

# Simulacro de Examen: AWS Certified AI Practitioner (AIF-C01)

Dominio 4: IA Responsable (150 Preguntas)

Material de Preparación Profesional

## Instrucciones

Este cuestionario contiene 150 preguntas de opción múltiple. Selecciona la respuesta correcta para cada una. La clave de respuestas se encuentra al final del documento.

## Preguntas

1. ¿Cuál es la definición de IA Responsable según AWS?
  - a) Un conjunto de algoritmos que no requieren supervisión humana.
  - b) Directrices y principios para garantizar que los sistemas de IA funcionen de manera segura, fiable y ética.
  - c) El proceso de comprar hardware de bajo consumo.
  - d) Un tipo de licencia de software de código abierto.
2. ¿Qué dimensión de la IA responsable busca que los modelos traten a todos de forma imparcial?
  - a) Solidez.
  - b) Equidad (Fairness).
  - c) Gobernanza.
  - d) Privacidad.
3. La capacidad de explicar en términos humanos por qué un modelo rechazó un préstamo se conoce como:
  - a) Transparencia.
  - b) Explicabilidad.
  - c) Robustez.
  - d) Inclusividad.
4. ¿Qué garantiza la "Solidez" (Robustness) en un sistema de IA?
  - a) Que el sistema sea tolerante a errores y funcione de forma consistente.

- b) Que el sistema sea de código abierto.
  - c) Que el sistema use cifrado de 256 bits.
  - d) Que el sistema no tenga latencia.
5. El objetivo principal de la privacidad y seguridad en IA responsable es:
- a) Reducir el coste de los tokens.
  - b) Proteger los datos del usuario y no exponer información de identificación personal (PII).
  - c) Compartir los datos con otras empresas.
  - d) Borrar los modelos cada semana.
6. ¿En qué consiste la "Gobernanza" en el contexto de la IA?
- a) En elegir al director de la empresa.
  - b) En cumplir y auditar la conformidad con los estándares y prácticas recomendadas del sector.
  - c) En programar en Python exclusivamente.
  - d) En usar solo servicios gratuitos de AWS.
7. ¿Qué práctica de transparencia es esencial para los usuarios finales?
- a) Mostrar el código fuente del modelo.
  - b) Asegurarse de que los usuarios sepan cuándo están interactuando con una IA.
  - c) Publicar el salario de los científicos de datos.
  - d) No dar ninguna información para evitar ataques.
8. ¿Cómo se mide la imparcialidad de un modelo de IA?
- a) Por su velocidad de respuesta.
  - b) Por el sesgo y la varianza de los resultados entre diferentes grupos demográficos.
  - c) Por el número de capas de la red neuronal.
  - d) Por el tamaño del archivo .json de salida.
9. ¿Qué ocurre si un modelo es más exacto para hombres que para mujeres debido a los datos de entrenamiento?
- a) El modelo es perfecto.
  - b) El modelo muestra un sesgo (Bias) por desequilibrio de clases.
  - c) El modelo necesita más RAM.
  - d) El modelo es multimodal.
10. Un conjunto de datos de entrenamiento donde un grupo demográfico tiene pocas muestras comparado con otro causa:
- a) Inferencia asincrónica.

- b) Desequilibrio de clases.
  - c) Alta temperatura.
  - d) Cifrado de datos.
11. ¿Qué riesgo legal importante surge de los resultados sesgados en la contratación?
- a) Pérdida de contraseñas.
  - b) Trato discriminatorio que puede llevar a demandas judiciales.
  - c) Aumento del consumo eléctrico.
  - d) El modelo se vuelve más lento.
12. ¿Qué es una "alucinación" en IA Generativa?
- a) Un error de red en AWS.
  - b) Cuando el modelo inventa información que parece real pero es falsa.
  - c) Una técnica para mejorar imágenes.
  - d) El nombre de un nuevo modelo de Amazon.
13. El caso de los abogados sancionados en Nueva York en 2023 se debió a:
- a) No pagar las tasas del tribunal.
  - b) Presentar citas legales falsas generadas por una IA (alucinaciones).
  - c) Usar software pirata.
  - d) No tener conexión a internet.
14. ¿Por qué las obras generadas íntegramente por IA no suelen tener protección de derechos de autor?
- a) Porque son gratis.
  - b) Porque legalmente no son obra de un ser humano.
  - c) Porque AWS posee los derechos.
  - d) Porque se borran automáticamente.
15. ¿Qué servicio de AWS ayuda a detectar sesgos durante la preparación de datos y después del entrenamiento?
- a) Amazon Polly.
  - b) Amazon SageMaker Clarify.
  - c) AWS Glue.
  - d) Amazon S3.
16. ¿Cómo trata SageMaker Clarify a los modelos para explicar sus predicciones?
- a) Analizando cada línea de código.
  - b) Como una "caja negra", observando entradas y salidas para determinar la importancia de las características.

- c) Como un sistema de reglas fijas.
  - d) Como una base de datos SQL.
17. ¿Qué técnica de seguridad en Bedrock filtra contenido de odio o violencia antes de que llegue al usuario?
- a) AWS Shield.
  - b) Barreras de protección (Guardrails).
  - c) Amazon Titan Lite.
  - d) Amazon VPC.
18. Si una petición de usuario es bloqueada por una barrera de protección:
- a) El modelo se apaga.
  - b) El usuario recibe una respuesta de infracción configurada y la petición no llega al modelo.
  - c) Se cobra el doble por el token.
  - d) El sistema se reinicia.
19. ¿Qué métrica de Clarify indica si una clase tiene una proporción injusta de rechazos frente a aceptaciones?
- a) Latencia.
  - b) Disparidad demográfica.
  - c) Precisión media.
  - d) RMSE.
20. La "Diferencia de Exhaustividad" (Recall Difference) entre grupos mide:
- a) El tiempo de entrenamiento.
  - b) Si el modelo predice correctamente los positivos con la misma frecuencia para todas las clases.
  - c) El número de palabras en el prompt.
  - d) El coste de la GPU.
21. ¿Qué es el "Model Disgorgement"?
- a) Una técnica para hacer el modelo más grande.
  - b) El proceso de eliminar los efectos de datos específicos o protegidos de un modelo entrenado.
  - c) Un ataque de denegación de servicio.
  - d) La traducción de modelos al español.
22. ¿Qué técnica de entrenamiento asegura que un dato individual no afecte significativamente la salida, protegiendo la privacidad?
- a) Privacidad Diferencial.

- b) RAG.
  - c) Fine-tuning masivo.
  - d) Tokenización.
23. ¿Cómo funciona el "Watermarking" (Marcas de agua) en LLMs?
- a) Poniendo un logo transparente en el texto.
  - b) Seleccionando palabras de una lista "verde" secreta para crear una prueba estadística de generación por IA.
  - c) Guardando el nombre del usuario en el prompt.
  - d) Usando tinta invisible.
24. ¿Qué característica de los datasets responsables garantiza que se representen poblaciones diversas?
- a) Velocidad.
  - b) Inclusividad.
  - c) Compresión.
  - d) Cifrado.
25. ¿Por qué es importante la "Sostenibilidad" al seleccionar modelos de IA?
- a) Para que el modelo dure 100 años.
  - b) Para minimizar el impacto ambiental y el consumo de energía del entrenamiento masivo.
  - c) Para que el modelo hable de ecología.
  - d) No es un factor relevante.
26. ¿Qué principio sugiere reutilizar modelos preentrenados en lugar de entrenar desde cero para ser más responsable?
- a) Transparencia.
  - b) Sostenibilidad.
  - c) Seguridad.
  - d) Diversidad.
27. ¿Qué ocurre si borras los datos originales de un FM después de que ya ha sido entrenado con ellos?
- a) El modelo olvida los datos instantáneamente.
  - b) El modelo no puede "olvidar" los datos ya incorporados en sus ponderaciones.
  - c) El modelo deja de funcionar.
  - d) AWS devuelve el dinero del entrenamiento.
28. ¿Qué servicio permite realizar una "Evaluación Humana" de modelos en la consola de Amazon Bedrock?

- a) Amazon Bedrock Model Evaluation (Módulo de evaluación).
  - b) AWS Glue DataBrew.
  - c) Amazon Rekognition.
  - d) Amazon SageMaker Canvas.
29. Los <sup>Est</sup>ereotipos de peticiones.<sup>en</sup> SageMaker Clarify miden:
- a) El número de letras mayúsculas.
  - b) La probabilidad de que el modelo incluya sesgos por raza, género o religión en su respuesta.
  - c) La velocidad de la conexión.
  - d) El formato de la fecha.
30. ¿Qué dimensión evalúa si el resultado del modelo cambia debido a errores tipográficos simples?
- a) Toxicidad.
  - b) Solidez semántica.
  - c) Conocimiento fáctico.
  - d) Exactitud.
31. La transparencia de un modelo mide:
- a) El brillo de la pantalla.
  - b) El grado en que las partes interesadas pueden entender cómo funciona y por qué produce sus resultados.
  - c) Si el modelo es gratuito o de pago.
  - d) El número de usuarios que pueden verlo.
32. ¿Cuál de estos modelos es considerado altamente <sup>Interpretable</sup>?
- a) Redes Neuronales Profundas.
  - b) Regresión Lineal.
  - c) Transformadores.
  - d) Modelos de Difusión.
33. ¿Qué caracteriza a un modelo que es una <sup>Caja Negra</sup>?
- a) Que no tiene interfaz gráfica.
  - b) Que sus mecanismos internos son tan complejos que no se pueden entender matemáticamente de forma sencilla.
  - c) Que está apagado.
  - d) Que solo acepta datos encriptados.
34. ¿Qué permite la <sup>Explicabilidad</sup> sin conocer los mecanismos internos exactos?

- a) Describir lo que hace el modelo observando la relación entre entradas y salidas.
  - b) Borrar el modelo.
  - c) Aumentar el número de GPUs.
  - d) Ver los pesos de cada neurona.
35. ¿Cuál es el principal "menoscabo" (trade-off) de mejorar la transparencia de un modelo?
- a) Aumenta el coste.
  - b) Suele implicar un menor rendimiento o capacidad en tareas complejas.
  - c) El modelo se vuelve más pesado.
  - d) No hay menoscabo.
36. Desde el punto de vista de la seguridad, ¿cuál es el riesgo de un modelo totalmente transparente?
- a) Que sea muy lento.
  - b) Es más susceptible a ataques ya que los hackers conocen sus vulnerabilidades internas.
  - c) Que nadie quiera usarlo.
  - d) Que consuma mucha RAM.
37. ¿Qué documentación de AWS detalla el diseño responsable y las limitaciones de sus APIs de IA?
- a) AWS Bills.
  - b) Tarjetas de servicio de IA (AI Service Cards).
  - c) Amazon Prime Video.
  - d) AWS CloudTrail.
38. ¿Para qué servicio de Bedrock existe una Tarjeta de Servicio de IA oficial?
- a) Solo para modelos de Meta.
  - b) Para Amazon Titan Text.
  - c) Para todos los modelos de internet.
  - d) No existen para Bedrock.
39. ¿Qué herramienta de SageMaker genera automáticamente documentación sobre el ciclo de vida, datasets y entrenamiento de un modelo?
- a) Fichas del modelo de SageMaker (Model Cards).
  - b) Amazon S3.
  - c) AWS Glue.
  - d) Amazon Lex.
40. Los "Valores de Shapley" en SageMaker Clarify se utilizan para:

- a) Calcular el precio de la factura.
  - b) Determinar la contribución (importancia) de cada característica a las predicciones del modelo.
  - c) Medir la temperatura de la sala.
  - d) Contar el número de tokens.
41. ¿Qué muestra un "Gráfico de dependencia parcial"?
- a) El número de servidores fallidos.
  - b) Cómo cambian las predicciones de un modelo según varían los valores de una característica (ej. edad).
  - c) El tiempo de descarga del dataset.
  - d) La lista de usuarios conectados.
42. ¿Qué es la IA centrada en el ser humano?
- a) IA que no necesita electricidad.
  - b) Diseño de sistemas que priorizan las necesidades y valores humanos, mejorando sus capacidades en lugar de reemplazarlos.
  - c) IA que solo habla con humanos.
  - d) Un robot físico de Amazon.
43. ¿Qué servicio de AWS permite enviar predicciones de baja confianza a revisores humanos?
- a) Amazon Polly.
  - b) Amazon Augmented AI (A2I).
  - c) Amazon Transcribe.
  - d) AWS Lambda.
44. En Amazon A2I, ¿quiénes pueden formar el grupo de revisores?
- a) Solo empleados de Amazon.
  - b) Equipo propio, proveedores externos o Amazon Mechanical Turk.
  - c) Solo robots.
  - d) El público general de Twitter.
45. ¿Qué significa RLHF?
- a) Real Life High Frequency.
  - b) Aprendizaje por refuerzo a partir de la retroalimentación humana.
  - c) Remote Logging Human Feedback.
  - d) Redes Lógicas de Humanos Fieles.
46. ¿Cuál es el objetivo principal de aplicar RLHF a un LLM?



- a) Que sea más rápido.
  - b) Garantizar que produzca contenido veraz, inofensivo y útil.
  - c) Reducir el número de parámetros.
  - d) Eliminar la necesidad de GPUs.
47. En el proceso de RLHF, ¿qué es el "Modelo de Recompensa"?
- a) Un modelo que da premios físicos.
  - b) Un modelo entrenado para predecir qué respuesta preferiría un humano.
  - c) Un modelo de pago por uso.
  - d) Un error del sistema.
48. ¿Qué herramienta facilita la recopilación de preferencias humanas para RLHF mediante la clasificación de respuestas?
- a) SageMaker Ground Truth.
  - b) Amazon RDS.
  - c) Amazon Bedrock Playgrounds.
  - d) AWS CloudWatch.
49. ¿Qué valor humano asegura que el modelo no genere contenido discriminatorio o ilegal?
- a) Amabilidad.
  - b) Inocencia/Seguridad (Harmlessness).
  - c) Velocidad.
  - d) Complejidad.
50. La colaboración interdisciplinaria con éticos y psicólogos es propia de:
- a) MLOps nivel 0.
  - b) IA centrada en el ser humano.
  - c) Regresión lineal simple.
  - d) Cómputo periférico.
51. ¿Qué riesgo de privacidad ocurre cuando un LLM revela PII presente en su entrenamiento?
- a) Inyección de prompts.
  - b) Filtración de datos (Data leakage).
  - c) Alucinación creativa.
  - d) Poda de parámetros.
52. ¿Cuál es una característica de la "Inocencia" en el contexto de RLHF?
- a) El modelo no sabe nada.
  - b) El modelo evita producir respuestas perjudiciales u ofensivas.

- c) El modelo es gratuito.
  - d) El modelo solo tiene un parámetro.
53. ¿Qué técnica busca que el modelo admita cuando no sabe algo?
- a) Honestidad.
  - b) Amabilidad.
  - c) Robustez.
  - d) Inocencia.
54. ¿Qué es el enfoque "Human-on-the-loop"?
- a) El humano decide cada acción.
  - b) El humano supervisa el sistema y puede intervenir si algo falla.
  - c) El humano no tiene poder de decisión.
  - d) Un error de programación.
55. ¿Cuál es la diferencia entre "Human-on-the-loop" y "Human-over-the-loop"?
- a) No hay diferencia.
  - b) "on-the-loop" es supervisión activa; "over-the-loop" es autoridad final sobre el sistema general.
  - c) "over-the-loop" es para niños.
  - d) Son marcas de hardware.
56. ¿Qué es la "Toxicidad Subjetiva"?
- a) Un virus en el servidor.
  - b) Contenido que puede ser ofensivo dependiendo del contexto cultural o personal.
  - c) Un error de sintaxis en Python.
  - d) El humo de los centros de datos.
57. ¿Qué técnica de Kearns ayuda a desalentar el "engaño" mediante IA en textos?
- a) Borrar internet.
  - b) Watermarking (marcado estadístico de tokens).
  - c) Cobrar por cada palabra.
  - d) Usar solo modelos viejos.
58. El término "Inglés es el nuevo lenguaje de programación" se refiere a:
- a) Que ya no habrá programadores.
  - b) La capacidad de los LLMs para generar código a partir de instrucciones en lenguaje natural.
  - c) Que AWS solo funciona en inglés.
  - d) Un error de traducción.

59. ¿Qué práctica organizacional recomienda Godden para evitar sesgos ciegos?
- a) Contratar solo ingenieros.
  - b) Crear equipos multidisciplinarios y diversos.
  - c) Usar un solo modelo para todo.
  - d) No evaluar los modelos.
60. ¿Por qué AWS recomienda definir Casos de uso específicos.<sup>en</sup> lugar de apps catch-all?
- a) Para gastar más dinero.
  - b) Para facilitar la transparencia, rendición de cuentas y mitigación de riesgos.
  - c) Porque los modelos no saben hacer muchas cosas.
  - d) Por limitaciones de espacio en S3.
61. ¿Qué métrica de Clarify compara la relación entre falsos negativos y falsos positivos entre grupos?
- a) Exactitud.
  - b) Igualdad de trato.
  - c) Latencia.
  - d) RMSE.
62. La Rendición de cuentas” (Accountability) significa:
- a) Contar el número de servidores.
  - b) Establecer líneas claras de responsabilidad para las decisiones del modelo.
  - c) Pagar las facturas a tiempo.
  - d) No cometer nunca errores.
63. ¿Qué servicio de AWS tiene una tarjeta de servicio de IA para la ”Detección de PII?
- a) Amazon Comprehend.
  - b) Amazon Polly.
  - c) Amazon Bedrock.
  - d) Amazon RDS.
64. ¿Cuál es el beneficio de la colaboración interdisciplinar en IA?
- a) Hacer más reuniones.
  - b) Captar perspectivas diversas para asegurar que la IA sea justa y ética.
  - c) Aumentar el coste del proyecto.
  - d) No tiene beneficios técnicos.
65. ¿Qué permite el .Análisis local.<sup>en</sup> explicabilidad?
- a) Ver el comportamiento de todo el modelo.

- b) Entender por qué se tomó una decisión sobre un único punto de datos específico.
  - c) Traducir el modelo a un idioma local.
  - d) Borrar los datos locales.
66. ¿Qué técnica permite a un modelo "auto-corregirse" criticando su propia salida?
- a) RAG.
  - b) Auto-refinamiento (Self-refinement).
  - c) Poda.
  - d) Inferencia por lotes.
67. ¿Qué dimensión de IA responsable asegura que los datos se usen con permiso?
- a) Robustez.
  - b) Consentimiento y Transparencia.
  - c) Velocidad.
  - d) Latencia.
68. ¿Por qué el sobreajuste es un problema de IA responsable?
- a) Porque gasta mucha CPU.
  - b) Porque el modelo puede funcionar mal para grupos no representados en el entrenamiento.
  - c) Porque el modelo se vuelve demasiado pequeño.
  - d) Porque el modelo deja de hablar español.
69. ¿Qué es un "Falso Rechazo" injusto?
- a) Cuando el servidor no responde.
  - b) Cuando el modelo predice default para un cliente solvente de un grupo minoritario.
  - c) Un error al borrar datos.
  - d) Cuando el usuario cancela la petición.
70. El uso de "Disclaimers" (avisos de descargo) en LLMs sirve para:
- a) Hacer el texto más largo.
  - b) Educar al usuario sobre las limitaciones y posibles errores del modelo.
  - c) Ocultar los fallos del sistema.
  - d) Cumplir con el formato PDF.
71. ¿Qué técnica de Kearns divide los datos de entrenamiento para poder eliminar el impacto de uno de ellos fácilmente?
- a) Sharding (Fragmentación).
  - b) RAG.

- c) Prompt Engineering.
  - d) Clustering.
72. ¿Qué servicio de AWS ayuda a auditar predicciones aleatorias como control de calidad?
- a) Amazon A2I.
  - b) Amazon Polly.
  - c) Amazon S3.
  - d) AWS KMS.
73. ¿Qué característica de SageMaker Clarify permite ver cómo influye la <sup>Edad</sup>.<sup>en</sup> la predicción final?
- a) Análisis de sentimientos.
  - b) Gráfico de dependencia parcial.
  - c) Conteo de palabras.
  - d) Cifrado de disco.
74. ¿Cuál es el riesgo de la Ingeniería Inversa.<sup>en</sup> modelos transparentes?
- a) Que el modelo se borre.
  - b) Que competidores o atacantes roben algoritmos propietarios basándose en las explicaciones.
  - c) Que el modelo hable al revés.
  - d) No hay riesgo.
75. El "Vaciado de modelo" (Model Disgorgement) es necesario cuando:
- a) El modelo es muy viejo.
  - b) Se descubre que el modelo usó datos protegidos o ilícitos en su entrenamiento.
  - c) El modelo tiene poca precisión.
  - d) Queremos cambiar de región en AWS.
76. ¿Qué técnica penaliza tokens repetidos para mejorar la calidad del texto?
- a) Temperatura.
  - b) Penalizaciones de presencia y frecuencia.
  - c) Stop sequence.
  - d) Cifrado SSL.
77. ¿Qué significa que un modelo tenga un "Sesgo de entrenamiento"?
- a) Que entrena muy lento.
  - b) Que los datos usados no representan fielmente la diversidad del mundo real.
  - c) Que el modelo usa mucha energía.
  - d) Que el modelo es secreto.

78. ¿Cuál es el beneficio de los proyectos de Código Abierto para la transparencia?
- a) Son siempre gratis.
  - b) El funcionamiento interno está abierto al escrutinio y auditoría de cualquier persona.
  - c) No requieren servidores de AWS.
  - d) Son más seguros por ser secretos.
79. ¿Qué servicio de AWS se usa para Identificar PII.<sup>en</sup> texto de forma responsable?
- a) Amazon Rekognition.
  - b) Amazon Comprehend.
  - c) Amazon RDS.
  - d) AWS Shield.
80. ¿Qué valor de RLHF asegura que el modelo sea educado con el usuario?
- a) Amabilidad (Helpfulness).
  - b) Honestidad.
  - c) Inocencia.
  - d) Latencia.
81. ¿Qué es la "Varianza."<sup>en</sup> el contexto de equidad?
- a) El precio del dólar.
  - b) Diferencias en el rendimiento del modelo a través de diferentes subconjuntos de datos.
  - c) Un error de red.
  - d) El número de capas neuronales.
82. ¿Qué técnica de Kearns ayuda a que un modelo no memorice datos específicos del entrenamiento?
- a) Privacidad Diferencial.
  - b) Aumento de datos.
  - c) Inferencia por lotes.
  - d) Poda.
83. ¿Cuál es el peligro de la "Toxicidad sutil?
- a) Que el modelo se apague.
  - b) Que pase los filtros automáticos y cause daño emocional o prejuicios de forma indirecta.
  - c) Que el servidor se caliente.
  - d) Ninguno.

84. ¿Qué describe mejor la "Honestidad" en un LLM?
- a) Que el modelo no diga palabras malsonantes.
  - b) Que el modelo proporcione información veraz y reconozca sus limitaciones de conocimiento.
  - c) Que el modelo no cobre de más.
  - d) Que el modelo sea de Amazon.
85. ¿Qué herramienta de AWS permite "Etiquetar datos" de forma masiva para entrenamiento responsable?
- a) Amazon SageMaker Ground Truth.
  - b) AWS Glue.
  - c) Amazon RDS.
  - d) AWS CloudFormation.
86. ¿Qué es un "Dataset Equilibrado"?
- a) Un dataset que pesa 1GB exacto.
  - b) Aquel que tiene una representación proporcional y justa de todas las clases y grupos.
  - c) Un dataset guardado en dos regiones.
  - d) Un dataset que solo tiene números.
87. La "Privacidad de los datos" en IA Generativa es un riesgo porque:
- a) El modelo puede memorizar y revelar secretos introducidos en los prompts.
  - b) El modelo se puede borrar solo.
  - c) El modelo gasta mucho dinero.
  - d) El modelo es demasiado inteligente.
88. ¿Qué técnica permite a un médico revisar el diagnóstico de una IA antes de dárselo al paciente?
- a) Inferencia serverless.
  - b) Human-in-the-loop (A2I).
  - c) Poda de árboles.
  - d) Cifrado KMS.
89. ¿Cuál es la principal ventaja de usar "Equipos Diversos" en el desarrollo de IA?
- a) Reducir el salario medio.
  - b) Identificar sesgos y problemas éticos que un equipo homogéneo podría pasar por alto.
  - c) Programar más rápido.
  - d) No hay ventaja técnica.

90. ¿Qué describe la "Inocencia" (Harmlessness) en un chatbot?
- a) Que el chatbot no sabe nada.
  - b) Que el chatbot evita dar instrucciones para fabricar armas o realizar actos ilegales.
  - c) Que el chatbot es gratis.
  - d) Que el chatbot es nuevo.
91. ¿Qué mide la "Especificidad" en un análisis de sesgo?
- a) La frecuencia con la que el modelo predice correctamente un resultado negativo.
  - b) El número de palabras del prompt.
  - c) El color de la salida.
  - d) El tiempo de carga.
92. ¿Qué es un "Dataset Inclusivo"?
- a) Uno que incluye todas las fotos de internet.
  - b) Uno que representa activamente poblaciones y experiencias diversas para evitar la exclusión.
  - c) Uno que está en formato .zip.
  - d) Uno que solo tiene datos de AWS.
93. ¿Cuál es un objetivo de la "Participación de las partes interesadas" (Stakeholders)?
- a) Ganar más dinero.
  - b) Incluir diversas perspectivas en la selección y despliegue del modelo para asegurar su utilidad social.
  - c) Que todos programen.
  - d) Vender más licencias.
94. ¿Qué métrica de Clarify analiza si el modelo predice resultados positivos de forma diferente para cada clase?
- a) Diferencia en proporciones positivas.
  - b) Latencia de red.
  - c) RMSE.
  - d) Conteo de parámetros.
95. ¿Cuál es el riesgo de la "Manipulación de datos" por IA?
- a) Que el ordenador se rompa.
  - b) Ataques que alteran la percepción de la realidad o influyen en decisiones de forma maliciosa.
  - c) Borrar el disco duro de AWS.
  - d) Que el modelo se vuelva lento.



96. ¿Qué técnica de Kearns .etiqueta.el texto como IA mediante probabilidad?
- a) Watermarking.
  - b) RAG.
  - c) Prompting.
  - d) Cifrado.
97. ¿Qué es el "Mimetismo de estilo".en GenAI?
- a) El modelo se viste como un humano.
  - b) La capacidad de la IA para copiar el estilo de un artista o escritor específico, planteando dudas de IP.
  - c) Un error gráfico.
  - d) El nombre de una red neuronal.
98. ¿Qué valor humano prioriza que la respuesta sea relevante para la pregunta del usuario?
- a) Amabilidad (Helpfulness).
  - b) Honestidad.
  - c) Inocencia.
  - d) Robustez.
99. ¿Qué describe el nivel 2 de MLOps en términos de responsabilidad?
- a) Todo es manual.
  - b) Automatización total que permite auditoría y trazabilidad completa de cada versión del modelo.
  - c) No usar seguridad.
  - d) Solo para apps de juegos.
100. ¿Qué componente de AWS ayuda a cumplir con la "Sostenibilidad.ºoptimizando el hardware?
- a) Instancias Graviton3.
  - b) Amazon S3.
  - c) AWS Glue.
  - d) Amazon RDS.
101. ¿Qué es la "Fidelidad".en la evaluación de Bedrock?
- a) El número de usuarios que repiten.
  - b) Si la respuesta se basa fielmente en los datos proporcionados (evita alucinaciones).
  - c) La calidad del audio.
  - d) El color de las imágenes.
102. ¿Por qué es importante la .Auditariedad.en sistemas bancarios de IA?

- a) Para gastar menos papel.
  - b) Para poder revisar y justificar legalmente cada decisión de crédito tomada por el modelo.
  - c) Para que el banco sea famoso.
  - d) Por estética.
103. ¿Qué es un "Logit" en términos de probabilidad de IA?
- a) Un registro de errores.
  - b) Puntuación cruda antes de convertirse en probabilidad para elegir una palabra.
  - c) El nombre de la base de datos.
  - d) Una marca de servidores.
104. ¿Qué describe el "Holistic Evaluation of Language Models" (HELM)?
- a) Un manual de yoga.
  - b) Un punto de referencia que evalúa múltiples dimensiones (exactitud, sesgo, toxicidad) simultáneamente.
  - c) Un servicio de seguridad perimetral.
  - d) El nombre de una red neuronal.
105. ¿Qué técnica de "Human-centered design" ayuda a que la IA sea fácil de usar?
- a) Participación de usuarios en el proceso de desarrollo.
  - b) No dejar que los humanos vean la IA.
  - c) Usar solo comandos de terminal.
  - d) Cobrar por cada uso.
106. ¿Cuál es el beneficio de la "Diversidad de orígenes de datos"?
- a) Hacer el archivo más grande.
  - b) Asegurar calidad e integridad evitando sesgos de una única fuente.
  - c) Reducir el tiempo de carga.
  - d) No tiene beneficios.
107. ¿Qué métrica penaliza a los modelos que son groseros o agresivos?
- a) Toxicidad (Toxicity).
  - b) Latencia.
  - c) Accuracy.
  - d) Recall.
108. ¿Qué es la "Inversión de Modelo"?
- a) Girar el servidor 180 grados.

- b) Un ataque para extraer datos privados del entrenamiento a partir de las respuestas del modelo.
  - c) Cambiar el orden de las palabras.
  - d) Comprar acciones de una empresa de IA.
109. ¿Qué valor asegura que el modelo no mienta sobre lo que sabe?
- a) Amabilidad.
  - b) Honestidad.
  - c) Inocencia.
  - d) Robustez.
110. ¿Cuál es la función de un "Modelo de Recompensa" en RLHF?
- a) Dar dinero a los usuarios.
  - b) Estimar matemáticamente la puntuación que un humano daría a una respuesta.
  - c) Borrar los datos malos.
  - d) Traducir el prompt.
111. ¿Qué riesgo existe al usar datos de Wikipedia para entrenar modelos médicos?
- a) Que el modelo se vuelva enciclopédico.
  - b) Que los datos no sean lo suficientemente precisos o verificados para decisiones críticas de salud.
  - c) Que Wikipedia cobre a AWS.
  - d) Ninguno.
112. ¿Qué servicio de AWS ayuda a documentar qué datos se usaron para entrenar un modelo?
- a) SageMaker Model Cards.
  - b) Amazon Polly.
  - c) AWS Shield.
  - d) Amazon S3 Glacier.
113. ¿Qué es el "Sesgo demográfico"?
- a) El número de personas en el mundo.
  - b) Trato desigual por parte de la IA basado en raza, género o edad.
  - c) Un error de red local.
  - d) La versión del sistema operativo.
114. ¿Cuál es el objetivo del "Human-over-the-loop"?
- a) Que el humano haga todo el trabajo.
  - b) Que el humano tenga la autoridad final y control total sobre el despliegue del sistema.

- c) Que el humano vuele sobre los servidores.
  - d) No existe ese término.
115. ¿Qué técnica busca reducir el "Olvido Catastrófico"?
- a) Refinamiento Multitarea.
  - b) Poda radical.
  - c) Borrado de memoria.
  - d) Inferencia serverless.
116. ¿Qué mide el "Área bajo la curva" (AUC) en un clasificador binario?
- a) El tamaño del dataset.
  - b) El rendimiento agregado del modelo en todos los umbrales posibles.
  - c) El precio de la GPU.
  - d) El número de usuarios activos.
117. ¿Cuál es una ventaja de los "Modelos de baja complejidad" para la banca?
- a) Son más bonitos.
  - b) Son altamente interpretables, cumpliendo con regulaciones de transparencia.
  - c) Son más caros.
  - d) No necesitan datos.
118. ¿Qué es la "Detección de PII" en Amazon Comprehend?
- a) Un juego de preguntas.
  - b) Identificar datos sensibles (nombres, tarjetas) para proteger la privacidad.
  - c) Un error de memoria.
  - d) El nombre de una red neuronal.
119. ¿Qué riesgo existe si un atacante conoce las "atribuciones de características" de un modelo?
- a) Que el modelo se apague.
  - b) Que pueda realizar ataques de evasión o inyección más efectivos al saber qué es lo que más importa al modelo.
  - c) Que el modelo hable más rápido.
  - d) Ninguno.
120. ¿Qué valor humano prioriza que el modelo admita sus errores?
- a) Honestidad.
  - b) Amabilidad.
  - c) Inocencia.
  - d) Robustez.

121. ¿Cuál es la función del Container de procesamiento.<sup>en</sup> SageMaker Clarify?
- a) Guardar comida.
  - b) Ejecutar los cálculos de métricas de sesgo y explicabilidad interactuando con S3 y el endpoint del modelo.
  - c) Transportar hardware.
  - d) Es una base de datos.
122. ¿Qué métrica mide la "probabilidad de que el modelo incluya sesgos por religión?
- a) Estereotipos de peticiones (SageMaker Clarify).
  - b) Latencia de red.
  - c) Puntuación BLEU.
  - d) Número de tokens.
123. El término "Caja Blanca" se refiere a:
- a) Un modelo pintado de blanco.
  - b) Un modelo totalmente interpretable donde se ven todos sus mecanismos internos.
  - c) Un modelo de Amazon.
  - d) Un error de visualización.
124. ¿Cuál es la prioridad de la IA Responsable respecto al empleo?
- a) Despedir a todos.
  - b) Evaluar el impacto de la automatización y buscar que la IA aumente las capacidades humanas.
  - c) No importa el empleo.
  - d) Sustituir a los jefes.
125. ¿Qué técnica ayuda a evitar que un LLM sea "grosero?
- a) Filtros de toxicidad en Guardrails.
  - b) Aumentar la temperatura.
  - c) Borrar el prompt.
  - d) Usar solo números.
126. ¿Qué significa que un modelo sea "Modular?
- a) Que es muy grande.
  - b) Que sus componentes se pueden modificar o ampliar fácilmente para adaptarlos a tareas específicas.
  - c) Que vive en un módulo de AWS.
  - d) Que es gratuito.
127. ¿Cuál es la función de los "Informes Visuales" de Clarify?

- a) Ver películas.
  - b) Proporcionar una forma fácil de entender los resultados del análisis de sesgo y explicabilidad para humanos.
  - c) Decorar la oficina.
  - d) No existen.
128. ¿Qué describe mejor el Consentimiento Informado?
- a) Que el usuario pulse "Aceptar" sin leer.
  - b) Que los interesados sepan exactamente cómo y para qué se usarán sus datos antes de darlos.
  - c) Que el modelo sepa todo.
  - d) Un error legal.
129. ¿Qué es la "Vulnerabilidad de inyección"?
- a) Una vacuna.
  - b) Cuando un usuario manipula el prompt para saltarse las reglas éticas del modelo.
  - c) Un error de hardware.
  - d) Un tipo de cifrado.
130. ¿Qué valor humano asegura que el modelo no incite al odio?
- a) Inocencia/Seguridad.
  - b) Amabilidad.
  - c) Honestidad.
  - d) Robustez.
131. ¿Por qué es vital la "Auditoría periódica" de datasets?
- a) Para gastar dinero.
  - b) Para identificar sesgos que puedan haber aparecido con el tiempo o cambios en los datos.
  - c) Para que el dataset brille.
  - d) No es necesaria.
132. ¿Qué métrica penaliza respuestas de IA demasiado cortas en traducción?
- a) Brevity Penalty (BLEU).
  - b) Accuracy.
  - c) Recall.
  - d) RMSE.
133. ¿Qué técnica de Kearns asegura que un modelo "ignore" datos personales?
- a) Privacidad Diferencial.

- b) Sharding.
  - c) RAG.
  - d) Prompting.
134. ¿Cuál es la ventaja de la IA Agéntica.<sup>en</sup> la toma de decisiones?
- a) No toma decisiones.
  - b) Puede evaluar múltiples estrategias y adaptarse a cambios para cumplir metas complejas.
  - c) Es más barata que un humano.
  - d) Siempre acierta.
135. ¿Qué significa que un modelo sea "Factua
- a) Que es muy rápido.
  - b) Que sus respuestas se basan en hechos reales y no son inventadas.
  - c) Que es de Facebook.
  - d) Que tiene muchas capas.
136. ¿Qué valor humano prioriza que el chatbot ayude de verdad al usuario?
- a) Amabilidad (Helpfulness).
  - b) Honestidad.
  - c) Inocencia.
  - d) Robustez.
137. ¿Qué describe mejor el "Bias de Etiquetado
- a) Poner mal las etiquetas físicas de los servidores.
  - b) Cuando los humanos que etiquetaron los datos de entrenamiento introdujeron sus propios prejuicios en ellos.
  - c) Un error de formato CSV.
  - d) No existe ese término.
138. ¿Cuál es la función de Amazon Augmented AI.<sup>en</sup> medicina?
- a) Operar a los pacientes.
  - b) Permitir que un médico valide las sugerencias de la IA en casos difíciles.
  - c) Pagar las facturas del hospital.
  - d) Es un robot cirujano.
139. ¿Qué métrica de Clarify mide si la tasa de "Falsos Negativos.<sup>es</sup> igual entre grupos?
- a) Igualdad de trato.
  - b) Latencia.
  - c) Exactitud.

- d) RMSE.
140. El concepto "Atención es todo lo que necesitas" (2017) es la base de:
- a) Regresión Lineal.
  - b) Transformadores.
  - c) Redes Neuronales Recurrentes.
  - d) Árboles de decisión.
141. ¿Cuál es la prioridad de AWS en "Privacidad de datos"?
- a) Ver los datos de los clientes.
  - b) Garantizar que los datos del cliente nunca se usen para entrenar modelos de otros sin permiso.
  - c) Borrar los datos al azar.
  - d) Compartir los datos en internet.
142. ¿Qué describe la técnica RAG para la veracidad?
- a) Borrar los datos falsos.
  - b) Fundamentar las respuestas en documentos reales externos para reducir alucinaciones.
  - c) Cambiar el nombre del modelo.
  - d) No mejora la veracidad.
143. ¿Qué valor humano asegura que el modelo no sea ofensivo?
- a) Inocencia/Seguridad.
  - b) Amabilidad.
  - c) Honestidad.
  - d) Robustez.
144. ¿Qué significa "alineación de Modelos" (Model Alignment)?
- a) Poner los servidores en línea recta.
  - b) Asegurar que el comportamiento de la IA coincida con las intenciones y valores humanos.
  - c) Cambiar la versión del modelo.
  - d) Poner el logo en el centro.
145. ¿Qué herramienta permite a expertos auditar el funcionamiento de la IA de Amazon?
- a) Tarjetas de Servicio de IA (AI Service Cards).
  - b) Amazon Polly.
  - c) AWS Shield.
  - d) Amazon RDS.



146. ¿Cuál es el último paso para garantizar una IA responsable?

- a) Publicar el modelo.
- b) Monitorización continua y revisión periódica de resultados y sesgos en producción.
- c) Borrar el código.
- d) No hacer nada más.

## Clave de Respuestas (Dominio 4)

1. b	31. b	61. b	91. b	121. b
2. b	32. b	62. b	92. a	122. b
3. b	33. b	63. a	93. b	123. b
4. a	34. a	64. b	94. b	124. b
5. b	35. b	65. b	95. b	125. b
6. b	36. b	66. b	96. b	126. b
7. b	37. b	67. b	97. b	127. b
8. b	38. b	68. b	98. b	128. a
9. b	39. a	69. b	99. b	129. a
10. b	40. b	70. b	100. b	130. b
11. b	41. b	71. a	101. a	131. b
12. b	42. b	72. a	102. b	132. b
13. b	43. b	73. b	103. b	133. a
14. b	44. b	74. b	104. b	134. b
15. b	45. b	75. b	105. b	135. b
16. b	46. b	76. b	106. b	136. b
17. b	47. b	77. b	107. b	137. b
18. b	48. a	78. b	108. b	138. b
19. b	49. b	79. b	109. b	139. b
20. b	50. b	80. a	110. b	140. b
21. b	51. b	81. b	111. b	141. b
22. a	52. b	82. a	112. a	142. a
23. b	53. a	83. b	113. b	143. b
24. b	54. b	84. b	114. b	144. a
25. b	55. b	85. a	115. a	145. b
26. b	56. b	86. b	116. b	146. a
27. b	57. b	87. a	117. b	147. b
28. a	58. b	88. b	118. b	148. a
29. b	59. b	89. b	119. a	149. b
30. b	60. b	90. b	120. b	150. b