

PRACTICE QUESTIONS

Cisco

810-110 AITECH

Cisco AI Technical Practitioner

SHARED BY [u/kind_byte_46](#) on aiexamcerts

EDITION Free Set 1

QUESTIONS 10

SHARED October 15, 2026

EXAM Cisco 810-110 AITECH, Cisco AI Technical Practitioner

Shared by u/kind_byte_46 on aiexamcerts

10 practice questions for the Cisco 810-110 AITECH exam, Cisco AI Technical Practitioner, each followed by its answer and a short explanation. More free sets and the full practice exam:

www.aiexamcerts.com/cisco/810-110-aitech. Spotted a mistake? Email support@aiexamcerts.com with the exam code 810-110 AITECH and the question number.

1. What is the key difference between a foundation model and a task-specific model?
 - A A foundation model only handles images, while a task-specific model only handles text
 - B A foundation model cannot be fine-tuned, while a task-specific model can
 - C A foundation model is broad and adaptable, while a task-specific model is narrow
 - D A foundation model always runs on a CPU, while a task-specific model needs a GPU

Answer: C

Foundation models learn general capabilities from broad data and can be adapted to many downstream tasks. A narrow task-specific model is built for one job, which makes the first option the best description. The claims about modality, fine-tuning, and hardware are inaccurate, and fine-tuning is common for foundation models.

2. The maximum amount of text, measured in tokens, that a model can consider in a single request is called the context _____.

Answer: window (also accepted: length)

The context window bounds how many input and output tokens a model can handle in one exchange, so long documents may need chunking or retrieval. Context length is a common synonym. Parameter count is a different measure of model size.

3. A prompt must return JSON that a downstream parser can read. Which instruction is most effective?
 - A Ask the model to be as friendly as possible in its reply
 - B Ask for a long essay and extract the JSON from it later by hand manually
 - C Request the fields in a random order each time
 - D Specify exact keys, types, and allowed values, and require only JSON

Answer: D

An explicit schema, types, and an output-only constraint give the parser a predictable target and make failures easy to detect. A friendly tone does not affect structure, which makes it the most tempting but ineffective option. Long essays and random field order make parsing harder.

4. Asking a model to reason step by step before giving a final answer can improve accuracy on multistep problems.

- A True
- B False

Answer: A (True)

Chain-of-thought style prompting encourages intermediate reasoning, which often helps with arithmetic, logic, and multistep tasks. The size of the gain varies by model and task, so results should be tested. The statement is True.

5. Which two controls reduce the impact of a compromised or misbehaving agent? (Choose two.)

- A Sandboxing tool execution so the agent cannot reach unrelated systems
- B Giving the agent administrator credentials for convenience
- C Requiring confirmation for destructive or external actions
- D Letting the agent edit its own permission list
- E Sharing one service account across all agents and tasks

Answer: A, C

Sandboxing and confirmation gates limit what a misbehaving agent can reach or change. Administrator credentials and shared service accounts widen the blast radius, and they are the most tempting convenience. Letting the agent edit its own permissions removes the boundary entirely.

6. A workflow extracts invoice fields with a model and writes them to an accounting system. Which design choice matters most?

- A Validate extracted fields against rules such as totals before writing
- B Write every extracted value to the ledger immediately because the model is accurate
- C Skip logging so the workflow runs faster
- D Store all invoices in one unstructured text field

Answer: A

Validating totals, dates, and required fields before writing protects the ledger from silent extraction errors. Assuming the model is accurate is the most tempting shortcut, since extraction still makes mistakes. Skipping logs removes traceability, and unstructured storage defeats the purpose of extraction.

7. Why should an agent's tools be granted the least privilege needed for the task?
- A So the agent runs faster because it has fewer permissions to check on each call
 - B So a mistake or injected prompt cannot cause more damage than needed
 - C So the model can learn the tool names during training
 - D So the agent does not need any logging

Answer: B

Limiting tool permissions contains the damage from errors or injected instructions, which is the core principle of least privilege. Speed is the most tempting misreason, since permission checks are rarely the bottleneck. Permissions do not teach the model tool names, and logging is still needed.

8. A workflow step that runs only after the previous step succeeds and passes its output forward is called a _____ step.

Answer: sequential (also accepted: dependent, chained)

Sequential or chained steps depend on earlier outputs, which keeps the flow predictable and easier to debug. Parallel steps run independently at the same time, which is the most tempting contrast. Chaining needs error handling so one failure does not corrupt later steps.

9. A model's outputs help decide loan applications. Which governance practice is most appropriate?
- A Deploy it without documentation because the model is proprietary
 - B Allow decisions without any appeal route to protect processing speed
 - C Document the purpose, test for disparate impact, and name an owner
 - D Measure only the average approval rate across all applicants

Answer: C

Documentation, fairness testing, a named owner, and an appeal process are core governance controls for consequential decisions. Being proprietary does not exempt a model from explanation and fairness duties, which makes it the most tempting excuse. Averages alone can hide disparities between groups.

10. Which capability is most useful when an AI assistant must answer questions about a company's internal documents?
- A A larger GPU allocated to the chat front end
 - B A faster internet connection for each employee
 - C A higher temperature setting in the sampling configuration
 - D A search index over approved documents that supplies passages

Answer: D

A search index over approved documents supplies relevant passages, which grounds answers in company sources. Extra GPU capacity and connection speed are infrastructure upgrades rather than a way to supply knowledge. A higher temperature only changes how varied the wording is, and it adds no information.