data was manipulated.

QUESTION NO: 11

A hospital intends to reuse a pre-trained image-classification model to prioritize chest X-rays for specialist review. The model was trained by another organization, and details of its original training data are limited.

Which TWO activities are MOST appropriate before relying on the model in the hospital?

Choose TWO options (2 out of 5)

A. Test the model using representative X-rays from the intended hospital context

B. Evaluate error patterns for relevant patient groups and image conditions

C. Accept the published overall accuracy as sufficient evidence of suitability

D. Assume that inherited bias cannot occur because the model is already trained

E. Retrain the model immediately without first establishing a target-context baseline

ANSWER: A B

Explanation:

Testing the model with representative images from the intended hospital context is necessary because performance established in an unknown source context does not demonstrate suitable performance for the target population, equipment, image protocols, and operating conditions. Evaluating results for relevant patient groups and image conditions can expose inherited bias or systematic weaknesses.

Documentation from the supplier may be useful, but it cannot replace independent target-context testing. Treating pre-training as proof of correctness ignores the risk of inherited defects and bias. Retraining automatically without evidence is not a test activity and may introduce additional uncontrolled changes.

QUESTION NO: 12

An autonomous warehouse robot uses an AI-based vision system to navigate around workers, shelves, and moving carts. The test strategy must address AI-specific quality risks in addition to conventional functional testing.

Which TWO test activities BEST address these risks?

Choose TWO options (2 out of 5)

A. Test navigation under shadows, glare, partial occlusion, and sensor noise

B. Test the robot's reaction when an unexpected obstacle enters its path

C. Execute only routes that have no moving obstacles

D. Require identical movement decisions on every repeated execution

E. Use code coverage as the sole acceptance criterion for safe navigation

ANSWER: A B

Explanation:

Testing the robot with varied and degraded visual conditions assesses robustness. Conditions such as shadows, glare, partial occlusion, and sensor noise can cause a vision model to behave differently from its behaviour under ideal test conditions. Testing safe reactions to unexpected obstacles is essential because the system operates autonomously in a dynamic physical environment where unsafe decisions can have serious consequences.

Executing only nominal routes provides insufficient evidence for robustness or safety. Requiring identical outputs for every execution disregards possible AI-system non-determinism and does not define safe behavioural limits. Code coverage can be useful for conventional software components, but it does not by itself demonstrate that the AI behaviour is safe in the relevant operational conditions.

QUESTION NO: 13

Which statement about automation bias is correct?

Choose ONE option (1 out of 4)

A. When testing AI-based systems, automation bias does not play a role in supporting test activities such as boundary value analysis

B. Automation bias affects the testing of AI-based systems that support users in their actions or decisions

C. Automation bias particularly affects testing of autonomous systems

D. Automation bias is tested with representative users, but human input quality is irrelevant

ANSWER: B

Explanation:

Automation bias affects the testing of AI-based systems that support users in their actions or decisions. Automation bias is the tendency for people to place excessive trust in the output, recommendations, or alerts of an automated system, potentially accepting them without sufficient independent review. It is particularly relevant to AI-enabled decision-support systems, in which a user remains responsible for taking an action or making a decision but is influenced by an AI-generated recommendation, prediction, ranking, or warning.

For AI testing, this human factor must be considered when defining realistic test scenarios and evaluating the system with representative users. Testers should investigate whether users appropriately understand the AI system's role, can recognize uncertainty or limitations in its outputs, and retain effective oversight rather than simply deferring to an automated recommendation. This supports the CT-AI syllabus emphasis on human factors and on testing AI systems in their intended operational context, including interactions between the system and its users. The relevant terminology and principles are covered by ISTQB's Certified Tester AI Testing certification materials. Further background on automation-related human performance considerations is available from the NIST AI Risk Management Framework.

QUESTION NO: 14

An image-recognition system controls access to a restricted facility. Security testing has identified image inputs that are deliberately modified in small ways but cause the system to grant access incorrectly.

Choose TWO options (2 out of 5)

Which TWO actions are MOST appropriate after identifying these adversarial examples?

A. Use the examples to assess and improve relevant defences against the vulnerability

B. Retain the examples as controlled regression tests for later system changes

C. Ignore the examples because they are unlikely to occur in ordinary use

D. Add every example directly to training data without reviewing its suitability

E. Disable all access decisions permanently

ANSWER: A B

Explanation:

The identified adversarial examples should be used to assess and improve the system's defences, such as input validation, preprocessing, model robustness, and safe handling of uncertain results. They should also be retained as controlled regression tests so that later changes can be checked for reintroduction of the vulnerability.

Ignoring the examples leaves the identified vulnerability unaddressed. Treating every adversarial example as ordinary valid training data without review can introduce unsuitable or manipulated data. Disabling all access decisions would avoid the immediate failure but would not provide a proportionate corrective action for the AI system.