SUGGERIMENTI DI AI POLICY

Informazioni da NON inserire mai in sistemi di AI pubbliche.

Queste categorie di informazioni non devono **MAI** essere inserite in prompt o caricamenti verso sistemi AI pubblici o non autorizzati:

- Segreti industriali e proprietà intellettuale non registrata
- Dati finanziari non ancora pubblicati (bilanci, previsioni, margini)
- Strategie competitive, piani di pricing, tattiche di mercato
- Informazioni su vulnerabilità tecniche o di sicurezza
- Dati personali identificabili (nomi, email, numeri di telefono, indirizzi)
- Dati sensibili ai sensi del GDPR (salute, origine etnica, orientamento politico, religioso, sessuale)
- Informazioni su fornitori, partner commerciali o beneficiari che potrebbero essere identificati
- Contratti o documenti legali contenenti termini riservati
- Credenziali di accesso, chiavi API, password o token

Esempi

Concessionaria automobilistica

SBAGLIATO:

Un responsabile vendite inserisce in ChatGPT:

“Analizza le performance del concessionario Milano Nord nel Q3 2024: 145 veicoli venduti, margine medio 8.3%, cliente top: Rossi Mario con 12 veicoli acquistati”

RISCHIO:

Esposizione di dati commerciali sensibili, performance specifiche di rete, identificazione di cliente top.

CORRETTO:

“Analizza questo scenario ipotetico: un punto vendita automotive ha venduto circa 150 veicoli in un trimestre con margini nella media di settore. Quali strategie potrebbero aumentare le vendite del 15% mantenendo la redditività?”

Esempi

Organizzazione non profit

SBAGLIATO:

Un operatore sociale chiede a ChatGPT:

“Scrivi una lettera di richiesta fondi raccontando la storia di Maria, 34 anni, rifugiata siriana ospitata nel nostro centro di Roma, con due figli minori e diagnosi di disturbo post-traumatico.”

RISCHIO:

Violazione della privacy, esposizione di dati sensibili (salute, status legale, origine etnica), identificabilità della persona.

CORRETTO:

“Scrivi una lettera emotivamente coinvolgente per una campagna di raccolta fondi che supporta l'integrazione di nuclei familiari in condizioni di vulnerabilità, senza identificare persone specifiche. Focus su impatto collettivo del progetto.”

Esempi

Studio di Architettura e Progettazione

SBAGLIATO:

Un architetto carica su un'IA generativa i file CAD o il computo metrico di un nuovo progetto non ancora approvato:

“Analizza questo computo metrico per la ristrutturazione della Villa [Nome Cliente] a [Località] e suggerisci dove possiamo tagliare i costi del 10% mantenendo questi standard di materiali.”

RISCHIO:

Violazione del segreto professionale e del diritto d'autore. I disegni tecnici e i dati economici della proprietà finiscono nel database pubblico, diventando potenzialmente accessibili a terzi o utilizzati per addestrare modelli concorrenti.

CORRETTO:

“Agisci come esperto di gestione costi nell'edilizia di lusso. Ho una lista di materiali (marmo di Carrara, infissi in triplo vetro, domotica avanzata) con i relativi prezzi unitari medi di mercato. Suggerisci soluzioni alternative o materiali ecosostenibili che garantiscano un risparmio del 10% senza compromettere l'estetica.”

Esempi

Scuola Pubblica e Istituti Scolastici

SBAGLIATO:

Un docente inserisce in una chat pubblica i dati per generare un piano didattico personalizzato (PDP):

“Scrivi un piano di studio per un alunno della classe 3ªC, con diagnosi di dislessia e difficoltà nell'attenzione, che nell'ultimo compito di storia ha preso 4.”

RISCHIO:

Esposizione di dati sensibilissimi (salute e rendimento di un minore). Questi dati diventano "conoscenza pubblica" e violano i protocolli di sicurezza scolastica e il GDPR.

CORRETTO:

“Crea un'attività didattica di storia sulla Rivoluzione Industriale adatta a uno studente di scuola media con dislessia compensata. Utilizza mappe concettuali, font ad alta leggibilità e suggerisci esercizi basati sull'audio o su immagini per facilitare l'apprendimento inclusivo.”

Il falso senso di sicurezza: "E se non metto il nome?"

Molti credono che basti eliminare il nome proprio (es. chiamarlo "lo studente X") per essere al sicuro. In realtà, nel contesto di una scuola o di un ufficio, questo spesso non è sufficiente per tre motivi critici:

1. Re-identificazione per contesto: Se scrivi "lo studente della 3ªC con dislessia che ha preso 4 nel compito di storia del 12 gennaio", quell'individuo è perfettamente identificabile da chiunque conosca la classe. L'IA aggrega questi dettagli e, se incrociati con altri dati pubblici o violazioni future, contribuiscono a creare un profilo digitale sensibile della persona senza il suo consenso.

2. Dati Sensibili (Categorie Particolari): Descrivere diagnosi cliniche, difficoltà di apprendimento o situazioni familiari disagiate significa comunque immettere "Dati Sensibili" (secondo il GDPR e
l'AI Act) in sistemi pubblici. Questi sistemi usano le informazioni per addestrarsi, rendendo quel caso clinico o umano parte di un database commerciale globale.

3. Proprietà Intellettuale e Metodo: Anche senza nomi, stai regalando a un'IA esterna il tuo "know-how" pedagogico o professionale. Stai insegnando alla macchina come risolvere quel problema specifico basandoti sulla tua esperienza, perdendo il controllo esclusivo sul tuo metodo di lavoro.

L'approccio corretto (Inversione del flusso)

Invece di dare all'IA i dati del problema per avere la soluzione, bisogna fornire il modello teorico per generare strumenti neutri:

SBAGLIATO:

"Aiutami a gestire questo caso: studente dislessico che non riesce a studiare Napoleone..." (L'IA impara il caso reale)."

CORRETTO:

"Agisci come esperto in didattica inclusiva. Quali sono le migliori strategie per spiegare l'epoca napoleonica a studenti con DSA? Prepara uno schema standard basato sulle linee guida ministeriali." (L'IA ti fornisce lo strumento, che tu applicherai privatamente al tuo caso).

Ricordiamo sempre che, anche senza dati sensibili, le impostazioni di default delle AI pubbliche possono:

- **memorizzare** le tue conversazioni per migliorare i modelli
- **conservare** cronologia accessibile da altri dispositivi
- **addestrare** l'AI su esempi del tuo settore/stile.

Impariamo dunque a configurare correttamente gli LLM che usiamo, disattivando la memorizzazione dei dati per impedire l'addestramento dei modelli.

Ricorda sempre che la responsabilità dell'utente non si esaurisce con la protezione dell'input: è altrettanto vitale prestare la massima attenzione al cosiddetto '**Fattore Output**'. Ovvero, verificare sempre criticamente ciò che l'IA produce, a causa del rischio di 'allucinazioni' o pregiudizi (bias), specialmente in ambiti delicati come quello scolastico o legale.

Speriamo che questa guida ti aiuti a vivere l'innovazione con più serenità, evitando i rischi legati al "falso senso di sicurezza" o a impostazioni di default poco trasparenti.

Ti va di parlarne?

Sappiamo che ogni realtà è diversa e che una policy generale non può rispondere a ogni dubbio specifico. Se vuoi capire come:

- Rendere sicuri i flussi di lavoro del tuo team;

- Configurare correttamente gli strumenti per impedire l'addestramento dei modelli sui tuoi dati;

- Sfruttare l'IA nella comunicazione video e web senza compromettere il tuo know-how.

Siamo qui per un confronto sincero. Non cerchiamo di venderti una soluzione standard, ma di capire se e come la nostra esperienza può supportare la tua visione.

Facciamo due chiacchiere?

Paola Furlan

Info@rivoluzione.online