ASTQB Certified Tester Testing with Generative AI v1.0 Sample Exam Answers

American Software Testing Qualifications Board

46 Points Possible : Passing Grade = 30 Points
GenAI – Sample Exam

1. (1 pt) Which of the following uses deep learning techniques to create new content?

a. Symbolic AI
b. Classical machine learning
c. Deep learning
d. Generative AI

D is correct. Generative AI uses deep learning techniques to create new data by learning and mimicking patterns from the training data. (GenAI-1.1.1 (K1) Recall different types of AI: symbolic AI, classical machine learning, deep learning, and generative AI)

2. (1 pt) You have used the same prompt multiple times, but you have received different responses for the LLM. Why is this happening?

a. Because the prompt is incorrect
b. Because LLMs have non-deterministic behavior
c. Because the training data was faulty
d. Because LLMs are trained to always vary the response

B is correct. LLMs exhibit non-deterministic behavior due to the probabilistic nature of their inference mechanisms. (GenAI-1.1.2 (K2) Explain the basics of generative AI and large language models)

ASTQB

3. (1 pt) How is tokenization related to embeddings?

a. Tokens are a numeric representation of how a given word is embedded into the model
b. Tokens represent characters, subwords, or words, and embeddings capture the semantic, syntactic, and contextual relationship of the tokens
c. Embeddings describe the text component in such a way that it can be tokenized for use within the model
d. Embeddings identify the non-deterministic behavior of a word component, which is then used to determine the positioning of its token using a vector to represent its position

B is correct. Because the syllabus says so. (GenAI-1.1.2 (K2) Explain the basics of generative AI and large language models)

4. (1 pt) You need an LLM that will help with decision-making based on a process similar to human logic. Which type of LLM should you use?

a. Foundation LLM
b. Instruction-tuned LLM
c. Reasoning LLM
d. Ad-hoc LLM

C is correct. You need a reasoning LLM because that training method is designed to emulate human-like reasoning responses. (GenAI-1.1.3 (K2) Distinguish between foundation, instruction-tuned and reasoning LLMs)

5. (1 pt) Which of the following is an advantage of using an LLM augmented with vision-language models for developing test cases?

a. There is no advantage because this type of model is used primarily for military applications
b. It can create test cases that can include text and visual cues
c. It can edit images much faster and more accurately than a human
d. It can generate executable code from wireframes and screenshots

B is correct. It can create test cases that include textual data and visual cues, which increases the overall test coverage. (GenAI-1.1.4 (K2) Summarize the basic principles of multimodal LLMs and vision-language models)

6. (1 pt) You have created test automation that will be used to test driver's license applications. There is a huge amount of data available for this testing, but it will take too long to test with all of that data. How can an LLM help with this problem?

a. It can create new synthetic data that will protect private information
b. It can take the existing data and augment that with data that is missing according to the data rules
c. It can take the existing data and break it into smaller sets that will contain interesting variances
d. It can delete data that is duplicated and replace it with new data

C is correct. It can look at the existing data, analyze it, and break it into smaller sets that will contain representative and interesting data. (GenAI-1.2.1(K2) Give examples of key LLM capabilities for test tasks)

7. (1 pt) Your organization is developing a test management tool that will use AI to generate, classify, and store test cases based on requirements, general testing knowledge, and previous issues found. Which of the following interaction models would be best to use?

a. Chatbots talking to the LLM
b. Chatbots talking to other chatbots that have been trained on your system
c. APIs of the LLM to allow the test task information to pass into and out of the LLM
d. APIs of the LLM to pass the test case information into the LLM to be stored there

C is correct. Using the APIs of the LLM will allow maximum usage of the LLM by passing the necessary data to and from it. D doesn't make sense because the test case information would be stored in the test management system, not in the LLM. A and B are incorrect because chatbots would not be the most efficient or flexible access. (GenAI-1.2.2 (K2) Compare interaction models when using GenAI for software testing)

8. (1 pt) Which of the following are the six components of an effective prompt?

a. Prompt Chain, Persona, Instruction, Assumptions, Constraints, Output Format
b. Role, Bias, Context, Instruction, Input Data, Constraints, Output Data
c. Persona, Background, Instruction, Input Data, Output Data, Assumptions
d. Role, Context, Instruction, Input Data, Constraints, Output Format

D is correct. These are the components. (GenAI-2.1.1 (K2) Give examples of the structure of prompts used in generative AI for software testing)

9. (1 pt) You have provided a structured prompt to a GenAI model regarding the creation of test cases. You have specified that you want to receive test cases in the following format:

- **Description of test**
- **Reference to source**
- **Steps to complete**
- **Expected outcome**

This is an example of what component of a structured prompt?

a. Role
b. Context
c. Instruction
d. Output format

D is correct. This is the output format that you are specifying. (GenAI-2.1.1 (K2) Give examples of the structure of prompts used in generative AI for software testing)

10. (1 pt) If you have a test process with complex tasks and you will need to check intermediate results before proceeding, which prompting technique is the most appropriate?

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

A is correct. Prompt chaining allows the next prompt to be tuned based on the results of the preceding step. (GenAI-2.1.2 (K2) Differentiate core prompting techniques for software testing)

11. (1 pt) Which of the following is true regarding a user prompt for an LLM?

a. It is set once at the start of the conversation
b. It is usually set by the developer
c. It usually changes with each interaction
d. It is user-independent, so the response will be consistent regardless of how the user words the prompt

C is correct. The user prompt is the interaction with the system, and this usually changes with each interaction. A and B are correct for system prompts. (GenAI-2.1.3 (K2) Distinguish between system prompts and user prompts)

12. (2 pts) You are working on a replacement of an e-commerce system. You have been given a large set of test conditions to test. There is not adequate time to test all the conditions. You have a testing-trained LLM that you can use. How can you best use it in this situation?

a. The LLM can create an estimate for how long testing will take based on similar test conditions. This schedule can then be incorporated into the overall project schedule to allow adequate time for testing.
b. The LLM can be fed information regarding previous defects, end-user requirements, and payment regulations to let it determine the relative risk of all the test conditions. This can then be used to prioritize the testing.
c. The LLM can generate and execute all the test cases based on the test conditions, which will save significant time.
d. The LLM can generate test data to cover every possible purchasing scenario. This will ensure testing will not miss any corner cases, and that will help the schedule by eliminating any testing escapes.

B is correct. Since there isn't enough time, the tests will have to be prioritized and executed in priority order. When time runs out, at least the most important tests will have been executed. A is not correct because there's already not enough time – making a schedule proving that won't help. C is not correct. The LLM could generate the test cases, but it won't know how to execute them. It might be able to develop test automation from the generated test cases, but that's not likely to shorten the time required for testing. D is not correct because it will make testing take longer. (GenAI-2.2.1 (K3) Apply generative AI to test analysis tasks)

13. (2 pts) You have been given requirements in the form of wireframes of the GUI and user stories. You want to use these to create the test conditions for a set of tests. How should you use prompts to evaluate the inputs and create the test conditions?

a. Apply multimodal prompts
b. Set well-defined system parameters to define the graphics to be used
c. Apply prompt chaining
d. Use meta prompting so the prompts can be continuously refined

A is correct. In this case, you will definitely need multimodal prompts because the data you are providing is both graphics and text. B is not correct because this is not something that is defined in system parameters. C is not correct because, also, prompt chaining may be useful; you first need to get the information into the system via A. D is not correct because it's not a matter of generating additional prompts. (GenAI-2.2.2 (K3) Apply generative AI to test design and test implementation tasks)

14. (2 pts) You have been testing versions of the same software over several years. Rather than the quality improving, it seems to just stay the same, with the same high number of defects being caught during testing for each release. How could AI help with this problem?

a. It could analyze test failures and prioritize the tests that failed last time so they are executed first for each new release.
b. It could use meta prompting to review each test case to determine if it is an effective test. Ineffective tests could then be removed from the overall test suite.
c. It could create large sets of test data that are used across releases of the product to ensure that data variance is not causing the repeated failures.
d. It could conduct root cause analysis over all the defects found across releases to determine the common causes. This information could then be used to improve development and analysis processes.

D is correct. We don't know what's causing these issues, and we are catching them in testing, but they shouldn't be occurring in the first place. Root cause analysis will tell us what is breaking and why. (GenAI-2.2.3 (K3) Apply generative AI to automated regression testing)

15. (2 pts) You are working on a critical project that is under the scrutiny of senior management. As a result, simple and clear reporting of testing progress is paramount. You are using the test reporting capabilities of your test management system, but any analysis of the issues and trends has to be done manually by you. How could AI help with this situation?

a. AI could analyze all the open defects and verify that the prioritization is correct based on keywords in the defect report.
b. AI could track which test cases are finding the most defects, so you could highlight those on your reports.
c. AI could analyze the defect and test case execution information and provide trending information and root cause analysis mapped to the use cases for the software.
d. AI would automatically email the existing dashboards to the pertinent executives to ensure they always have the latest information available. For example, defect reports would go to the development management, defect reports to the test management, etc.

C is correct. AI can do analysis of the large amount of information to pull out the necessary trending and root cause analysis to provide more information other than just the numbers. By mapping the findings to the use cases, the importance of the failures would be easier to assess. A would be useful, but doesn't address the problem in the question. B might be useful, but it would only give part of the picture because you wouldn't know if those defects mattered. D is not correct because that would be annoying and this information is readily available already. (GenAI-2.2.4 (K3) Apply generative AI to test control and monitoring tasks)

16. (2 pts) You are working on testing a new system. The requirements are high-level and lack detail. What prompting technique should you use to get started to get some ideas for how to progress with test case generation?

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

C is correct. In this case meta prompting is a good starting point. You have the general description of what the software should do and what a test case should accomplish, so those inputs can be used. AI can then respond, and together you can build more test cases. (GenAI-2.2.5 (K3) Select and apply appropriate prompting techniques for a given context and test task)

17. (1 pt) You are using GenAI to create test cases for you for a heavily rules-based application. You have requested that it use decision tables to build the test cases. Which metric would be the best one to use to determine how well it has done with this request?

a. Accuracy
b. Recall
c. Diversity
d. Time Efficiency

B is correct. The recall metric evaluates the completeness of the generated output with respect to a specific objective, in this case, using decision tables. (GenAI-2.3.1 (K2) Understand the metrics for evaluating the results of Generative AI on test tasks)

18. (1 pt) You have been iteratively modifying your prompt for GenAI test case creation. You are still getting the same types of issues with the output, even though you keep refining the prompts and adding more information. What technique can you employ to use those issues to better evaluate and refine the prompts?

a. Iterative prompt modification
b. A/B testing of prompts
c. Output analysis
d. Adjusting prompt specificity

C is correct. At this point, you need to do some analysis on the issues to determine why you are seeing them occur repeatedly. (GenAI-2.3.2 (K2) Give examples of techniques for evaluating and iteratively refining prompts)

19. (1 pt) Which of the following is a result of a reasoning error?

a. An output that favors certain types of information, approaches, or assumptions
b. An output that is based on misinterpretations of logical structures
c. An output that appears plausible but is factually incorrect or irrelevant
d. An output that is only partially complete

B is correct. This is an example of a reasoning (GenAI-3.1.1 (K1) Recall the definitions of hallucinations, reasoning errors and biases in Generative AI systems)

20. (2 pts) You are using GenAI to develop a test plan. You have fed it the requirements, high-level project plan, schedules, and tester profiles. Despite your instructions, when you ask it to generate the test schedule, it creates a new team member and puts them on the schedule to make the testing effort fit the project timeline. What is the problem, and what should you do about it?

a. This is likely a bias problem. You should give it a wider variety of requirements for similar systems so it generates a larger scope of tests.
b. This is likely a hallucination. You should instruct the AI model that this is incorrect so it learns.
c. This is likely a reasoning error. You should run the test cases against the test objects to verify if the expected results are valid.
d. This is likely a temperature problem. You should lower the temperature to allow a more specific focus.

B is correct. This is a hallucination where the AI model has created a new team member for you. While this might be handy, it won't really help with the schedule and the AI model must be taught that this is incorrect. (GenAI-3.1.2 (K3) Identify hallucinations, reasoning errors and biases in LLM output)

21. (1 pt) In what way does comparing results across several different models help reduce hallucinations, reasoning errors, and biases?

a. You can combine the results into a superset of data, which will reduce the errors in the data
b. You can select the data that occurs with the highest frequency and know that it is the most accurate
c. You can compare different outputs from the same prompts to detect output errors and reliability
d. You can get the results faster by using multiple models concurrently

C is correct. By comparing the results from different models using the same prompts, you can tune the prompts, compare the outputs, and pick the most reliable results. (GenAI-3.1.3 (K2) Summarize mitigation techniques for GenAI hallucinations, reasoning errors and biases in software test tasks)

22. (1 pt) How does setting a random seed value help improve reproducibility in LLM outputs?

a. By setting a lower seed value, the temperature setting is automatically increased, which narrows the probability distribution
b. By setting a higher seed value, more seeds are allowed to be used, which will promote larger output sets from the LLM, which can then be used to extract consistent data, which improves reproducibility
c. By setting a variable seed value, the random sequence values are increased, which results in a larger set of data, which is easier to recreate
d. By setting a seed value, the random number generator has the same starting point, which will ensure the same pseudo-random sequence is used, which improves reproducibility

D is correct. The seed value is used for the random number generator. If it always starts from the same seed value, it is more likely to use the same pseudo-sequence, which will improve the reproducibility of the outputs. (GenAI-3.1.4 (K1) Recall mitigation techniques for non-deterministic behavior of LLMs)

23. (1 pt) If you have a GenAI tool that is processing sensitive customer data, what must you ensure?

a. It cannot store the data in an unencrypted form because it will be readable by anyone
b. It cannot store the data where it might be accessed by an unauthorized user, including another GenAI tool
c. It cannot modify and then restore the data because it will be incorrect and might be misleading for other users
d. It cannot access any sensitive data for any use

B is correct. Presumably, it's accessing the data because it needs to, but the risk is that it might store it somewhere that it could be accessed by another GenAI tool. (GenAI-3.2.1 (K2) Explain key data privacy and security risks associated with using generative AI in software testing)

24. (1 pt) You've been working on an application that processes employee leave requests. The data you are using contains the birth date of the employee, but you have replaced all the birth dates with a default date of 1/1/89. You are using GenAI within your company's databases to build new test data. You have now discovered that the test data contains birthdates again, and they are the real dates. This is an example of what type of data privacy issue?

a. Lack of control over data usage
b. Vulnerability to security attacks
c. Evidence of a malicious introduction of false data
d. Unintentional data exposure

D is correct. This is an unintentional exposure of private data – in this case, the birthdates of the employees. A is not correct because the control is still within the company, but there is clearly access to some of the HR data. C and D are not correct because there's no indication that this was malicious. (GenAI-3.2.2 (K2) Give examples of data privacy and vulnerabilities in using Generative AI in software testing)

25. (1 pt) What type of attack vector is being used when the results of AI-generated data are rated incorrectly, intentionally?

a. Data exfiltration
b. Request manipulation
c. Data poisoning
d. Malicious code generation

C is correct. This is an example of data poisoning. (GenAI-3.2.2 (K2) Give examples of data privacy and vulnerabilities in using Generative AI in software testing)

26. (1 pt) What is the purpose of data anonymization?

a. To only use and process the data that is legally permitted and use only the smallest amount needed
b. To mask or replace any private information with non-identifiable data
c. To implement strong encryption and strictly control data access
d. To ensure everyone is trained in the proper handling of data and is minimizing access

B is correct. (GenAI-3.2.3 (K2) Summarize mitigation strategies to protect data privacy and enhance security in Generative AI for software testing)

27. (1 pt) Which of the following GenAI activities takes about as much power as charging a smartphone?

a. Generating a letter to your bank
b. Creating “business” penguin pictures for a presentation
c. Creating only one penguin image
d. Generating a financial year summary report

C is correct. Creating just one image uses about the same power as charging a cell phone. (GenAI-3.3.1 (K2) Explain the impact of task characteristics and model usage on the energy consumption of Generative AI in software testing)

28. (1 pt) Which of the following establishes a legal framework addressing AI risks and classifying applications by risk level?

a. ISO/IEC 42001:2023
b. ISO/IEC 23053:2022
c. EU AI Act
d. NIST AI Risk Management Framework

C is correct. D, which sounds right, concentrates on managing the risk, focusing on fairness, transparency, and security, so it is not the correct answer. (GenAI-3.4.1 (K1) Recall examples of AI regulations, standards, and best practice frameworks relevant to Generative AI in software testing)

29. (1 pt) Which part of the architecture in an LLM-powered test infrastructure is responsible for integrating multiple data sources?

a. The front-end
b. The LLM itself
c. The back-end
d. The data controller

C is correct. The back-end integrates multiple data sources. (GenAI-4.1.1 (K2) Explain key architectural components and concepts of LLM-powered test infrastructure)

30. (1 pt) What are the two steps in RAG processing?

a. Input / Retrieval
b. Processing / Output
c. Access / Processing
d. Retrieval / Generation

D is correct. (GenAI-4.1.2 (K2) Summarize Retrieval-Augmented Generation)

31. (1 pt) What does "orchestration" mean regarding LLM-powered agents?

a. It is when several agents work together in a coordinated manner to more efficiently solve complex problems
b. It is when more than three LLM-powered agents work independently to generate outputs that are then reviewed by another agent for accuracy and post-processing
c. It is when RAG is an integral part of an LLM-powered test infrastructure
d. It is a term for background processing that uses a combination of autonomous and semi-autonomous agents

A is correct. B, C, and D are fictional. (GenAI-4.1.3 (K2) Explain the role and application of LLM-powered agents in automating test processes)

32. (1 pt) How can overfitting compromise your model during fine-tuning?

a. It can make the LLM too generalized to be applicable and efficient
b. It can introduce hallucinations and biases
c. It can make the model unable or inaccurate when processing data that is new to it
d. It can create a lack of transparency regarding LLM decision-making

C is correct. Overfitting is when the model has become too specific and is unable to deal with data that is not within what it already knows. (GenAI-4.2.1 (K2) Explain the fine-tuning of language models for specific test tasks)

33. (1 pt) Which of the following LLM deployment types is most concerned with data privacy?

a. Chatbots
b. Commercial testing tools with AI components
c. In-house developed commercial test tools with AI components
d. Data privacy is a critical concern for all of these deployment types

D is correct. It depends more on the data being handled than the process that is handling it, so potentially any of these could be most critical. (GenAI-4.2.2 (K2) Explain LLMOps and its role in deploying and managing LLMs for test tasks)

34. (1 pt) How is the risk of using copyrighted data without proper authorization categorized when using shadow AI?

a. Information security weakness
b. Data privacy weakness
c. Compliance issues
d. Vague intellectual property

D is correct. This is an issue because these tools may have unclear licensing agreements and may use copyright data without proper authorization. (GenAI-5.1.1 (K1) Recall the risks of shadow AI)

35. (1 pt) Why is high-quality input data important when implementing an LLM-powered testing approach?

a. It will be faster with good data that produces fewer errors
b. It will produce reliable results
c. It will allow the largest possible data set to be built
d. It will ensure access is protected and limited to only the correct users

B is correct. (GenAI-5.1.2 (K2) Explain the key aspects to consider when defining a Generative AI strategy for software testing)

36. (1 pt) You are working to convince your manager that a GenAI tool would be helpful in generating test reports. One of his issues is that he feels he could hire interns to write the test reports, and that would be faster than incorporating a new tool. You are sure that it will be able to write the reports faster and more accurately, freeing critical testing time.

What is an important capability of the LLM that you should evaluate to be sure it will prove you right?

a. Model performance
b. Fine-tuning potential
c. Direct cost
d. Community and support

A is correct. You need to be sure the model will perform efficiently and accurately in the report generation tasks. (GenAI-5.1.3 (K2) Summarize key criteria for selecting LLMs/SLMs for software test tasks in a given context)

37. (1 pt) What is the first key phase when adopting GenAI into the test organization?

a. Discovery
b. Initiation and Training
c. Usage Definition
d. Exploitation and Iteration

A is correct. (GenAI-5.1.4 (K1) Recall key phases in the adoption of Generative AI in a test organization)

38. (1 pt) Why does a tester need to understand "prompt engineering" in order to work effectively with GenAI?

a. Because they will have a better grasp of the internal layers of the neural network
b. Because they need to understand what user prompts will need to be tested and the relative risk of each
c. Because they will be asked to generate new user prompts based on the AI-designed UI
d. Because they will need to design and refine the input prompts to guide the LLM

D is correct. (GenAI-5.2.1 (K2) Explain the essential skills and knowledge areas required for testers to work effectively with generative AI in test processes)

39. (1 pt) When conducting GenAI training meetings with the testing team, what is an important learning method?

a. Discussing challenges and their resolutions
b. Highlighting erroneous prompts used within the team
c. Processing AI outputs to better analyze manual testing gaps
d. Promoting the use of AI methods and nominating an "AI champion of the week"

A is correct. B, C and D might also be informative, but one of the key methods highlighted in the syllabus is to discuss challenges as a team. (GenAI-5.2.2 (K1) Recall strategies for cultivating AI skills within test teams to support the adoption of Generative AI in test activities)

40. (1 pt) Who is usually responsible for ensuring that test teams maintain both traditional testing competencies and AI literacy?

a. Each tester
b. The senior testers
c. The test manager
d. The AI master

C is correct. This is usually the role of the test manager to be sure that the team as a whole has these capabilities. Individuals are responsible for their own capabilities and ensuring they are staying current with the industry changes. (GenAI-5.2.3 (K1) Recognize how test processes and responsibilities shift within a test organization when adopting Generative AI for testing)