

Gli studenti devono rispondere a 1 dei 2 casi presentati. Di seguito sono riportate le indicazioni relative alla lunghezza del saggio e ai requisiti di formattazione. L'articolo allegato fornisce informazioni di base rilevanti. Gli studenti possono utilizzare qualsiasi risorsa fornita in classe o trovata online. Le risorse online devono essere basate su prove provenienti da fonti sottoposte a revisione paritaria.

Caso 1: Modelli mentali condivisi in un centro operativo di sicurezza umano-AI

Un centro operativo nazionale per la sicurezza informatica (SOC) ha implementato un sistema di supporto decisionale basato sull'intelligenza artificiale che correla i registri, assegna priorità agli avvisi e raccomanda azioni di risposta in caso di incidenti gravi. Dopo diversi attacchi gravi, le analisi dimostrano che gli analisti, i responsabili della gestione degli incidenti e il sistema di intelligenza artificiale spesso operano con una comprensione incompatibile di ciò che sta accadendo, delle priorità e di chi dovrebbe fare cosa. Alcuni analisti trattano i punteggi di rischio dell'IA come decisioni definitive, altri li ignorano in gran parte; i responsabili degli incidenti si lamentano di non poter vedere "ciò che l'IA pensa stia accadendo", mentre l'IA raccomanda spesso azioni che non corrispondono ai playbook esistenti o alle intenzioni del comandante. I debriefing rivelano ripetutamente sorprese dell'automazione, interpretazioni contrastanti degli stessi dati e confusione su obiettivi, ruoli e piani. Il management ritiene che il problema principale non sia solo la fiducia nell'automazione, ma anche il disallineamento dei modelli mentali condivisi e della consapevolezza condivisa della situazione all'interno del team umano-AI.

Basandoti sull'articolo allegato sulla consapevolezza della situazione e sui modelli mentali condivisi nei sistemi uomo-IA, analizza i malfunzionamenti del SOC, spiega in che modo le lacune nella consapevolezza della situazione individuale e di squadra contribuiscono a questi problemi e mostra come modelli mentali condivisi più allineati potrebbero migliorare le prestazioni. Proponi e giustifica cambiamenti concreti alla progettazione dell'interfaccia, ai ruoli e alle procedure del team, alla formazione o alle esercitazioni che potrebbero aiutare gli analisti umani e il sistema di IA a sviluppare, mantenere e aggiornare modelli mentali condivisi durante incidenti dinamici. Valuta criticamente i probabili vantaggi e limiti della tua proposta, comprese le possibili conseguenze indesiderate, e descrivi come valuteresti empiricamente se la consapevolezza condivisa della situazione e le prestazioni del team siano effettivamente migliorate.

Endsley, M. R. (2023). Supportare i team uomo-IA: trasparenza, spiegabilità e consapevolezza della situazione. *Computers in Human Behavior*, 140, 107574.

Caso 2:

Compito d'esame: Cultura organizzativa, salute mentale e minacce interne in un'azienda tecnologica

Un'azienda tecnologica che fornisce servizi cloud ha subito diversi incidenti di sicurezza che hanno coinvolto dipendenti con privilegi. In un caso, un amministratore di sistema ha copiato grandi quantità di dati di produzione su un dispositivo personale poco prima di dimettersi. Un'indagine interna non ha rilevato alcuna intenzione dolosa, ma ha rivelato alti livelli di stress, burnout e percezione di ingiustizia tra il personale tecnico. I sondaggi tra i dipendenti mostrano una scarsa sicurezza psicologica, un debole sostegno da parte della leadership alla salute mentale e la convinzione che "la sicurezza riguardi principalmente gli strumenti, non le persone". Allo stesso tempo, il team di sicurezza ha iniziato a considerare i dipendenti insoddisfatti principalmente come potenziali minacce interne, il che ha danneggiato ulteriormente la fiducia.

Il consiglio di amministrazione ha concluso che la cultura e le pratiche di leadership dell'organizzazione potrebbero aumentare sia il rischio psicosociale che quello di sicurezza. Lei è stato chiamato come consulente esterno per progettare una strategia completa e incentrata sull'uomo per la mitigazione delle minacce interne che non si basi esclusivamente sulla sorveglianza o sui controlli tecnici.

Il tuo compito è quello di proporre un programma integrato che affronti la cultura organizzativa, il comportamento della leadership, la salute mentale e le pratiche di sicurezza come elementi interconnessi. Utilizza teorie e ricerche empiriche pertinenti su minacce interne, stress lavorativo e burnout, giustizia organizzativa e clima di sicurezza. Descrivi interventi concreti a livello individuale, di team e organizzativo e valuta criticamente le tensioni etiche tra privacy e monitoraggio, spiegando come valuteresti l'efficacia e i potenziali effetti indesiderati del programma proposto nel tempo.

Khan, N., J. Houghton, R., & Sharples, S. (2022). Comprendere i fattori che influenzano le minacce interne involontarie: un quadro di riferimento per contrastare i rischi involontari. *Cognition, Technology & Work*, 24(3), 393-421.

Formattazione del saggio

Il saggio può contenere **3000 ± 10% parole** nel corpo principale (abstract e riferimenti non contano) in carattere Times New Roman 12. Il saggio deve essere digitato e scritto a doppia spaziatura su carta di formato standard (A4) con margini di 1" (standard in Word Document) su tutti i lati. Includere una pagina intestazione (nota anche come "intestazione corrente") nella parte superiore di ogni pagina. L'intestazione corrente è una

versione abbreviata del titolo del tuo elaborato e non può superare i 50 caratteri, spazi e punteggiatura inclusi. Il mancato rispetto di questo requisito può comportare un voto inferiore. Il conteggio delle parole non

le immagini senza testo. Le tabelle saranno conteggiate nel conteggio finale delle parole. Se nelle tabelle sono utilizzate immagini, queste saranno conteggiate come 500 parole. Il saggio dovrà avere una pagina del titolo con il titolo e il conteggio delle parole, un abstract, un corpo principale e dei riferimenti. Questo saggio è un testo accademico.

Per ulteriori informazioni, consultare <https://taltech.ee/en/formatting-guidelines>

Per questo saggio, deve essere utilizzato lo stile APA 7th (vedere: <https://apastyle.apa.org/style-grammar-guidelines/references/examples>)

Altre risorse per APA 7th (io personalmente uso questa): https://owl.purdue.edu/owl/research_and_citation/apa_style/apa_style_introduction.html

-

Valutazione

Vedere <https://taltech.ee/en/grading-system> per la politica di valutazione di Taltech

§ 14. Valutazione del rendimento accademico

(1) I metodi e i criteri di valutazione definiti nel programma devono essere comunicati agli studenti prima dell'inizio degli studi e non devono essere modificati durante il semestre. I metodi di valutazione definiscono le modalità di attestazione dell'acquisizione di conoscenze e competenze (ad esempio, esame orale o scritto, valutazione di idoneità/non idoneità, saggio, relazione, lavoro di gruppo, questionario). Se per la valutazione dei risultati dell'apprendimento vengono utilizzati diversi metodi, la loro peso relativo nella determinazione del voto finale deve essere specificato nel programma. Un criterio di valutazione deve specificare il livello e l'ambito di conoscenza attesi che possono essere dimostrati dai metodi di valutazione.

(3) In caso di valutazione graduata, il raggiungimento dei risultati di apprendimento viene valutato in base alla seguente scala:

A (5) – "eccellente" – raggiungimento eccezionale e particolarmente approfondito dei risultati di apprendimento, insieme a creatività e competenza consumata nell'applicazione delle abilità e delle conoscenze;

B (4) – "molto buono" – ottimo raggiungimento dei risultati di apprendimento, insieme a competenza nell'applicazione delle abilità e delle conoscenze in modo mirato e creativo. Alcuni dettagli delle conoscenze e delle abilità possono presentare errori che non sono né sostanziali né gravi;

C (3) – "buono" – buon raggiungimento dei risultati di apprendimento, insieme a competenza nell'applicazione delle abilità e delle conoscenze in modo pertinente. Una certa imprecisione e incertezza sono evidenti nella profondità e nei dettagli delle conoscenze e delle abilità;

D (2) – "soddisfacente" – raggiungimento sufficiente dei risultati di apprendimento, insieme all'applicazione delle conoscenze e delle competenze in modo tipico; in situazioni atipiche sono evidenti sia l'incertezza che la mancanza di conoscenze e competenze.

E (1) – "scarso" – raggiungimento minimo accettabile dei risultati di apprendimento più importanti, insieme a un'applicazione limitata delle conoscenze e delle competenze in situazioni tipiche; in situazioni atipiche situazioni atipiche sono evidenti sia una notevole incertezza che una mancanza di conoscenze e abilità; F

(0) – "bocciato" – raggiungimento di conoscenze e abilità al di sotto dello standard minimo.