

**I QUADERNI
DEL CONSUMATORE**

Le truffe nell'era dell'AI



PREMESSA

Le truffe cambiano forma, ma continuano a fare leva sulla fiducia e sulla fretta con cui spesso prendiamo una decisione. L'intelligenza artificiale non ha inventato questi meccanismi, però può renderli più difficili da riconoscere, perché permette di costruire messaggi credibili e adattati alla persona che li riceve.

Quando un tentativo di inganno utilizza informazioni reali e un linguaggio naturale, gli indizi ai quali eravamo abituati possono non bastare. Il rischio riguarda i consumatori e raggiunge anche i luoghi di lavoro, dove una richiesta falsa può sembrare provenire da una persona conosciuta o da un'organizzazione affidabile.

La responsabilità della sicurezza non può ricadere soltanto sull'attenzione individuale. Prendersi il tempo di verificare resta necessario, ma servono anche sistemi capaci di riconoscere comportamenti sospetti e protezioni aggiornate rispetto a minacce che cambiano rapidamente.

Per Adiconsum, conoscere il modo in cui vengono costruite queste frodi è parte della tutela del consumatore. Comprendere i meccanismi dell'inganno aiuta a valutare con maggiore cautela una richiesta e a non confondere la qualità formale di un messaggio con la sua attendibilità.

Con questo quaderno vogliamo offrire strumenti accessibili per affrontare un rischio che diventa più sofisticato senza per questo essere inevitabile. La tecnologia impiegata per ingannare richiede una tutela capace di evolvere con la stessa rapidità.

Buona lettura.



Quando l'AI viene usata per truffare

L'intelligenza artificiale ha grandi potenzialità, che richiedono però una guida rigorosa per essere indirizzate verso **usi utili e responsabili**. Il problema nasce quando lo strumento viene impiegato consapevolmente a **scopo fraudolento**. In questi casi **l'AI può rendere più credibili i tentativi di truffa**, automatizzare attività che prima richiedevano tempo e competenze specifiche e adattare messaggi e contenuti alle informazioni disponibili sulla vittima.

I rischi riguardano la sicurezza dei cittadini, delle aziende e perfino dei sistemi che proteggono le nostre abitazioni. Per contrastarli non basta affidarsi all'attenzione individuale. Occorrono strumenti di difesa capaci di individuare contenuti sospetti e vulnerabilità prima che vengano sfruttati, anche ricorrendo alla stessa intelligenza artificiale per riconoscere e bloccare le minacce.

Ingegneria sociale e truffe personalizzate

L'ingegneria sociale **sfrutta meccanismi psicologici e suggestioni** per orientare le azioni della vittima.

Non riguarda soltanto i cittadini, ma coinvolge spesso anche aziende e istituzioni. È una delle tecniche di truffa più versatili ed efficaci e può essere utilizzata per ottenere informazioni, denaro o persino creare una forma di dipendenza psicologica.

Dai santoni ai finti impiegati di banca, fino alle frodi più complesse ai danni di società ed enti, ricorre quasi sempre qualche elemento di **ingegneria sociale**, applicato con maggiore o minore efficacia.

Nella storia troviamo esempi di attacchi basati sull'ingegneria sociale sia nella **Bibbia che nei grandi classici come l'Iliade e l'Odissea**.

Ma che cosa ha a che fare una pratica così antica con l'**intelligenza artificiale**?

Le tendenze più recenti emerse nel dark web indicano che l'intelligenza artificiale sta assumendo un ruolo sempre più rilevante nell'ingegneria sociale. In particolare, sono gli agenti di intelligenza artificiale a rappresentare il rischio più grave¹.

In pratica l'intelligenza artificiale agentica gestisce (quasi totalmente) **la truffa da sola**: prepara i testi, invia le email, crea siti di phishing, ma **il vero problema è che l'attacco non si limita a riproporre sempre le solite strategie**, ma **si evolve continuamente sulla base dei feedback che riceve**.

¹ Un agente AI è un sistema capace di svolgere attività per raggiungere un obiettivo con un certo grado di autonomia.

phishing

Oltre al normale **phishing**, l'AI agentica è in grado di realizzare attacchi più specifici, conosciuti come **spear phishing²** e **whaling³**, automatizzando la raccolta preliminare di informazioni attraverso **l'OSINT - Open Source INTElligence**.

In poco tempo l'intelligenza artificiale rintraccia **contatti, numeri di telefono, opinioni espresse sui social** ed altre informazioni utili a rendere la truffa più credibile ed efficace.

Fino a pochi anni fa lo spear phishing era meno frequente, perché, se realizzato da una persona, richiedeva tempo per **rac cogliere i dati, metterli in relazione e costruire un messaggio su misura**. L'AI può ridurre notevolmente questi tempi, favorendo un aumento degli attacchi mirati, che potrebbero non colpire più soltanto i tradizionali **"bersagli di interesse"**, ma un numero molto più ampio di persone.

Questo cambia radicalmente le carte in tavola: dobbiamo prepararci a **truffe più sofisticate e complesse**. Solitamente il suggerimento più gettonato è quello di **tenere alto il livello di attenzione**; una pratica scontata, ma non così facile da realizzare proprio per una **caratteristica propria degli esseri umani**.

2 Attacco di phishing mirato nel quale l'email o l'SMS di truffa non sono generici, ma costruiti sulla base di informazioni riferite a una vittima specifica. Questa pratica richiede una raccolta preliminare di dati sulla persona.

3 Forma di spear phishing rivolta ai dirigenti di un'azienda.



Perché abbassiamo la guardia

Ci sono azioni che compiamo nella vita di tutti i giorni **in maniera automatica**, quasi senza pensarci e con poca attenzione e altre a cui dedichiamo **un'attenzione meticolosa**, quasi paranoica. Per capire meglio questa doppia modalità nel comportamento umano, possiamo pensare ad un esempio formulato dal premio nobel **Richard H. Thaler** nel contesto dell'economia comportamentale, ma che si adatta benissimo anche alla **psicologia delle truffe**.

Immaginiamo che nella nostra mente convivano due personaggi: **Homer Simpson** e il signor **Spock di Star Trek**. Quando il primo è al comando prendiamo **decisioni istintive** in modo automatico quasi senza pensarci. Quando prevale il secondo, **ogni scelta viene valutata con maggiore attenzione**.

A seconda del nostro carattere e delle situazioni possiamo essere più simili a Homer Simpson in alcune circostanze e più vicini a Spock in altre. **È proprio quando prevale la reazione automatica che diventiamo più vulnerabili alle truffe.**

Il problema è che chi realizza truffe farà di tutto per fare emergere l'Homer Simpson presente in ciascuno di noi. Quanto più la truffa è sofisticata, tanto più può diventare difficile riconoscerla.

Per questo, oltre al **senso critico**, abbiamo bisogno di un **supporto** che può provenire dalla tecnologia stessa e **agevolare il passaggio dalla reazione istintiva all'azione razionale.**

Vulnerabilità informatiche e cybersicurezza

Un ulteriore tassello che si aggiunge allo scenario di utilizzo mai levolo dell'AI è proprio quello relativo alla scoperta di vulnerabilità informatiche.

Individuare una vulnerabilità non è qualcosa di semplice, anzi richiede spesso l'intervento di hacker esperti o una circostanza favorevole.

Nel campo dell'AI, però, tutto cambia rapidamente. Nell'aprile 2026 il modello **Mythos di Anthropic**⁴ ha dimostrato di poter individuare e sfruttare vulnerabilità molto **più velocemente di un team umano**.

Mythos si è mostrato così potente da poter **mettere in ginocchio anche la sicurezza degli istituti di credito**, tanto che la BCE stessa ha convocato le banche europee per raccogliere riflessioni su come affrontare i possibili rischi per la sicurezza collegati allo sviluppo dell'intelligenza artificiale⁵.

Mythos **è soltanto una delle numerose intelligenze artificiali specializzate** nella cybersicurezza e destinate a cambiare profondamente il settore.

Questi strumenti possono offrire **garanzie elevate di sicurezza**, ma anche **rappresentare un grave rischio**, a seconda di chi li utilizza e delle finalità per cui vengono impiegati. Sono irrinunciabili se vogliamo contrastare efficacemente il cybercrimine, ma richiedono altrettanta attenzione nel loro sviluppo e nel loro

4 Anthropic è una benefit corporation statunitense di intelligenza artificiale, particolarmente specializzata su tematiche di rischio ed etica.

5 Fonte: <https://teleborsa.ansa.it/notiziario/finanza/bce-convoca-le-banche-per-approfondire-rischi-claude-mythos/>



utilizzo. **Non possiamo permetterci di restare impreparati**, perché minacce che oggi possono apparire ancora teoriche o circoscritte potrebbero presto diventare concrete, **mettendo in seria difficoltà chi non avrà gli strumenti necessari per fronteggiarle**.

È particolarmente interessante quanto emerge dal report realizzato da **IBM X-Force**⁶:

Nel **2025** il numero di **vulnerabilità monitorate da X-Force** si è avvicinato a **40.000** e oltre la metà poteva essere sfruttata **senza autenticazione**.⁷

6 IBM X-Force è il gruppo di IBM specializzato nell'analisi e nel contrasto delle minacce informatiche.

7 Fonte: X-Force Threat Intelligence Index 2026

IL PROGETTO

Il progetto **“DICO Sì. Diritti, Consapevolezza, Opportunità, Strumenti, Inclusione (III edizione)”**, promosso da Adiconsum e finanziato dal **Ministero delle Imprese e del Made in Italy** nell’ambito delle **Iniziative a vantaggio dei consumatori**, ai sensi dell’articolo 148, comma 1, della legge 23 dicembre 2000, n. 388 e del Decreto direttoriale 12 maggio 2025, nasce per aiutare i cittadini-consumatori a orientarsi nei cambiamenti prodotti dalla digitalizzazione e dall’intelligenza artificiale, che incidono ormai sulle scelte di acquisto e sull’accesso ai servizi.

La terza edizione dedica particolare attenzione al divario digitale legato all’età e al genere, con attività rivolte soprattutto alle persone over 60 e alle donne. Il progetto intende:

- rendere più comprensibili le nuove tecnologie e il modo in cui influenzano la vita quotidiana;
- accompagnare i consumatori verso un uso del digitale più consapevole, prestando attenzione alla sicurezza e alla tutela dei propri diritti;
- diffondere una cultura della sostenibilità digitale, attenta agli effetti dell’innovazione sulle persone e sull’ambiente;
- favorire il confronto tra cittadini, esperti, istituzioni, imprese e mondo della ricerca.

Il progetto prevede un’indagine nazionale sull’impatto dell’intelligenza artificiale sui consumi, cinque incontri formativi online e materiali pratici dedicati all’uso sicuro delle tecnologie. È previsto anche un gioco interattivo per conoscere il funzionamento degli algoritmi, mentre due forum e una rete di sportelli territoriali offrono occasioni di confronto e assistenza diretta.

Per maggiori informazioni visita il sito www.adiconsum.it



III EDIZIONE

Finanziato dal MIMIT - D.D. 12 maggio 2025

*Facciamo
crescere
le tue tutele
di consumatore!*

Firma il tuo **5xmille** ad Adiconsum
C.F. 96107650580

segui su



www.adiconsum.it



ADICONSUM

Associazione Difesa Consumatori APS

dal 1987