AWS Certified AI Practitioner (AIF-C01)

100 Questions & Answers

Welcome to your complete AWS Certified AI Practitioner (AIF-C01) practice questions collection. This set is designed not just to quiz you, but to teach, reinforce, and sharpen your ***exam readiness using real-world AWS AI scenarios.***

Learning Objectives and Expectations

You'll get:

- Realistic, scenario-based questions modeled after the AWS certification exam structure
- Organized in **batches of 10 questions, followed by 10 answers** and explanations
- Clear answer keys with short explanations to reinforce core concepts, AWS services, and decision logic

AWS AI Practitioner (AIF-C01) Domains

Each domain carries a specific weight. Domains 2 and 3 (Generative AI and Foundation Model Applications) are the largest:

- **Domain 1:** Fundamentals of AI and ML – 20%
- **Domain 2:** Fundamentals of Generative AI – 24%
- **Domain 3:** Applications of Foundation Models – 28%
- **Domain 4:** Guidelines for Responsible AI – 14%
- **Domain 5:** Security, Compliance, and Governance – 14%

Quick Reminder: How the Exam Works

- **Number of Questions:** 65
- **Format:** Multiple choice, multiple response, matching, and scenario-based
- **Time Limit:** 90 minutes
- **Passing Score:** 700/1000
- **Test Provider:** Pearson VUE (online or onsite)

Questions By Domain

Domain	Title	Questions Assigned	Question Numbers
Domain 1	Fundamentals of AI and ML (20%)	20 Questions	Q1–2, Q12, Q14, Q27, Q31, Q34–35, Q42, Q52, Q56, Q63, Q66, Q72–73, Q74, Q94, Q97
Domain 2	Fundamentals of Generative AI (24%)	24 Questions	Q3–4, Q10, Q16, Q21, Q26, Q36, Q45–46, Q55, Q58, Q60, Q64, Q70, Q75, Q79–80, Q84–85, Q88, Q98–100
Domain 3	Applications of Foundation Models (28%)	28 Questions	Q5–6, Q9, Q13, Q19–20, Q22–23, Q25, Q30, Q33, Q37, Q47, Q50, Q57, Q59, Q61, Q67, Q69, Q71, Q76–77, Q82, Q86–87, Q90–91, Q95
Domain 4	Responsible AI (14%)	14 Questions	Q7, Q11, Q15, Q28–29, Q32, Q38, Q48–49, Q53–54, Q83, Q89, Q92
Domain 5	Security, Compliance, and Governance (14%)	14 Questions	Q8, Q17–18, Q24, Q39–41, Q43–44, Q51, Q62, Q65, Q68, Q78, Q93

Remember — You Don’t Need to Be Perfect to Pass!

The AWS AI Practitioner passing score is **700/1000**, which means you can **miss 15–20 questions** and still pass. Focus on concepts, AWS AI service selection, and responsible model deployment — trust your understanding and eliminate wrong answers.

Questions 1–10

Q1.
Which AWS service allows you to build, train, and deploy custom machine learning models at scale?

A) Amazon Lex
B) Amazon Bedrock
C) Amazon SageMaker
D) AWS Lambda

Q2.
Which of the following best describes unsupervised learning?

A) Predicting numeric values using labeled data
B) Identifying patterns or clusters in unlabeled data
C) Using rewards and penalties to train an agent
D) Applying labeled data to train a model

Q3.
A retail company wants to summarize customer reviews. Which AWS service is best suited for this task?

A) Amazon Rekognition
B) Amazon Comprehend
C) Amazon Translate
D) Amazon Polly

Q4.
What is one advantage of using Retrieval-Augmented Generation (RAG) with a foundation model?

A) Reduces latency by skipping the model altogether
B) Improves prediction speed by pruning the model
C) Provides up-to-date, context-specific answers using external data
D) Decreases API cost by caching prompt results

Q5.
Which type of model training is used to specialize a foundation model for a specific task or dataset?

A) Pre-training
B) Prompt chaining
C) Fine-tuning
D) Embedding

Q6.
Which AWS service provides access to foundation models like Claude, Jurassic, or Titan via API without managing infrastructure?

A) Amazon SageMaker
B) Amazon Kendra
C) Amazon Bedrock
D) AWS Lambda

Q7.
In the context of AI, what is “bias” most accurately defined as?

A) An overuse of GPU resources during training
B) A measure of model complexity
C) Systematic error that unfairly influences model outcomes
D) A type of deep learning optimization

Q8.
Which service would help you build a multilingual chatbot with speech recognition?

A) Amazon Translate + Amazon Comprehend
B) Amazon Polly + Amazon Rekognition
C) Amazon Lex + Amazon Transcribe
D) Amazon SageMaker + Amazon Textract

Q9.
Which AWS service helps generate speech from text, useful for voice-enabled applications?

A) Amazon Polly
B) Amazon Transcribe
C) Amazon Lex
D) Amazon Q

Q10.
What is a “token” in the context of large language models?

A) A multi-factor authentication method
B) A pre-processed image input
C) A segment of text used during model inference
D) An access credential for invoking APIs

Answers 1–10

A1.
Answer: C) Amazon SageMaker
Explanation: SageMaker is AWS’s end-to-end ML platform for building, training, and deploying custom models.

A2.
Answer: B) Identifying patterns or clusters in unlabeled data
Explanation: Unsupervised learning finds structure in data without labeled outputs.

A3.
Answer: B) Amazon Comprehend
Explanation: Comprehend can perform NLP tasks like sentiment analysis and summarization.

A4.
Answer: C) Provides up-to-date, context-specific answers using external data
Explanation: RAG enhances foundation models by retrieving relevant documents before generating responses.

A5.
Answer: C) Fine-tuning
Explanation: Fine-tuning adapts a pre-trained model to new tasks using labeled domain-specific data.

A6.
Answer: C) Amazon Bedrock
Explanation: Bedrock provides fully managed API access to multiple foundation models.

A7.
Answer: C) Systematic error that unfairly influences model outcomes
Explanation: Bias in AI typically refers to unintended, unfair impacts on decisions.

A8.
Answer: C) Amazon Lex + Amazon Transcribe
Explanation: Lex handles chatbot logic; Transcribe converts speech to text.

A9.
Answer: A) Amazon Polly
Explanation: Polly converts text to speech in lifelike voices.

A10.
Answer: C) A segment of text used during model inference
Explanation: Tokens are subunits of text (words, subwords) processed by LLMs.

Questions 11–20

Q11.
Which AWS service can help identify personal information like credit card numbers or names in a training dataset stored in Amazon S3?

A) AWS IAM
B) Amazon Macie
C) Amazon Textract
D) AWS CloudTrail

Q12.
What is the purpose of SageMaker Clarify?

A) Launching endpoint servers for foundation models
B) Evaluating model speed and GPU usage
C) Identifying bias and explaining model predictions
D) Serving prompts to Amazon Bedrock

Q13.
A developer needs a vector database to store text embeddings for semantic search. Which AWS service is most appropriate?

A) Amazon DynamoDB
B) Amazon DocumentDB
C) Amazon OpenSearch with k-NN
D) Amazon Polly

Q14.
Which of the following is a common risk of generative AI in production?

A) Deployment via HTTPS
B) High recall
C) Model hallucination
D) Insufficient GPU logging

Q15.
What is the primary role of Amazon Lex in AI-powered applications?

A) Sentiment analysis
B) Speech synthesis
C) Building conversational interfaces and chatbots
D) Object detection

Q16.
What type of prompting helps models reason more clearly by breaking tasks into logical steps?

A) Zero-shot prompting
B) Prompt chaining
C) Chain-of-thought prompting
D) Embedding prompting

Q17.
Which AWS service allows a user to catalog, govern, and manage AI-related datasets?

A) Amazon DataZone
B) Amazon Inspector
C) Amazon S3
D) AWS Artifact

Q18.
What is a *model card* in SageMaker used for?

A) Encrypting model input/output
B) Documenting model use, data, metrics, and ethical considerations
C) Launching auto-scaling endpoints
D) Selecting the correct foundation model from Amazon Bedrock

---

Q19.
Which AWS service helps with identifying and logging all API activity related to your AI infrastructure?

A) Amazon Macie
B) AWS Config
C) AWS CloudTrail
D) Amazon Inspector

Q20.
Which AWS tool helps ensure that only approved users and roles can access an Amazon Bedrock model endpoint?

A) AWS KMS
B) AWS IAM
C) Amazon Polly
D) Amazon Route 53

Answers 11–20

A11.
Answer: B) Amazon Macie
Explanation: Macie detects and classifies sensitive data in S3, such as PII.

A12.
Answer: C) Identifying bias and explaining model predictions
Explanation: SageMaker Clarify supports fairness and explainability through metrics and SHAP values.

A13.
Answer: C) Amazon OpenSearch with k-NN
Explanation: OpenSearch supports vector search for embeddings using k-Nearest Neighbors.

A14.
Answer: C) Model hallucination
Explanation: A common risk where the model outputs plausible but false or made-up content.

A15.
Answer: C) Building conversational interfaces and chatbots
Explanation: Lex provides voice/text-based interfaces powered by NLP.

A16.
Answer: C) Chain-of-thought prompting
Explanation: Helps improve reasoning by guiding the model step-by-step.

A17.
Answer: A) Amazon DataZone
Explanation: DataZone enables cataloging, governance, and sharing of data assets across teams.

A18.
Answer: B) Documenting model use, data, metrics, and ethical considerations
Explanation: Model cards improve transparency and compliance by summarizing key model details.

A19.
Answer: C) AWS CloudTrail
Explanation: Tracks all API actions across your AWS account, including those for AI services.

A20.
Answer: B) AWS IAM
Explanation: IAM controls access and permissions for Bedrock, SageMaker, and all AWS resources.

Questions 21–30

Q21.
A user needs to generate multiple image variations from a single text prompt. Which type of model would be best for this task?

A) Recurrent neural network
B) Diffusion model
C) Transformer model
D) Logistic regression

Q22.
Which AWS service helps convert handwritten forms into structured digital data?

A) Amazon Textract
B) Amazon Polly
C) Amazon Lex
D) AWS Glue

Q23.
In a generative AI application, which AWS service provides no-code tools to experiment with foundation models?

A) Amazon Q
B) Amazon Bedrock Studio
C) Amazon PartyRock
D) AWS Lambda

Q24.
Which of the following is a benefit of using AWS-managed foundation models instead of training your own?

A) Requires more infrastructure setup
B) Lower initial compute costs and faster time-to-value
C) Allows unlimited control of the model's internal layers
D) Offers higher hallucination risk due to limited access

Q25.
Which metric is best suited to evaluate the quality of machine translation output?
A) ROUGE

B) BLEU
C) AUC
D) F1 Score

Q26.
What is one major benefit of in-context learning over fine-tuning?

A) It updates the model's weights for new tasks
B) It costs less and doesn't require retraining
C) It guarantees higher accuracy
D) It allows GPU scaling

Q27.
A machine learning model performs well on training data but poorly on new data. What is the most likely issue?

A) Underfitting
B) Data drift
C) Overfitting
D) High bias

Q28.
What is the purpose of Amazon SageMaker Feature Store?

A) Manage IAM roles for training jobs
B) Store structured logs for model performance
C) Serve consistent features for training and inference
D) Replace Glue for visual data transformation

Q29.
What is a key feature of the AWS Shared Responsibility Model as it applies to AI workloads?

A) AWS is responsible for securing custom model code
B) Customers are responsible for physical data center access
C) AWS secures the infrastructure, customers secure configurations and data
D) Customers rely on AWS to guarantee model explainability

Q30.
Which AWS tool enables human reviewers to verify low-confidence model predictions before results are finalized?

A) Amazon Lex
B) Amazon A2I
C) AWS Macie
D) Amazon Q

Answers 21–30

A21.
Answer: B) Diffusion model
Explanation: Diffusion models generate high-quality images from noise guided by text prompts.

A22.
Answer: A) Amazon Textract
Explanation: Textract extracts text and structure (like tables, forms) from scanned documents.

A23.
Answer: C) Amazon PartyRock
Explanation: PartyRock allows users to experiment with Bedrock-hosted models via a no-code interface.

A24.
Answer: B) Lower initial compute costs and faster time-to-value
Explanation: Managed models save cost and time since training and infra are already done by AWS.

A25.
Answer: B) BLEU
Explanation: BLEU evaluates translation accuracy based on n-gram overlap with reference text.

A26.
Answer: B) It costs less and doesn't require retraining
Explanation: In-context learning uses well-structured prompts without changing the model itself.

A27.
Answer: C) Overfitting
Explanation: Overfitting happens when a model learns training data too well but fails to generalize.

A28.
Answer: C) Serve consistent features for training and inference
Explanation: Feature Store provides a centralized source for ML features used in both stages.

A29.
Answer: C) AWS secures the infrastructure, customers secure configurations and data
Explanation: The Shared Responsibility Model splits infra vs user responsibilities.

A30.
Answer: B) Amazon A2I
Explanation: A2I (Augmented AI) lets humans review ML outputs before they are accepted.

Questions 31–40

Q31.
Which AWS service would you use to translate chatbot messages from English to Spanish in real-time?

A) Amazon Comprehend
B) Amazon Polly
C) Amazon Translate
D) Amazon Lex

Q32.
What is the main advantage of using Amazon SageMaker JumpStart for a new AI project?

A) It provides serverless access to third-party models.
B) It allows full model interpretability by default.
C) It offers pre-built models and solutions for quick deployment.
D) It replaces the need for IAM roles in ML workflows.

Q33.
Which method is commonly used to measure the similarity between user input and stored content in a vector database?

A) Hash matching
B) k-Nearest Neighbors (k-NN)
C) Confusion matrix
D) ROC curve

Q34.
What is an *embedding* in the context of AI and vector databases?

A) A label used for supervised learning
B) A data format used for model deployment
C) A dense vector representing the meaning of data
D) A specific GPU setting for SageMaker endpoints

Q35.
A model always predicts the majority class but fails to detect rare cases. What metric is most likely misleading in this case?

A) Precision
B) F1 Score
C) Accuracy
D) Recall

Q36.
What is the purpose of using SHAP values in model evaluation?

A) To encrypt data before inference
B) To calculate how each feature impacts a model's output
C) To standardize tokenization of inputs
D) To classify model bias as legal or illegal

Q37.
Which AWS service supports storing AI model outputs securely with version control and access logging?

A) AWS Lambda
B) AWS Glue
C) Amazon S3
D) AWS CloudTrail

Q38.
What does Amazon Rekognition primarily analyze?

A) Document structure
B) Audio waveforms
C) Image and video content
D) Token-based language prompts

Q39.
How does Amazon Q for Business support enterprise AI use cases?

A) By running real-time image recognition
B) By providing GPU acceleration for model training

C) By answering natural language questions using internal company data
D) By launching distributed training jobs on SageMaker

Q40.
Which of the following defines *prompt injection* in generative AI applications?

A) Feeding the model with missing values
B) Tricking the model into executing unintended behavior via crafted input
C) Replacing GPU memory during inference
D) Automatically chaining model outputs to data pipelines

Answers 31–40

A31.
Answer: C) Amazon Translate
Explanation: Translate provides real-time, high-quality language translation.

A32.
Answer: C) It offers pre-built models and solutions for quick deployment
Explanation: JumpStart accelerates AI adoption by providing ready-to-use templates and models.

A33.
Answer: B) k-Nearest Neighbors (k-NN)
Explanation: Vector search compares embeddings using k-NN for similarity ranking.

A34.
Answer: C) A dense vector representing the meaning of data
Explanation: Embeddings convert text/images into vectors for comparison and semantic search.

A35.
Answer: C) Accuracy
Explanation: High accuracy can be misleading when rare classes are missed (i.e., imbalanced data).

A36.
Answer: B) To calculate how each feature impacts a model's output
Explanation: SHAP values explain model predictions by showing feature contribution.

A37.
Answer: C) Amazon S3

Explanation: S3 securely stores data, including models and logs, with versioning and access control.

A38.
Answer: C) Image and video content
Explanation: Rekognition detects labels, faces, and unsafe content in media.

A39.
Answer: C) By answering natural language questions using internal company data
Explanation: Amazon Q for Business is designed for internal Q&A across enterprise documents.

A40.
Answer: B) Tricking the model into executing unintended behavior via crafted input
Explanation: Prompt injection exploits LLMs by injecting commands that override the intended prompt.

Questions 41–50

Q41.
Which AWS service can automate data transformations and prep for machine learning with a visual interface?

A) AWS Glue DataBrew
B) Amazon S3
C) Amazon Q
D) Amazon Polly

Q42.
What is the purpose of a system prompt in a generative AI application?

A) To define the format of the output
B) To provide instructions that guide the model's role or behavior
C) To set GPU configuration
D) To compress the input tokens

Q43.
Which AWS service would help detect data drift or anomalies in live model input and predictions?

A) Amazon Rekognition
B) SageMaker Model Monitor
C) AWS Config
D) Amazon Kendra

Q44.
An organization wants to classify documents by language and topic. Which combination of services is best?

A) Amazon Polly and Amazon Textract
B) Amazon Translate and Amazon S3
C) Amazon Comprehend and Amazon S3
D) Amazon Bedrock and Amazon Q

Q45.
Which type of model evaluation metric is most suitable for summarization tasks?

A) ROUGE
B) BLEU
C) Perplexity
D) MAE

Q46.
How does Amazon Bedrock simplify access to foundation models?

A) It auto-trains any model uploaded by the user
B) It hosts open-source models in SageMaker endpoints
C) It provides API access to multiple foundation models without managing infrastructure
D) It requires a local container environment for use

Q47.
What is the purpose of a vector database in Retrieval-Augmented Generation (RAG)?

A) To reduce training time by compressing models
B) To generate GPU-friendly model files
C) To store embeddings and enable similarity search for grounding responses
D) To support IAM token storage

Q48.
Which AWS tool enables ongoing audit tracking and policy alignment for AI projects?

A) Amazon Q
B) AWS Artifact
C) AWS Audit Manager
D) Amazon Comprehend

Q49.
What is the main purpose of Amazon Titan in the AWS AI ecosystem?

A) It offers containerized GPU inference for local deployments
B) It's AWS's proprietary foundation model family, accessible via Bedrock
C) It provides OCR and handwriting recognition
D) It's an Amazon robotics platform for automation

Q50.
In what situation would using human-in-the-loop (via Amazon A2I) be most appropriate?

A) To store model parameters in S3
B) When model outputs must be verified for accuracy before release
C) When you want to increase token limits in Bedrock
D) To convert speech to text

Answers 41–50

A41.
Answer: A) AWS Glue DataBrew
Explanation: DataBrew lets users visually transform, clean, and prepare data for ML without code.

A42.
Answer: B) To provide instructions that guide the model's role or behavior
Explanation: System prompts establish tone, behavior, and rules for the AI's responses.

A43.
Answer: B) SageMaker Model Monitor
Explanation: Model Monitor tracks model inputs and outputs for drift, helping detect issues post-deployment.

A44.
Answer: C) Amazon Comprehend and Amazon S3
Explanation: Comprehend analyzes text for language and topic classification; S3 stores the documents.

A45.
Answer: A) ROUGE
Explanation: ROUGE compares overlaps between generated and reference summaries.

A46.
Answer: C) It provides API access to multiple foundation models without managing infrastructure
Explanation: Bedrock is a managed service that provides simple access to FMs from various providers.

A47.
Answer: C) To store embeddings and enable similarity search for grounding responses
Explanation: Vector databases hold embeddings to support retrieval in RAG workflows.

A48.
Answer: C) AWS Audit Manager
Explanation: Audit Manager helps you map and maintain compliance controls for audit readiness.

A49.
Answer: B) It's AWS's proprietary foundation model family, accessible via Bedrock
Explanation: Titan models support generative AI use cases and are natively integrated with Bedrock.

A50.
Answer: B) When model outputs must be verified for accuracy before release
Explanation: Amazon A2I allows for human intervention to validate or correct AI results.

Questions 51–60

Q51.
What is the benefit of using Amazon SageMaker Studio for AI development?

A) It replaces IAM permissions
B) It offers a browser-based IDE with integrated ML tools
C) It automatically generates prompts for Bedrock
D) It converts images to embeddings

Q52.
Which AWS service would best support training and deploying a machine learning model with custom Python code?

A) Amazon Lex
B) Amazon SageMaker
C) Amazon Translate
D) Amazon Polly

Q53.
A company wants to track changes in model performance over time. What AWS service helps with this?

A) Amazon Textract
B) SageMaker Model Monitor
C) Amazon Lex
D) Amazon Transcribe

Q54.
Which of the following AWS services provides an interface to manage and govern data shared across teams?

A) AWS IAM
B) Amazon Macie
C) Amazon DataZone
D) Amazon Translate

Q55.
What is the main benefit of using Amazon Comprehend's custom classification feature?

A) It allows training of vision models on image datasets
B) It enables model quantization for faster inference
C) It lets you categorize documents based on your own business labels
D) It encrypts output summaries for compliance

Q56.
Which of the following best defines "zero-shot prompting"?

A) The model is fine-tuned before inference
B) The model answers a task without seeing any examples
C) The model is prompted after receiving labeled training data
D) The prompt contains multiple few-shot examples

Q57.
What is a major advantage of using Amazon Q (Business) over a general LLM interface?

A) It supports GPU-based training workloads
B) It automatically clusters training data
C) It grounds answers in enterprise documents and internal knowledge
D) It supports image classification natively

Q58.
Which AWS service provides a knowledge base-backed AI chatbot using Bedrock models and RAG?

A) Amazon Lex
B) Amazon Rekognition
C) Amazon Q
D) Amazon S3

Q59.
A security team wants to ensure AI data in S3 is encrypted at rest. Which AWS service helps manage encryption keys?

A) Amazon A2I
B) AWS Artifact

C) AWS KMS
D) Amazon Bedrock

Q60.
What is the benefit of using Guardrails in Amazon Bedrock?

A) To detect and prevent drift in training pipelines
B) To convert structured data into image format
C) To restrict harmful, unsafe, or non-compliant model outputs
D) To perform live speech translation

Answers 51–60

A51.
Answer: B) It offers a browser-based IDE with integrated ML tools
Explanation: SageMaker Studio centralizes data science workflows with Jupyter notebooks and built-in tools.

A52.
Answer: B) Amazon SageMaker
Explanation: SageMaker supports bringing your own code and containers for ML model training and deployment.

A53.
Answer: B) SageMaker Model Monitor
Explanation: Tracks prediction quality and input distribution over time to identify drift or degradation.

A54.
Answer: C) Amazon DataZone
Explanation: DataZone supports cataloging, sharing, and governance of datasets across teams.

A55.
Answer: C) It lets you categorize documents based on your own business labels
Explanation: Comprehend custom classification lets you train a classifier on your own label taxonomy.

A56.
Answer: B) The model answers a task without seeing any examples
Explanation: Zero-shot prompting relies on pretrained capabilities and instructions alone.

A57.
Answer: C) It grounds answers in enterprise documents and internal knowledge
Explanation: Amazon Q uses RAG to answer based on your company's data.

A58.
Answer: C) Amazon Q
Explanation: Q integrates Bedrock models with internal data sources for grounded, enterprise chat.

A59.
Answer: C) AWS KMS
Explanation: Key Management Service handles encryption key creation, storage, and access.

A60.
Answer: C) To restrict harmful, unsafe, or non-compliant model outputs
Explanation: Bedrock Guardrails filter and control generative model responses based on safety rules.

Questions 61–70

Q61.
Which AWS service is specifically designed to help build, test, and deploy AI solutions using an end-to-end managed environment?

A) Amazon Textract
B) Amazon Q
C) Amazon SageMaker
D) AWS Lambda

Q62.
What is the main reason organizations use Retrieval-Augmented Generation (RAG) instead of relying only on a foundation model's internal knowledge?

A) It reduces GPU power requirements
B) It improves model explainability
C) It enables grounding in up-to-date or proprietary information
D) It increases BLEU score in translation tasks

Q63.
In a classification model, which metric is most useful when trying to reduce false negatives in rare event detection?

A) Accuracy
B) Recall
C) BLEU
D) MAE

Q64.
What does a model card typically include? *(Choose TWO.)*

A) GPU utilization metrics
B) Intended use cases
C) Dataset details and sources
D) Token cost per query

Q65.
Which AWS service helps prevent unauthorized access to AI endpoints by controlling permissions?

A) AWS IAM
B) Amazon Translate
C) Amazon Polly
D) Amazon CloudWatch

Q66.
What is one reason to use prompt engineering instead of fine-tuning?

A) It reduces latency
B) It allows reuse of the same model without additional training
C) It guarantees full explainability
D) It converts model outputs into tokens

Q67.
Which AWS service is best suited for real-time speech-to-text transcription?

A) Amazon Transcribe
B) Amazon Lex
C) Amazon Polly
D) Amazon Kendra

Q68.
A team wants to perform image analysis, including detecting objects and facial attributes. Which service should they use?

A) Amazon Bedrock
B) Amazon Comprehend
C) Amazon Rekognition
D) Amazon A2I

Q69.
What is the main risk of not monitoring model inputs after deployment?

A) GPU underuse
B) Increased training cost
C) Input drift leading to degraded model performance
D) Longer prompt tokenization

Q70.
What AWS feature supports safe handling of sensitive prompts and outputs in generative AI applications?

A) SageMaker Model Cards
B) Amazon Macie
C) Bedrock Guardrails
D) AWS CodeCommit

Answers 61–70

A61.
Answer: C) Amazon SageMaker
Explanation: SageMaker provides a complete ML platform for model building, training, and deployment.

A62.
Answer: C) It enables grounding in up-to-date or proprietary information
Explanation: RAG lets the model reference external documents for current or custom data.

A63.
Answer: B) Recall
Explanation: Recall focuses on capturing true positives — critical when false negatives are costly.

A64.
Answer: B) Intended use cases and C) Dataset details and sources
Explanation: Model cards document purpose, data origin, performance, and limitations.

A65.
Answer: A) AWS IAM
Explanation: IAM enforces fine-grained access control to AI resources like Bedrock endpoints.

A66.
Answer: B) It allows reuse of the same model without additional training
Explanation: Prompt engineering lets you tailor outputs with no retraining or cost overhead.

A67.
Answer: A) Amazon Transcribe
Explanation: Transcribe converts live or recorded speech into accurate text.

A68.
Answer: C) Amazon Rekognition
Explanation: Rekognition analyzes images and videos for faces, objects, and text.

A69.
Answer: C) Input drift leading to degraded model performance
Explanation: If input data shifts from what the model was trained on, accuracy may drop.

A70.
Answer: C) Bedrock Guardrails
Explanation: Guardrails help filter and manage prompts and outputs for compliance and safety.

Questions 71–80

Q71.
What AWS service enables access to pre-trained models from Anthropic, Stability AI, and AI21 Labs?

A) Amazon SageMaker
B) Amazon Bedrock
C) Amazon Comprehend
D) Amazon EC2

Q72.
What is a foundational concept of human-centered design in AI systems?

A) Reducing memory usage in training
B) Prioritizing GPU acceleration
C) Explaining outputs in a way users can understand
D) Minimizing vector database usage

Q73.
Which AWS service is primarily used to manage machine learning features and avoid duplication between training and inference?

A) AWS Glue
B) Amazon S3
C) SageMaker Feature Store
D) Amazon Lex

Q74.
An organization wants to control how long ML data is retained and ensure deletion after a specific period. Which feature helps?

A) S3 Bucket Lifecycle Policies
B) Amazon Lex intents
C) AWS Lambda
D) SageMaker model tuning

Q75.
What is the purpose of Amazon A2I in AI workflows?

A) Create multi-modal embeddings for large models
B) Allow human review of low-confidence or sensitive ML predictions
C) Accelerate foundation model training
D) Detect personally identifiable information in documents

Q76.
What does BLEU score evaluate in NLP tasks?

A) Output tone
B) Model inference time
C) Translation accuracy based on reference overlap
D) Ethical use of AI

Q77.
Which AWS service provides compliance documentation (e.g., SOC, ISO certifications)?

A) AWS CloudTrail
B) AWS Artifact
C) Amazon Rekognition
D) Amazon S3

Q78.
A financial company wants to trace which dataset was used to train a particular ML model. Which governance tool supports this need?

A) Amazon Polly
B) Amazon Translate
C) Model Cards
D) Amazon Kendra

Q79.
What is *latent space* in the context of foundation models?

A) A model's reserved memory area
B) A range of GPU utilization values

C) A compressed internal representation of features or concepts
D) A multi-user chat interface for AI responses

Q80.
Which AWS service would best support a chatbot that needs to understand spoken commands and reply with generated speech?

A) Amazon Textract
B) Amazon Q
C) Amazon Transcribe + Amazon Polly
D) Amazon Comprehend + Amazon S3

Answers 71–80

A71.
Answer: B) Amazon Bedrock
Explanation: Bedrock provides access to foundation models from multiple providers via API.

A72.
Answer: C) Explaining outputs in a way users can understand
Explanation: Human-centered design emphasizes clarity, trust, and user comprehension.

A73.
Answer: C) SageMaker Feature Store
Explanation: Stores and serves features consistently across ML pipelines.

A74.
Answer: A) S3 Bucket Lifecycle Policies
Explanation: These automate data archiving and deletion for governance and compliance.

A75.
Answer: B) Allow human review of low-confidence or sensitive ML predictions
Explanation: A2I adds human-in-the-loop validation to critical model decisions.

A76.
Answer: C) Translation accuracy based on reference overlap
Explanation: BLEU measures how closely machine translation matches a human reference.

A77.
Answer: B) AWS Artifact
Explanation: Provides downloadable compliance and security documents.

A78.
Answer: C) Model Cards
Explanation: Document model inputs, training data, use cases, and limitations.

A79.
Answer: C) A compressed internal representation of features or concepts
Explanation: Foundation models use latent space to understand and relate input data.

A80.
Answer: C) Amazon Transcribe + Amazon Polly
Explanation: Transcribe converts speech to text; Polly converts generated replies into audio.

Questions 81–90

Q81.
Which AWS service helps you define custom vocabulary and language models for speech recognition tasks?

A) Amazon Translate
B) Amazon Polly
C) Amazon Transcribe
D) Amazon Kendra

Q82.
What does the *Shared Responsibility Model* mean in the context of AWS AI services?

A) AWS and users co-own model weights
B) AWS handles physical security; users manage data, configuration, and access
C) AWS manages all aspects of AI, including ethics
D) Users only manage billing and logs

Q83.
Which tool would help a company detect bias in both datasets and model predictions?

A) Amazon Comprehend
B) SageMaker Clarify
C) AWS Glue
D) Amazon Bedrock

Q84.
Why might an organization use *continuous pretraining* with foundation models?

A) To delete unused tokens
B) To keep the model updated with recent data
C) To reset model weights
D) To avoid needing GPUs

Q85.
What is the function of *instruction tuning* in foundation models?

A) Adjusts latency parameters
B) Teaches the model to better follow user commands
C) Converts binary to text
D) Optimizes GPU memory

Q86.
Which AWS service supports chunking text input and adding grounding context for large prompts?

A) Amazon Lex
B) Amazon Comprehend
C) Amazon Bedrock
D) AWS Lambda

Q87.
Which term best describes the risk of a model memorizing and reproducing training data verbatim?

A) Overfitting
B) Latency
C) Prompt injection
D) Token duplication

Q88.
An AI assistant produces different responses each time for the same prompt. What parameter is likely causing this?

A) Accuracy
B) Recall
C) Temperature
D) Latency

Q89.
In responsible AI development, what is one purpose of using *model explainability* tools?

A) To reduce GPU usage
B) To measure token frequency

C) To help understand feature impact on model predictions
D) To optimize latency in real-time endpoints

Q90.
A company needs to prevent public access to its SageMaker inference endpoints. What should it use?

A) Public API Gateway
B) Amazon Macie
C) AWS PrivateLink
D) Amazon Polly

Answers 81–90

A81.
Answer: C) Amazon Transcribe
Explanation: Transcribe supports custom vocabularies and language models for tailored speech recognition.

A82.
Answer: B) AWS handles physical security; users manage data, configuration, and access
Explanation: The Shared Responsibility Model divides control between AWS (infra) and the customer (data, permissions, models).

A83.
Answer: B) SageMaker Clarify
Explanation: Clarify detects bias and provides SHAP-based explainability during and after training.

A84.
Answer: B) To keep the model updated with recent data
Explanation: Continuous pretraining extends a model's knowledge beyond its original training cutoff.

A85.
Answer: B) Teaches the model to better follow user commands
Explanation: Instruction tuning helps models learn to respond well to prompts phrased as instructions.

A86.
Answer: C) Amazon Bedrock

Explanation: Bedrock supports prompt engineering, chunking, and RAG for grounding context.

A87.
Answer: A) Overfitting
Explanation: Overfitting can cause models to memorize training data instead of generalizing.

A88.
Answer: C) Temperature
Explanation: Higher temperature adds randomness to model output, affecting determinism.

A89.
Answer: C) To help understand feature impact on model predictions
Explanation: Explainability helps interpret why the model made a certain prediction.

A90.
Answer: C) AWS PrivateLink
Explanation: PrivateLink allows private, VPC-internal access to SageMaker and Bedrock endpoints.

Questions 91–100

Q91.
What is the role of Amazon Kendra in a Retrieval-Augmented Generation (RAG) pipeline?

A) Train a new foundation model from scratch
B) Manage cloud cost for AI pipelines
C) Retrieve relevant documents or passages from a data source
D) Generate embeddings from images

Q92.
Which of the following is a typical component of a RAG system on AWS? *(Choose TWO.)*

A) Amazon Bedrock
B) Amazon Textract
C) Amazon OpenSearch
D) AWS CodePipeline

Q93.
What AWS service provides a no-code interface for creating apps powered by Bedrock foundation models?

A) Amazon Rekognition
B) Amazon Lex
C) Amazon Q
D) Amazon PartyRock

Q94.
What type of learning uses labeled input-output pairs to train a model?

A) Reinforcement learning
B) Supervised learning
C) Unsupervised learning
D) Federated learning

Q95.
Which AWS service allows users to transcribe streaming audio with speaker identification?

A) Amazon Polly
B) Amazon Transcribe
C) Amazon Comprehend
D) Amazon SageMaker

Q96.
Why would an AI team use AWS Config in an ML pipeline?

A) To build no-code AI assistants
B) To validate prompts for Bedrock
C) To monitor compliance of infrastructure configurations
D) To perform model hyperparameter tuning

Q97.
Which type of bias occurs when the training dataset does not represent the target population?

A) Algorithmic bias
B) Data drift
C) Sampling bias
D) Tuning bias

Q98.
What is one limitation of few-shot prompting compared to fine-tuning?

A) It requires GPU access
B) It needs a full labeled dataset
C) It has limited task-specific learning retention
D) It supports only classification models

Q99.
A company needs to host a multilingual question-answering chatbot using proprietary documentation. Which combination is best?

A) Amazon Translate + Amazon S3
B) Amazon Kendra + Amazon Bedrock

C) Amazon Polly + AWS Lambda
D) AWS Glue + Amazon SageMaker

Q100.
Which AWS service allows teams to detect and redact sensitive data from documents or text inputs before model training?

A) Amazon Bedrock
B) Amazon Rekognition
C) Amazon Macie
D) Amazon Comprehend

Answers 91–100

A91.
Answer: C) Retrieve relevant documents or passages from a data source
Explanation: Kendra is used to search and retrieve documents based on semantic or keyword relevance.

A92.
Answer: A) Amazon Bedrock and C) Amazon OpenSearch
Explanation: Bedrock provides the generative model; OpenSearch stores embeddings for vector search in RAG.

A93.
Answer: D) Amazon PartyRock
Explanation: PartyRock is a no-code app builder that uses Bedrock models to build generative AI apps.

A94.
Answer: B) Supervised learning
Explanation: Supervised learning uses labeled input-output pairs to learn mappings.

A95.
Answer: B) Amazon Transcribe
Explanation: Transcribe can transcribe speech with features like speaker identification and vocabulary filtering.

A96.
Answer: C) To monitor compliance of infrastructure configurations
Explanation: AWS Config tracks configuration changes and compliance of AWS resources.

A97.
Answer: C) Sampling bias
Explanation: Sampling bias arises when the training data does not represent the broader population.

A98.
Answer: C) It has limited task-specific learning retention
Explanation: Few-shot prompting is flexible but doesn't permanently adapt the model like fine-tuning.

A99.
Answer: B) Amazon Kendra + Amazon Bedrock
Explanation: Kendra retrieves relevant data, Bedrock generates grounded responses using that context.

A100.
Answer: C) Amazon Macie
Explanation: Macie scans for sensitive data (like PII) and helps with redaction and classification.