

Simulazione d'Esame - Risposte

Set di simulazione A

Versione 1.1

ISTQB® Testing with Generative AI Syllabus

Specialist Level

Compatibile con il Syllabus versione 1.1

International Software Testing Qualifications Board



Avviso di copyright

Avviso sul Copyright © International Software Testing Qualifications Board (di seguito chiamato ISTQB®).

ISTQB® è un marchio registrato di International Software Testing Qualifications Board.

Tutti i diritti riservati.

Gli autori dichiarano con la presente di trasferire il copyright a ISTQB®. Gli autori (come attuali titolari del copyright) e ISTQB® (come futuro titolare del copyright) hanno concordato le seguenti condizioni di utilizzo:

Possono essere copiati estratti di questo documento per un uso non commerciale se la fonte è riconosciuta.

Qualsiasi azienda Training Provider accreditata alla Formazione può utilizzare questa Simulazione d'Esame nei loro corsi di formazione se gli autori e ISTQB® sono riconosciuti come fonti e proprietari del copyright della Simulazione d'Esame, e a condizione che qualsiasi pubblicità di tale corso di formazione sia fatta dopo che l'accreditamento ufficiale dei materiali di formazione è stato ricevuto da un Member Board riconosciuto da ISTQB®.

Qualsiasi individuo o gruppo di individui può utilizzare questa Simulazione d'Esame negli articoli e nei libri, a condizione che gli autori e ISTQB® siano riconosciuti come fonti e possessori del copyright della Simulazione d'Esame.

È proibito qualsiasi altro utilizzo di questa Simulazione d'Esame senza prima avere ottenuto l'approvazione scritta di ISTQB®.

Qualsiasi Member Board riconosciuto da ISTQB® può tradurre questa Simulazione d'Esame a condizione di riprodurre il sopra menzionato Avviso di Copyright nella versione tradotta della Simulazione d'Esame.

Responsabilità del documento

Exam Working Group ISTQB® è responsabile di questo documento.

Questo documento è mantenuto da un team dedicato di ISTQB®, che consiste nei Syllabus Working Group e nell'Exam Working Group.

Riconoscimenti

Questo documento è stato prodotto da un core team di ISTQB®: Abbas Ahmad, Alessandro Collino, Bruno Legeard.

Il core team ringrazia il team di review Exam Working Group, il Syllabus Working Group e i Member Board per i loro suggerimenti e contributi.

Traduzione italiana a cura di ITA-STQB (www.ita-stqb.org) – info@ita-stqb.org

Revisione della traduzione italiana a cura di: Raffaele Bozzini, Salvatore Reale.

Cronologia delle revisioni

Simulazione d'esame - Modello di layout delle risposte utilizzato: Versione 2.12 Data: 27 novembre 2024

Versione	Data	Osservazioni
V1.0	2024/12/30	CT-GenAI Set di simulazione A 1.0 Alpha Release
V1.0	2025/04/17	CT-GenAI Set di simulazione A 1.0 Beta Release
V1.0	2025/06/10	CT-GenAI Set di simulazione A 1.0 Release
V1.1	2026/04/27	CT-GenAI Set di simulazione A 1.1 Aggiornamento minore

Indice dei contenuti

Avviso di copyright	2
Responsabilità del documento	2
Ringraziamenti	2
Cronologia delle revisioni	4
Indice	5
Introduzione	7
Scopo del presente documento	7
Istruzioni	7
Risposte	8
Risposte	9
1	9
2	10
3	10
4	11
5	12
6	12
7	13
8	13
9	14
10	15
11	15
12	16
13	17
14	17
15	18
16	18
17	19
18	20
19	20
20	21
21	21
22	22
23	23
24	24
25	25
26	25
27	26
28	27
29	28
30	28
31	29
32	30
33	30
34	31
35	32
36	33
37	34
38	35
39	36
40	36
Appendice: Risposte alle domande aggiuntive	37
A1	37
A2	38

A3..... 39

Introduzione

Scopo del documento

Le domande di questa simulazione, le risposte e le relative giustificazioni associate alle domande di questa simulazione d'esame sono state create da un team di esperti in materia e da scrittori esperti di domande, con l'obiettivo di:

- Supportare i Member Board e l'Exam Board ISTQB® nelle loro attività di scrittura delle domande
- Fornire ai Training Provider e ai candidati all'esame esempi di domande d'esame

Queste domande non possono essere utilizzate senza essere modificate in nessun esame ufficiale.

Si noti che gli esami reali possono includere un'ampia varietà di domande, e questa simulazione d'esame **non** intende includere esempi di tutti i possibili tipi, stili o lunghezze di domande; inoltre, questa simulazione d'esame può essere più difficile oppure meno difficile di qualsiasi esame ufficiale.

Istruzioni

In questo documento potete trovare:

- Tabella delle Risposte, includendo per ogni risposta corretta:
 - Livello di conoscenza (Livello K), obiettivo di apprendimento (LO) e valore del punteggio
- Insieme di risposte, includendo per tutte le domande:
 - Risposta corretta
 - Giustificazione per ogni opzione di risposta
 - Livello di conoscenza (Livello K), obiettivo di apprendimento (LO) e valore del punteggio
- Insieme di risposte aggiuntive, includendo per tutte le domande [non si applica a tutte le simulazioni d'esame]:
 - Risposta corretta
 - Giustificazione per ogni opzione di risposta
 - Livello di conoscenza (Livello K), obiettivo di apprendimento (LO) e valore del punteggio
- *Le domande sono contenute in un documento separato*

Tabella Riassuntiva delle Risposte

Numero domanda (#)	Risposta corretta	LO	Livello K	Punti
1	d	GenAI-1.1.1	K1	1
2	c	GenAI-1.1.2	K2	1
3	b	GenAI-1.1.2	K2	1
4	d	GenAI-1.1.3	K2	1
5	b	GenAI-1.1.4	K2	1
6	a, e	GenAI-1.2.1	K2	1
7	d	GenAI-1.2.2	K2	1
8	b	GenAI-2.1.1	K2	1
9	c	GenAI-2.1.1	K2	1
10	b	GenAI-2.1.2	K2	1
11	a	GenAI-2.1.3	K2	1
12	a	GenAI-2.2.1	K3	2
13	b	GenAI-2.2.2	K3	2
14	c	GenAI-2.2.3	K3	2
15	c	GenAI-2.2.4	K3	2
16	a	GenAI-2.2.5	K3	2
17	a	GenAI-2.3.1	K2	1
18	a	GenAI-2.3.2	K2	1
19	c	GenAI-3.1.1	K1	1
20	d	GenAI-3.1.2	K3	2

Numero domanda (#)	Risposta corretta	LO	Livello K	Punti
21	b	GenAI-3.1.3	K2	1
22	b	GenAI-3.1.4	K1	1
23	d	GenAI-3.2.1	K2	1
24	d	GenAI-3.2.2	K2	1
25	a	GenAI-3.2.2	K2	1
26	b	GenAI-3.2.3	K2	1
27	c	GenAI-3.3.1	K2	1
28	b, d	GenAI-3.4.1	K1	1
29	a	GenAI-4.1.1	K2	1
30	a	GenAI-4.1.2	K2	1
31	c	GenAI-4.1.3	K2	1
32	b	GenAI-4.2.1	K2	1
33	b	GenAI-4.2.2	K2	1
34	d	GenAI-5.1.1	K1	1
35	b	GenAI-5.1.2	K2	1
36	b	GenAI-5.1.3	K2	1
37	a	GenAI-5.1.4	K1	1
38	c	GenAI-5.2.1	K2	1
39	c	GenAI-5.2.2	K1	1
40	a	GenAI-5.2.3	K1	1

Risposte

Numero domanda (#)	Risposta corretta	Spiegazione / Motivazione	Obiettivo di apprendimento (LO)	Livello K	Numero di punti
1	d	Considerando che: - L'AI simbolica utilizza un sistema rule-based per imitare il processo decisionale umano, rappresentando la conoscenza tramite simboli e regole logiche (1B). - Il machine learning classico utilizza un approccio data-driven che richiede la preparazione dei dati, la selezione delle funzionalità e l'addestramento del modello (2D). - Il deep learning utilizza reti neurali (strutture di machine learning) per apprendere automaticamente le funzionalità dai dati (3A). - L'AI Generativa utilizza tecniche di deep learning per creare nuovi dati apprendendo e imitando i modelli dai suoi dati di addestramento (4C). Quindi: a) Non è corretta. b) Non è corretta. c) Non è corretta. d) È corretta.	GenAI-1.1.1	K1	1

2	c	a) Non è corretta perché le finestre di contesto non controllano la sequenza temporale, ma limitano l'ambito del testo che può essere considerato simultaneamente. b) Non è corretta perché le finestre di contesto influenzano l'ambito all'interno dell'input corrente, non le capacità di riferimento tra documenti. c) È corretta. Quando il testo supera le dimensioni della finestra di contesto, il modello non può considerare contemporaneamente tutte le parti del documento. Man mano che il modello elabora nuovi token, deve, in modo efficace, "dimenticare" o scartare i token che non rientrano nei limiti della finestra di contesto. d) Non è corretta perché le finestre di contesto non determinano gli approcci di analisi, ma limitano la quantità di testo che può essere elaborata contemporaneamente.	GenAI-1.1.2	K2	1
3	b	a) Non è corretta. Descrive gli embedding, non la tokenizzazione. b) È corretta. La tokenizzazione comporta la suddivisione del testo in unità più piccole (token) che rappresentano i building block delle attività di generazione del linguaggio naturale e consente ai LLM di comprendere e generare testo. Si veda anche la definizione di "Tokenizzazione" nel syllabus (vedere "Appendice D - Termini Specifici dell'AI Generativa"). c) Non è corretta. Descrive la funzione generale dei LLM, non la tokenizzazione. d) Non è corretta. Si riferisce al modo in cui i LLM generano testo, non alla tokenizzazione.	GenAI-1.1.2	K2	1

4	d	Considerando che: i. Non è corretto. Sebbene i LLM di base siano in grado di generare test case, non eccellono intrinsecamente in questa attività senza un input strutturato. Generare test case senza un input strutturato significa travisare le loro capacità. ii. Non è corretto. La creazione di test script basati su modelli richiede l'esecuzione di istruzioni esplicite, che è il ruolo dei LLM ottimizzati con istruzioni, non dei LLM di ragionamento. I LLM di ragionamento sono specializzati nell'inferenza logica e nella risoluzione di problemi, non nell'adesione rigida ai modelli. iii. Non è corretto. I LLM ottimizzati con istruzioni sono progettati per seguire prompt strutturati, non per prendere decisioni autonome. Prioritizzare i test in base al feedback richiede ragionamento, che va oltre il loro ambito di competenza. iv. È corretto. I LLM di ragionamento sono progettati esplicitamente per sintetizzare più fonti di dati ed eseguire inferenze logiche, risoluzione di problemi e processi decisionali. Rilevare le tendenze e prioritizzare l'effort di test è in linea con il loro ruolo nell'analisi contestuale e con la descrizione fornita. v. È corretto. I LLM ottimizzati con istruzioni sono specificamente addestrati a seguire le istruzioni, compreso il rispetto dei formati, degli stili e delle regole sintattiche richiesti. La generazione di test case conformi alla sintassi del linguaggio Gherkin è un'attività che ben si adatta alle loro capacità. Quindi: a) Non è corretta. b) Non è corretta. c) Non è corretta. d) È corretta.	GenAI-1.1.3	K2	1
---	---	---	-------------	----	---

5	b	a) Non è corretta. I modelli linguistico-visivi sono un sottoinsieme dei LLM multimodali, non il contrario. b) È corretta. I modelli linguistico-visivi integrano specificamente dati visivi e testuali, rendendoli un sottoinsieme dei LLM multimodali. c) Non è corretta. I modelli linguistico-visivi sono strettamente correlati ai LLM multimodali e si focalizzano sia sui dati visivi che su quelli testuali. d) Non è corretta. I LLM multimodali e i modelli linguistico-visivi hanno ambiti distinti e non sono intercambiabili.	GenAI-1.1.4	K2	1
6	a, e	a) È corretta. I LLM possono analizzare e chiarire i requisiti identificando ambiguità e incongruenze. b) Non è corretta. La generazione di codice applicativo completo non è una capacità principale dei LLM nelle attività di test. c) Non è corretta. I LLM possono supportare la test automation suggerendo miglioramenti ai test script e identificando modelli di progettazione, ma non eseguono direttamente i test script né li automatizzano completamente senza la supervisione umana. d) Non è corretta. I LLM non possono eseguire testing esplorativo manuale perché si tratta di un processo intuitivo e adattivo che richiede creatività, esperienza e capacità decisionale umana. e) È corretta. I LLM possono generare dati di test diversificati, comprese combinazioni e limiti.	GenAI-1.2.1	K2	1

7	d	a) Non è corretta. I chatbot AI sono più adatti per interazioni ad hoc, non per attività di test specifiche. b) Non è corretta. I chatbot AI e le applicazioni di testing alimentate da LLM hanno scopi distinti e non sono identici in termini di funzionalità o configurabilità. c) Non è corretta. Le applicazioni di testing alimentate da LLM si focalizzano sull'integrazione nei processi di test, non sui prompt conversazionali. I chatbot AI non richiedono l'integrazione negli strumenti di test e nei processi di test perché la loro funzione principale è quella di facilitare le interazioni ad hoc piuttosto che eseguire attività di test specifiche. d) È corretta. I chatbot AI forniscono interfacce conversazionali per attività ad hoc, mentre le applicazioni di testing alimentati da LLM offrono soluzioni personalizzate per esigenze specifiche.	GenAI-1.2.2	K2	1
8	b	a) Non è corretta. Il contesto fornisce informazioni di base sull'ambiente di test o sul sistema sottoposto a testing, non dati specifici (di input) da analizzare. Sebbene si faccia riferimento al testing delle prestazioni, vengono elencate fonti di dati effettive anziché fornire informazioni di base. b) È corretta. La descrizione elenca fonti di dati specifiche (test report, log di monitoraggio, benchmark delle prestazioni) che saranno elaborate dal LLM per eseguire l'attività di analisi. c) Non è corretta. I vincoli delineano restrizioni o considerazioni speciali su come l'attività dovrebbe essere eseguita. La descrizione elenca fonti di dati effettive piuttosto che restrizioni sull'approccio di analisi. d) Non è corretta. Il formato di output specifica la struttura e le caratteristiche attese della risposta del LLM. La descrizione elenca fonti di dati effettive, senza alcun riferimento a come dovrebbero essere presentati i risultati.	GenAI-2.1.1	K2	1

9	c	a) Non è corretta. La riga indicata non descrive quale attività eseguire, ma piuttosto come presentare i risultati. Le istruzioni indicano al LLM cosa fare, mentre questa riga indica al LLM come formattare l'output. b) Non è corretta. La riga indicata non impone restrizioni o limitazioni al processo di analisi, ma piuttosto specifica il formato di presentazione desiderato. I vincoli potrebbero includere aspetti come "escludere problemi estetici", mentre questa riga riguarda la formattazione dell'output. c) È corretta. La riga indicata corrisponde alla definizione del syllabus delle specifiche del formato di output che guidano il modo in cui il LLM dovrebbe formattare l'output. d) Non è corretta. La riga indicata non fornisce informazioni di base come il contesto della specifica dei requisiti. Riguarda invece la formattazione dell'output.	GenAI-2.1.1	K2	1
---	---	---	-------------	----	---

10	b	a) Non è corretta. Il prompt chaining in realtà è definito dalla scomposizione di un'attività in prompt sequenziali, non dalla "fornitura di esempi". Il few-shot prompting è la tecnica che fornisce esempi, non quella che "scompone le attività in sottoattività". Il meta prompting fa affidamento sul LLM per rivedere i prompt da solo, non sul tester che "perfeziona i prompt manualmente". b) È corretta. Il few-shot prompting fornisce una guida con esempi, il prompt chaining scompone le attività in passaggi intermedi (prompt multipli) e il meta prompting utilizza l'AI per perfezionare i propri prompt in modo iterativo. c) Non è corretta. Lo scopo principale del meta prompting è il perfezionamento automatico dei prompt da parte del LLM, non la "scomposizione dell'attività in passaggi". Il prompt chaining è caratterizzato dalla scomposizione step-by-step, non dall'"uso di esempi" come affermato. Il few-shot prompting è il metodo che fornisce esempi al modello e non è incentrato sull'"ottimizzazione manuale dei prompt". d) Non è corretta. Il prompt chaining è definito dalla scomposizione graduale, non dalla "guida senza esempi". Il meta prompting fa affidamento sul LLM per generare/perfezionare i prompt, non solo sulla formulazione definita dal tester.	GenAI-2.1.2	K2	1
11	a	a) È corretta. Il prompt di sistema rimane costante durante tutta la sessione di interazione e stabilisce il framework fondamentale per il modo in cui il LLM dovrebbe rispondere. b) Non è corretta. Descrive la funzione di un prompt utente, non di un prompt di sistema. c) Non è corretta. Il prompt di sistema non si adatta dinamicamente, ma rimane costante. d) Non è corretta. Il prompt di sistema è nascosto e non include input visibili da parte dell'utente.	GenAI-2.1.3	K2	1

12	a	Considerando che: i. È corretto perché garantisce che il processo inizi con la generazione delle condizioni di test dai requisiti, in linea con la descrizione del syllabus che afferma che l'analisi dei test con la GenAI comporta la generazione di condizioni di test basate sulla base di test, ad esempio sui requisiti. ii. È corretto perché chiarisce la necessità di criteri di accettazione (condizioni di test) per migliorare gli output generati dal LLM e segue le indicazioni del syllabus secondo cui i LLM possono prioritizzare le condizioni di test in base al livello di rischio quando viene fornito un contesto adeguato. iii. È corretto perché indirizza il LLM a eseguire l'analisi della copertura, affrontando tutti gli aspetti dei requisiti, il che corrisponde all'affermazione del syllabus secondo cui i LLM possono eseguire l'analisi della copertura per determinare se tutti gli aspetti della base di test sono coperti. iv. Non è corretto perché affidarsi a input minimi senza indicazioni non è in linea con la tecnica del prompt chaining né garantisce l'efficacia. Questa opzione cerca di fare tutto in un unico passaggio invece di suddividere il processo. v. Non è corretto perché l'identificazione dei difetti non è fondamentale per l'obiettivo di test di generare condizioni di test prioritate e identificare le lacune della copertura. Lo scenario afferma specificamente che i requisiti sono stabili e già accuratamente sottoposti a review per individuare eventuali difetti (che in genere includono ambiguità e incongruenze). Quindi: a) È corretta. b) Non è corretta. c) Non è corretta. d) Non è corretta.	GenAI-2.2.1	K3	2
----	---	---	-------------	----	---

13	b	a) Non è corretta. Sebbene specifichi l'uso di esempi predefiniti, non richiede esplicitamente l'uso della sintassi "Given-When-Then", fondamentale per i test case in stile Gherkin. Indica inoltre l'affidamento a vaghe best practice nella definizione dei vincoli e non garantisce specificamente l'allineamento con il criterio di accettazione. b) È corretta. È completo e utilizza esempi predefiniti per guidare il LLM. c) Non è corretta. Manca l'enfasi sull'uso di esempi predefiniti e indica l'affidamento a vaghe best practice nella definizione delle istruzioni. d) Non è corretta. Si focalizza sui casi limite, ma trascura la copertura completa e l'uso di esempi come guida.	GenAI-2.2.2	K3	2
14	c	a) Non è corretta. Affronta il clustering e il controllo incrociato, ma traslascia altri passaggi fondamentali come la separazione dei risultati dei test. b) Non è corretta. Migliora la chiarezza dei ruoli, ma non amplia le istruzioni né affronta i passaggi strutturati. c) È corretta. Amplia le istruzioni per includere tutti i passaggi critici dell'analisi strutturata. Separare i risultati attesi dai risultati effettivi aiuta a individuare efficacemente le discrepanze, il clustering dei problemi facilita una migliore prioritizzazione e riduce la ridondanza, mentre evidenziare le discrepanze aiuta a focalizzare l'attenzione sulle scoperte più importanti. d) Non è corretta. Introduce vincoli irrilevanti, allineando in modo errato il prompt con i requisiti dell'attività.	GenAI-2.2.3	K3	2

15	c	a) Non è corretta. Chiarisce il ruolo ma non aggiunge alcuna istruzione concreta che migliori l'accuratezza, la praticabilità o l'interpretabilità delle metriche fornite. b) Non è corretta. Aggiunge un'attività di analisi del rischio che potrebbe distrarre il modello dalle metriche fondamentali e continua a non offrire alcun meccanismo per rendere chiari i risultati agli stakeholder. c) È corretta. L'ampliamento del formato di output con un riassunto in linguaggio semplice che interpreta le metriche e delinea i passi successivi supporta direttamente la comprensione e la praticabilità da parte degli stakeholder (vedere anche il syllabus nella sezione 2.2.4: "Visualizzazione e reporting delle metriche di test migliorati"). d) Non è corretta. Si limita a ribadire un vincolo esistente e non fornisce indicazioni specifiche su come il LLM dovrebbe raggiungere l'interpretabilità a livello degli stakeholder.	GenAI-2.2.4	K3	2
16	a	a) È corretta. Il few-shot prompting è ideale in questo scenario perché consente di fornire alcuni esempi (i test case esistenti con input e risultati attesi) per guidare il LLM. Illustrando come queste regole di trasformazione vengono applicate per generare nuovi test case, il few-shot prompting può aiutare il LLM a comprendere e replicare il processo per generare ulteriori test case. b) Non è corretta. Sebbene sia possibile utilizzare il prompt chaining, ciò potrebbe complicare inutilmente questa attività molto semplice. c) Non è corretta. Sebbene sia possibile utilizzare il meta-prompting, non è così diretto e specifico come il few-shot prompting per affrontare questa attività molto semplice. d) Non è corretta. Il zero-shot prompting non sarebbe efficace poiché non utilizzerebbe i test case esistenti come esempi.	GenAI-2.2.5	K3	2

17	a	a) È corretta. La diversità garantisce una copertura completa dei casi limite e il tasso di successo dell'esecuzione dei test valuta l'affidabilità dei test script API (vedere la descrizione e l'esempio della metrica "diversità" forniti dal syllabus nella tabella nella sezione 2.3.1). b) Non è corretta. Sebbene l'accuratezza e la completezza siano importanti, affidarsi principalmente all'efficienza in termini di tempo non consente di valutare appieno la copertura o l'affidabilità dell'esecuzione dei test. c) Non è corretta. La precisione e l'adeguatezza contestuale sono importanti, ma non tengono conto della diversità né valutano in modo approfondito il successo dell'esecuzione dei test per i test script API. d) Non è corretta. La rilevanza e l'adeguatezza contestuale sono fondamentali, ma senza considerare metriche come il tasso di successo dell'esecuzione dei test o l'accuratezza, si tralasciano aspetti critici della valutazione.	GenAI-2.3.1	K2	1
----	---	---	-------------	----	---

18	a	a) È corretta. L'analisi dell'output esamina gli output dei LLM alla ricerca di imprecisioni o incongruenze. Classificando i risultati attesi errati che contraddicono i requisiti, si rivela perché il prompt ha fuorviato il LLM e fornisce indicazioni concrete per il perfezionamento del prompt. b) Non è corretta. Il testing A/B dei prompt è più adatto per confrontare diverse versioni dei prompt che per diagnosticare la causa dei risultati attesi errati. c) Non è corretta. Sebbene la regolazione della lunghezza e della specificità dei prompt possa migliorare la qualità delle risposte aggiungendo o riducendo il contesto, non affronta direttamente la causa dei risultati attesi errati. d) Non è corretta. Sebbene l'utilizzo delle informazioni raccolte dai tester sull'utilità e la chiarezza degli output generati possa aiutare a perfezionare i prompt per soddisfare meglio le esigenze di testing nel mondo reale, non affronta direttamente la causa dei risultati attesi errati.	GenAI-2.3.2	K2	1
19	c	a) Non è corretta. Descrive errori di ragionamento, non allucinazioni. b) Non è corretta. Descrive bias nell'output dell'AI, non allucinazioni. c) È corretta. Le allucinazioni si verificano quando il LLM genera output che è di fatto errato o irrilevante per una determinata attività. Vedere il syllabus nella sezione 3.1.1. d) Non è corretta. Descrive i bias dovuti alla sottorappresentazione nei dati di addestramento, non le allucinazioni.	GenAI-3.1.1	K1	1

20	d	a) Non è corretta. La gestione del carrello è esplicitamente menzionata nel briefing del progetto, rendendo questo test case rilevante. b) Non è corretta. L'applicazione di codici sconto è esplicitamente menzionata nel briefing del progetto, rendendo questo test case rilevante. c) Non è corretta. Le e-mail di conferma dell'ordine sono esplicitamente menzionate nel briefing del progetto, quindi questo test case è valido. d) È corretta. La gestione della lista dei desideri non è menzionata nel briefing del progetto, rendendo questo il più probabile test case che rappresenta un'allucinazione.	GenAI-3.1.2	K3	2
21	b	a) Non è corretta. L'effort per perfezionare i LLM per le attività di test è associato a un ulteriore addestramento di questi modelli su un set di dati mirato, che consente loro di apprendere le conoscenze e le sfumature specifiche del dominio necessarie per eseguire tali attività di test. L'uso di formati di dati di input chiari e strutturati non influisce su questo effort. b) È corretta. Quando i dati di input sono presentati in modo chiaro e strutturato, i potenziali malintesi sono ridotti al minimo. L'ambiguità deriva spesso da dati poco chiari e mal strutturati. c) Non è corretta. La rilevanza del contesto dipende più dalla fornitura delle giuste informazioni contestuali che dal semplice utilizzo di formati di dati di input chiari e strutturati. d) Non è corretta. L'uso di formati di dati di input chiari e strutturati non aumenta la creatività dei LLM nella generazione di output. In particolare, l'uso di questi formati non incoraggia risposte innovative, poiché impone ai LLM di generare risposte che rispettino i formati specificati.	GenAI-3.1.3	K2	1

22	b	a) Non è corretta. Il tasso di apprendimento è un'impostazione (scelta prima dell'inizio dell'addestramento) che controlla il processo di apprendimento delle reti neurali. Questa impostazione determina quanto modificare i pesi del modello durante l'addestramento. Non è correlata alla variabilità dell'inferenza. Aumentarla può portare a una convergenza più rapida durante l'addestramento, ma non influisce sulla variabilità degli output durante l'inferenza. b) È corretta. La temperatura è un parametro che controlla la casualità dell'output agendo sulla distribuzione di probabilità durante l'inferenza. Abbassando la temperatura si riduce la casualità, ottenendo output più coerenti. c) e d) non sono corrette. Sebbene l'impostazione di un seme casuale possa migliorare la riproducibilità, di per sé non restringe la distribuzione di probabilità durante l'inferenza. Inoltre, il fatto che il suo valore sia alto o basso è irrilevante.	GenAI-3.1.4	K1	1
----	---	--	-------------	----	---

23	d	a) Non è corretta. Si tratta di una preoccupazione della privacy, poiché l'"esposizione involontaria dei dati" è legata agli output dell'AI Generativa (GenAI) che rivelano accidentalmente informazioni sensibili. b) Non è corretta. Si tratta di una preoccupazione della privacy, poiché comporta la mancanza di controllo sulle modalità di archiviazione o elaborazione dei dati sensibili. c) Non è corretta. Si tratta di una preoccupazione della privacy legata alla non conformità con normative come il General Data Protection Regulation (GDPR), il che comporta rischi legali. d) È corretta. Un LLM non espone dati sensibili reali se ha allucinazioni durante la generazione di dati di test sintetici, purché non sia stato addestrato su dati sensibili reali. In questo caso, le allucinazioni sarebbero puramente sintetici e basati su modelli e strutture appresi dal modello durante l'addestramento. Il LLM potrebbe generare involontariamente dati di test sintetici che corrispondono a dati sensibili reali (e questa è sicuramente una preoccupazione), ma non esporrebbe dati sensibili reali. Tuttavia, questa generazione involontaria è molto improbabile.	GenAI-3.2.1	K2	1
----	---	---	-------------	----	---

24	d	a) Non è corretta. La generazione di codice dannoso comporta la "manipolazione di un LLM per generare backdoor (ad esempio, chiamate a comandi esterni) durante l'uso", non l'iniezione di dati falsi nei set di dati di addestramento. b) Non è corretta. La manipolazione del contesto comporta "l'invio di richieste progettate per estrarre dati di addestramento confidenziali", non l'iniezione di dati falsi nei set di dati di addestramento. c) Non è corretta. La manipolazione delle richieste comporta "l'introduzione di dati che interrompono l'output dell'AI" durante l'utilizzo in runtime, non l'iniezione di dati falsi nei set di dati di addestramento. d) È corretta. Secondo il syllabus nella sezione 3.2.2, l'avvelenamento dei dati comporta la "manipolazione dei dati di addestramento" con l'esempio specifico di "fornire valutazioni false durante la valutazione dei risultati di un test report generato dall'AI". La domanda descrive questo vettore di attacco: un attacker inietta risultati di (esecuzione dei) test falsificati nel set di dati di addestramento, il che manipola direttamente i dati di addestramento per compromettere la capacità del LLM di raccomandare con precisione strategie ottimali di copertura dei test.	GenAI-3.2.2	K2	1
----	---	---	-------------	----	---

25	a	Considerando che: • La manipolazione del contesto si basa sull'invio di richieste progettate per estrarre dati di addestramento riservati. (1C) • La manipolazione delle richieste si basa sull'introduzione di un'immagine che disturba l'output del LLM. (2D) • L'avvelenamento dei dati si basa sulla manipolazione delle raccomandazioni attraverso valutazioni false utilizzate nell'addestramento. (3A) • La generazione di codice dannoso si basa sulla manipolazione di un LLM per generare backdoor (ad esempio, chiamate a comandi esterni) durante l'uso. (4B) Quindi: a) È corretta. b) Non è corretta. c) Non è corretta. d) Non è corretta.	GenAI-3.2.2	K2	1
26	b	a) Non è corretta. Sebbene valutare gli output confrontando più LLM sia una pratica utile, in questo caso ci si focalizza sulla qualità degli output in termini di accuratezza senza affrontare i rischi relativi alla privacy dei dati. b) È corretta. L'anonimizzazione è una strategia efficace per mitigare i rischi della privacy dei dati. c) Non è corretta. L'accesso illimitato ai dati sensibili aumenta il rischio di violazione dei dati sensibili. d) Non è corretta. Disabilitare la crittografia dei dati sensibili indebolisce la sicurezza e aumenta il rischio di furto di dati sensibili.	GenAI-3.2.3	K2	1

27	c	a) Non è corretta. La generazione di immagini richiede molta più energia e produce molte più emissioni di CO₂ rispetto alla generazione di testo. Il consumo energetico e le emissioni di CO₂ sono generalmente correlati, a meno che non venga specificato un fattore chiave (come le differenze nelle fonti energetiche). b) Non è corretta. Al contrario, le ricerche alimentate dalla GenAI consumano molta più energia rispetto alle ricerche web tradizionali. c) È corretta. La generazione di immagini richiede risorse computazionali significativamente maggiori rispetto alla generazione di testo, rendendola molto più energivora. d) Non è corretta. L'impatto cumulativo della generazione di testo su milioni di utenti non è trascurabile.	GenAI-3.3.1	K2	1
----	---	---	-------------	----	---

28	b, d	a) Non è corretta. ISO/IEC 25010:2023 è uno standard (non menzionato nel syllabus, ma menzionato nel syllabus CTFL) che definisce un modello di qualità del prodotto relativo alle proprietà qualitative dei prodotti ICT/software. Non tratta l'uso della GenAI nel testing del software. b) È corretta. ISO/IEC 23053:2022 è uno standard (menzionato nella sezione 3.4.1 del syllabus) che fornisce un framework per la qualità dei dati, la trasparenza e la tolleranza ai guasti quando si utilizza la GenAI per il testing. c) Non è corretta. ISO/IEC/IEEE 29119-2:2021 è una parte (volume) dello standard (non menzionato nel syllabus, ma menzionato nel syllabus CTFL) che tratta i processi di test. Non affronta l'uso della GenAI nel testing del software. d) È corretta. ISO/IEC 42001:2023 è uno standard (menzionato nella sezione 3.4.1 del syllabus) che specifica i requisiti per la gestione dei sistemi basati sull'AI all'interno di un'organizzazione, fornendo le best practice per la GenAI nel testing del software e promuovendo la coerenza e l'affidabilità. e) Non è corretta. ISO/IEC/IEEE 29119-3:2021 è una parte (volume) dello standard (non menzionato nel syllabus, ma menzionato nel syllabus CTFL) che tratta la documentazione di test. Non affronta l'uso della GenAI nel testing del software.	GenAI-3.4.1	K1	1
----	------	---	-------------	----	---

29	a	a) È corretta. Il back-end è responsabile del recupero dei dati dai database relazionali e vettoriali, della loro combinazione con gli input dell'utente e della preparazione del prompt adattato per il LLM. b) Non è corretta. Il front-end funge da user interface per l'invio degli input, ma non prepara il prompt per il LLM. c) Non è corretta. Il componente di autenticazione garantisce la sicurezza dell'accesso, ma non è correlato al recupero dei dati o alla preparazione del prompt. d) Non è corretta. Il componente di post-elaborazione perfeziona l'output generato dal LLM, ma non gestisce la preparazione del prompt.	GenAI-4.1.1	K2	1
30	a	a) È corretta. La RAG recupera automaticamente le specifiche rilevanti e i blocchi di test case storici e li invia al LLM, che quindi genera test case accurati e sensibili al contesto. b) Non è corretta. La RAG è più efficace quando viene richiesto di recuperare informazioni specifiche mirate dal database vettoriale e non dall'intero set di dati. c) Non è corretta. L'esecuzione di una review manuale e il perfezionamento della query aggiungono un effort inutile, che contraddice il design della RAG di automatizzare il recupero e utilizzare i dati recuperati in modo dinamico nella generazione della risposta del LLM. d) Non è corretta. Affidarsi ai dati interni del LLM ignora la capacità fondamentale della RAG di migliorare le risposte integrando informazioni esterne e aggiornate.	GenAI-4.1.2	K2	1

31	c	a) Non è corretta. Gli agenti autonomi e semi-autonomi possono essere implementati come sistemi a agente singolo e/o multi-agente per migliorare sia l'efficienza che la qualità nei processi di test. Tuttavia, i miglioramenti principali apportati da questi agenti riguardano l'aumento dell'efficienza e della qualità nei processi di test grazie alla loro capacità di bilanciare l'autonomia con la supervisione umana. b) Non è corretta. Sebbene gli agenti autonomi e semi-autonomi incorporino verifiche per garantire la qualità, tali verifiche sono anche progettate per essere efficienti. L'uso di controlli automatizzati mira infatti a snellire i processi di test e a migliorare la qualità senza sacrificare l'efficienza (e viceversa), che a sua volta viene migliorata. c) È corretta. Gli agenti autonomi aumentano l'efficienza eseguendo attività di test all'interno dei processi di test con un intervento umano minimo e utilizzando controlli automatizzati progettati per essere efficienti. Gli agenti semi-autonomi prevedono una supervisione umana strategica, che garantisce la qualità dei risultati dei test. Questi due tipi di agenti consentono di implementare un approccio equilibrato per ottenere sia efficienza che qualità, utilizzando la loro capacità di operare con diversi livelli di interazione umana. d) Non è corretta. L'eliminazione completa della verifica non è né realistica né auspicabile. La verifica rimane fondamentale anche quando si utilizzano tecnologie avanzate come gli agenti autonomi e semi-autonomi alimentati da LLM.	GenAI-4.1.3	K2	1
----	---	---	-------------	----	---

32	b	a) Non è corretta. Questa è una descrizione corretta del fine-tuning, che migliora le prestazioni di un modello in un dominio specifico attraverso un addestramento specifico per l'attività. b) È corretta. Questa affermazione è errata perché il fine-tuning non sostituisce le conoscenze generali del modello e l'overfitting rimane una potenziale sfida. c) Non è corretta. Questa è una descrizione corretta, poiché il fine-tuning comporta la regolazione dei parametri utilizzando set di dati mirati. d) Non è corretta. Questa è una descrizione corretta delle sfide associate al fine-tuning, poiché i set di dati di alta qualità sono fondamentali per un adattamento efficace.	GenAI-4.2.1	K2	1
33	b	a) Non è corretta. Large Language Model Operations (LLMOps) si focalizza sulla gestione dei LLM, non sulla prevenzione del loro utilizzo. b) È corretta. LLMOps gestisce i LLM durante tutto il loro ciclo di vita, occupandosi di privacy, sicurezza e costi del testing. c) Non è corretta. LLMOps si applica a varie applicazioni, non solo alle soluzioni basate su chatbot. d) Non è corretta. LLMOps non mira ad automatizzare completamente tutte le attività di test o ad eliminare la supervisione umana.	GenAI-4.2.2	K2	1

34	d	a) Non è corretta. L'utilizzo di strumenti GenAI non approvati non garantisce il rispetto delle politiche organizzative in materia di dati. b) Non è corretta. L'AI ombra introduce rischi legati a accordi di licenza poco chiari, che possono portare a controversie in materia di proprietà intellettuale. c) Non è corretta. L'AI ombra aumenta, anziché ridurre, il rischio di controversie in materia di proprietà intellettuale. d) È corretta. Gli strumenti di AI non approvati spesso non dispongono di misure di sicurezza robuste, aumentando il rischio di violazioni dei dati o accessi non autorizzati.	GenAI-5.1.1	K1	1
----	---	--	-------------	----	---

35	b	a) Non è corretta. Sebbene le certificazioni per specifici LLM possano essere utili (e con maggiore disponibilità in futuro), i programmi di formazione dovrebbero essere progettati per garantire che i membri del team di test dispongano delle competenze tecniche necessarie per utilizzare gli strumenti GenAI in maniera efficace. b) È corretta. La selezione di LLM che sono compatibili con l'infrastruttura di test esistente e i requisiti di scalabilità sono un aspetto fondamentale di una strategia di GenAI. Gli strumenti di test e gli ambienti di test sono elementi fondamentali dell'infrastruttura di test. c) Non è corretta. La qualità dei dati e un input strutturato e sicuro sono fondamentali per ottenere risultati affidabili con la GenAI. È necessario disporre di una quantità adeguata di dati di input di alta qualità (in base all'obiettivo del test), non della maggior quantità possibile di dati. d) Non è corretta. La sezione 5.1.2 afferma che una strategia di GenAI per il testing del software deve raccogliere metriche specifiche per l'attività al fine di misurare l'efficacia degli output del LLM. Le metriche elencate nella sezione 2.3.1 (ad esempio, rilevanza, tasso di successo dell'esecuzione, efficienza temporale) sono adattate alle attività di test piuttosto che prese in prestito integralmente dal classico machine learning supervisionato.	GenAI-5.1.2	K2	1
----	---	---	-------------	----	---

36	b	a) Non è corretta. Un criterio principale di selezione consiste nel valutare le prestazioni del modello per le attività di test rispetto ai benchmark dell'organizzazione utilizzando metriche come quelle presentate nel syllabus stesso (vedere il syllabus nella sezione 5.1.3). Il criterio specificato può essere rilevante se le attività di test prevedono la generazione di codice. Tuttavia, non è applicabile ad altri tipi di attività di test. b) È corretta. Un criterio principale di selezione consiste nel valutare i costi ricorrenti (vedere il syllabus nella sezione 5.1.3) e questa risposta fa riferimento a un tipico esempio di costo ricorrente. c) Non è corretta. Valutare il LLM rispetto ai benchmark della community disponibili pubblicamente per garantire la piena compatibilità con essi non è uno dei criteri principali menzionati nel syllabus (nella sezione 5.1.3) per la selezione di un LLM appropriato per attività di test specifiche. d) Non è corretta. Un criterio principale di selezione consiste nel valutare i costi ricorrenti (vedere il syllabus nella sezione 5.1.3), ma questa risposta fa riferimento a un tipico esempio di costo non ricorrente.	GenAI-5.1.3	K2	1
----	---	--	-------------	----	---

37	a	Il syllabus descrive le seguenti 3 fasi principali: • Scoperta • Avvio e definizione dell'utilizzo • Utilizzo e iterazione Quindi: a) È corretta: menziona esplicitamente tutte e tre queste fasi principali. b) Non è corretta: non menziona esplicitamente nessuna di queste 3 fasi (menziona solo alcuni obiettivi e/o attività associati a queste fasi). c) Non è corretta: non menziona esplicitamente nessuna di queste 3 fasi (menziona solo alcuni obiettivi e/o attività associati a queste fasi). d) Non è corretta: non menziona esplicitamente nessuna di queste 3 fasi (menziona solo alcuni obiettivi e/o attività associati a queste fasi).	GenAI-5.1.4	K1	1
----	---	--	-------------	----	---

38	c	a) Non è corretta. Le allucinazioni nei LLM sono sfide intrinseche alle attuali tecnologie di AI e i tester non possono impedire il verificarsi di allucinazioni ed errori di ragionamento. I tester dovrebbero invece essere in grado di (identificare e) mitigare i rischi di allucinazioni ed errori di ragionamento (nonché i bias) durante il testing con la GenAI. b) Non è corretta. Sebbene la competenza nella test automation sia preziosa, non affronta specificamente l'integrazione della GenAI nei processi di test. c) È corretta. Il syllabus indica la "valutazione delle capacità degli LLM" come competenza principale. Ciò comporta naturalmente il confronto tra i modelli candidati e la selezione di quello che può essere adattato o perfezionato per particolari attività di test. Questo è un esempio diretto delle conoscenze/competenze di cui i tester hanno bisogno per lavorare in modo efficace con i LLM. d) Non è corretta. Si riferisce alle competenze richieste per lo sviluppo dei LLM, come a ricercatori di AI e data scientist, non ai tester che utilizzano la GenAI per attività di test. Questi tester si focalizzano sulla garanzia della qualità dei dati di test che utilizzano direttamente quando interagiscono con i LLM.	GenAI-5.2.1	K2	1
----	---	---	-------------	----	---

39	c	a) Non è corretta. Sebbene i corsi tenuti da esperti esterni con esercitazioni pratiche possano essere utili, affidarsi principalmente a essi sminuisce l'importanza della pratica interna e della creazione di una comunità. Inoltre, non è consigliabile integrare l'AI in tutte le attività di test quotidiane in modo improvviso. b) Non è corretta. Sebbene la sperimentazione sia parte del processo di apprendimento, una sperimentazione indipendente e senza struttura potrebbe non portare a uno sviluppo delle competenze coerente o efficace. c) È corretta. Focalizzarsi sullo sviluppo di competenze pratiche attraverso attività strutturate, peer learning e comunità di condivisione delle conoscenze è un approccio consigliato. d) Non è corretta. Affidarsi principalmente a corsi teorici tenuti da esperti esterni riduce l'importanza dello sviluppo di competenze pratiche e know-how attraverso la condivisione all'interno dell'organizzazione.	GenAI-5.2.2	K1	1
40	a	a) È corretta. I tester diventano specialisti del testing assistiti dall'AI, perfezionando i prompt e verificando gli output dell'AI. b) Non è corretta. Le responsabilità dei test manager vengono aggiornate per includere lo sviluppo di una strategia di test basata sull'AI, il risk management basato sull'AI e la supervisione dei processi di test basati sull'AI. Pertanto, la loro attenzione rimane focalizzata sul test management e non sulla comprensione del funzionamento interno (tecnicamente complesso) delle tecnologie di GenAI. c) Non è corretta. I tester non spostano la loro attenzione sulla supervisione dei processi di test basati sull'AI. Questa supervisione è responsabilità dei test manager. d) Non è corretta. I test manager devono bilanciare le capacità umane e quelle dell'AI per ottenere risultati efficienti.	GenAI-5.2.3	K1	1

Appendice: Risposte alle domande aggiuntive

Numero domanda (#)	Risposta corretta	Spiegazione / Motivazione	Obiettivo di apprendimento (LO)	Livello K	Numero di punti												
A1	N/A	Il processo RAG può essere riassunto come segue: 1. Suddivisione dei documenti in blocchi per un recupero mirato. 2. Pulizia ed elaborazione dei blocchi per garantire la coerenza. 3. Codifica dei blocchi in embedding ed archiviazione in un database vettoriale. 4. Recupero dei blocchi rilevanti durante l'elaborazione delle query degli utenti. 5. Generazione di risposte combinando i dati recuperati con le conoscenze del LLM. Pertanto, la risposta corretta è: <table><tr><th colspan="2">Risposta</th></tr><tr><td>Primo</td><td>Suddividere i documenti di grandi dimensioni in blocchi più piccoli</td></tr><tr><td></td><td>Pulire ed elaborare i blocchi di documento</td></tr><tr><td></td><td>Archiviare gli embedding in un database vettoriale</td></tr><tr><td></td><td>Recuperare i blocchi rilevanti in base alla similarità semantica</td></tr><tr><td>Ultimo</td><td>Generare una risposta utilizzando i blocchi recuperate e il LLM</td></tr></table>	Risposta		Primo	Suddividere i documenti di grandi dimensioni in blocchi più piccoli		Pulire ed elaborare i blocchi di documento		Archiviare gli embedding in un database vettoriale		Recuperare i blocchi rilevanti in base alla similarità semantica	Ultimo	Generare una risposta utilizzando i blocchi recuperate e il LLM	GenAI-4.1.2	K2	1
Risposta																	
Primo	Suddividere i documenti di grandi dimensioni in blocchi più piccoli																
	Pulire ed elaborare i blocchi di documento																
	Archiviare gli embedding in un database vettoriale																
	Recuperare i blocchi rilevanti in base alla similarità semantica																
Ultimo	Generare una risposta utilizzando i blocchi recuperate e il LLM																

A2	N/A	Considerando le definizioni: • Prompt Chaining - 1° elemento a destra: fa riferimento diretto all'uso iterativo del LLM per le attività di analisi dei test (Sezione 2.2.1). • Few-Shot Prompting - 2° elemento a destra: si collega esplicitamente ai test case generati da LLM con esempi (Sezione 2.2.2b). • Meta Prompting (C) - 4° elemento a destra: evidenzia la collaborazione tra LLM per il perfezionamento dei prompt (Sezione 2.1.2). • Zero-Shot Prompting (D) - 3° elemento a destra: enfatizza il ragionamento intrinseco del LLM per la prioritizzazione (Sezione 2.1.2). Quindi, la risposta corretta è: <pre> graph LR A[Prompt Chaining] --- B[Breaking down test analysis tasks into smaller steps requiring iterative LLM interactions and human verification] C[Few-Shot Prompting] --- D[Generating Gherkin-style test cases from user stories using the LLM with pre-defined examples] E[Meta Prompting] --- F[Prioritizing test cases based on risk by leveraging the LLM's inherent reasoning without examples] G[Zero-Shot Prompting] --- H[Interacting with the LLM to iteratively refine prompts for creating test oracles from ambiguous requirements] </pre>	GenAI-2.1.2	K2	1
-----------	-----	--	-------------	----	---

A3	N/A	Considerando le definizioni: • Scoperta – Le attività includono la formazione dei team di test sui concetti della GenAI, fornire l'accesso a LLM/SLM e la sperimentazione con use case iniziali per familiarizzare i tester con la GenAI e accrescere la loro sicurezza • Avvio e definizione dell'utilizzo – Si focalizza sull'identificazione e la prioritizzazione di use case pratici per la GenAI nel testing del software • Utilizzo e iterazione – È in atto un monitoraggio continuo dell'avanzamento della GenAI per il testing del software e degli strumenti correlati, nonché la misurazione e la gestione della trasformazione per garantire vantaggi sostenibili e scalabilità Quindi, la risposta corretta è: <table><tr><th>Activities</th><th>Key phases</th></tr><tr><td>Identifying and prioritizing practical use cases</td><td>Discovery</td></tr><tr><td>Training testers on Generative AI</td><td>Initiation and Usage Definition</td></tr><tr><td>Providing testers access to LLMs</td><td>Exploitation and Iteration</td></tr><tr><td>Experimenting with initial use cases to familiarize testers with Generative AI</td><td></td></tr><tr><td>Managing the evolution of test processes following the full integration of Generative AI</td><td></td></tr></table>	Activities	Key phases	Identifying and prioritizing practical use cases	Discovery	Training testers on Generative AI	Initiation and Usage Definition	Providing testers access to LLMs	Exploitation and Iteration	Experimenting with initial use cases to familiarize testers with Generative AI		Managing the evolution of test processes following the full integration of Generative AI		GenAI-5.1.4	K1	1
Activities	Key phases																
Identifying and prioritizing practical use cases	Discovery																
Training testers on Generative AI	Initiation and Usage Definition																
Providing testers access to LLMs	Exploitation and Iteration																
Experimenting with initial use cases to familiarize testers with Generative AI																	
Managing the evolution of test processes following the full integration of Generative AI																	