Intelligenza artificiale, boom di truffe anche in Europa

Hi tech. Crescono del 780% le nuove frodi attuate con intelligenza artificiale
Allarme di Bankitalia. Deloitte: «Solo negli Usa attesi danni per 40 miliardi»

Biagio Simonetta
MILANO

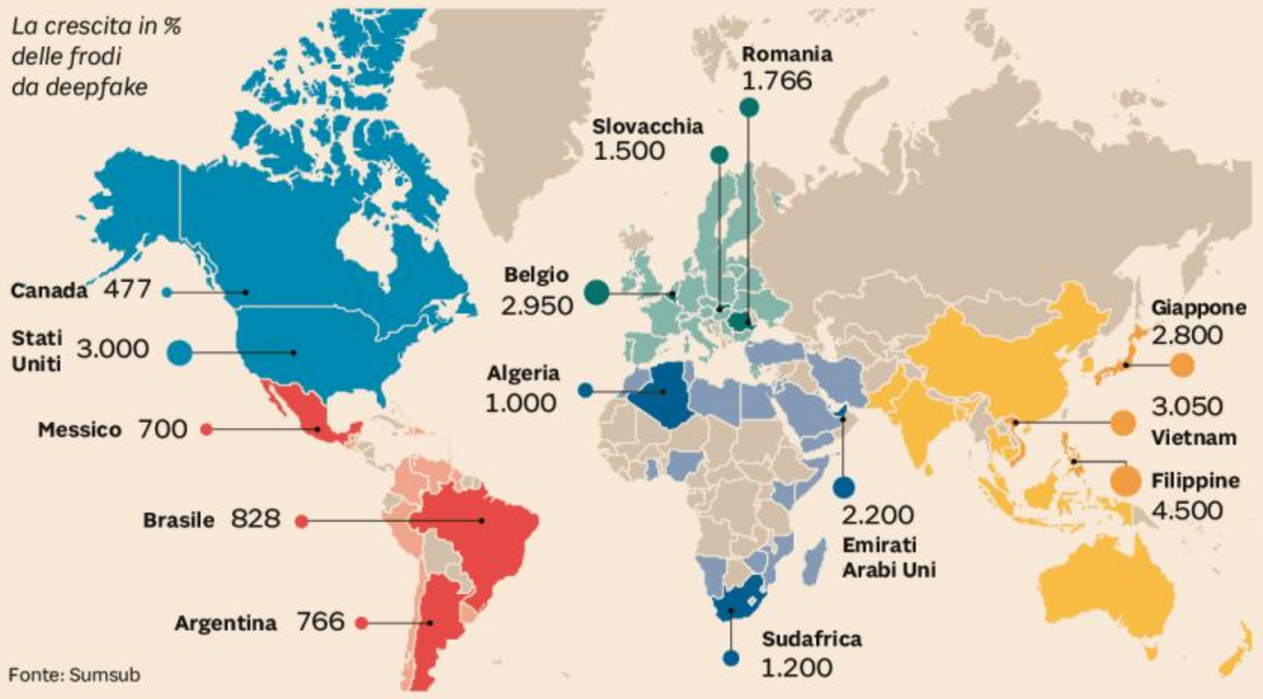
L'ultimo allarme lo ha lanciato Banca d'Italia qualche giorno fa, denunciando la «presenza in rete di videomessaggi che, in maniera artificiosa, riproducono l'immagine e la voce di esponenti di autorità competenti in materia finanziaria, tra cui la Banca d'Italia, e di altri vertici istituzionali e personalità note. Nessuno di questi video, anche nei casi in cui sono presenti espliciti riferimenti all'Istituto, è stato autorizzato dalla Banca d'Italia».

Un messaggio chiarissimo, che spiega rapidamente quale sia la nuova frontiera del crimine informatico che prende di mira la finanza: il deepfake, cioè la generazione - grazie all'intelligenza artificiale - di contenuti audio o video assolutamente falsi, ma anche assolutamente credibili. Contenuti generati con uno scopo ben preciso: utilizzare voce e volto di personalità della finanza per ingannare qualcuno, veicolando messaggi falsi finalizzati a una truffa.

Roba da hacker professionisti, starete pensando. E invece non è proprio così. Perché il boom dell'intelligenza artificiale generativa ha rimescolato le carte, riducendo la necessità di competenze. E oggi, grazie a nuovi tool che lavorano con la GenAI, più o meno tutti possono creare immagini apparentemente reali e far dire o fare qualsiasi cosa a personaggi più o meno noti.

Gli sviluppi di società come OpenAI, infatti, non hanno partorito solo chatbot intelligenti capaci di rispondere testualmente a qualsiasi nostra domanda, ma anche software capaci di creare una realtà che non esiste. Un trucco digitale raffinatissimo, che (come vedremo più avanti) solo la stessa intelligenza artificiale può aiutarci a svelare.

La mappa degli attacchi



Fonte: Sumsub

> **Ad aprile, anche i sistemi di sicurezza della JP Morgan sono stati ingannati da una voce clonata**

Il bonifico di Hong Kong

Il caso più clamoroso lo abbiamo raccontato proprio sul Sole 24 Ore quando a febbraio un dipendente di una società finanziaria britannica, con sede ad Hong Kong, è stato convinto - da un video deepfake nel quale compariva il suo *chief financial officer* - ad effettuare una serie di bonifici per una somma vicina ai 25 milioni di dollari. Oggi, a distanza di quasi quattro mesi da quel caso che fece molto rumore, sono emersi ulteriori particolari. Ad esempio che la società coinvolta si chiama Arup, multinazionale londinese con un fatturato da 2 miliardi di sterline. E che il dipendente caduto in trappola è un operatore finanziario che aveva partecipato a una videochiamata con persone ritenute il direttore finanziario dell'azienda e altri membri dello staff. Persone che gli avevano chiesto di effettuare un trasferimento di denaro. Ma il resto dei partecipanti a quella videochiamata, in realtà, erano dei deepfake. E circa 25 milioni di dollari sono stati trafugati.

Un'ondata crescente

Il punto è che il caso di Hong Kong è tutt'altro che isolato. Perché un'ondata crescente di truffe simili sta saccheggiando milioni di dollari alle aziende di tutto il mondo. E a giudicare dagli sviluppi, c'è poco da stare tranquilli, considerando che i criminali informatici oggi sfruttano l'intelligenza artificiale generativa.

A marzo scorso, i media statali cinesi, hanno raccontato di un caso simile a quello di Hong Kong nella provincia dello Shanxi, con una impiegata finanziaria indotta con l'inganno a trasferire 1,86 milioni di yuan (262mila dollari) sul conto di un truffatore dopo una videochiamata con un deepfake del suo capo.

Ad aprile, i sistemi di sicurezza della JP Morgan Chase Bank sono stati ingannati da una voce clonata con l'intelligenza artificiale (anche se per fortuna in questo caso si trattava di un esperimento di una giornalista del Wall Street Journal).

La situazione, insomma, rischia di finire presto fuori controllo. Ammesso che non lo sia già.

Il Center for Financial Services di Deloitte prevede che i contenuti falsi, creati con l'intelligenza artificiale generativa, daranno un nuovo impulso alle truffe finanziarie, che potrebbero raggiungere i 40 miliardi di dollari solo negli Stati Uniti entro il 2027, rispetto ai 12,3 miliardi di dollari del 2023, con un tasso di crescita annuo composto del 32%.

Sempre negli Stati Uniti, l'Internet Crime Complaint Center dell'FBI ha ricevuto nel 2023 più di 880mila denunce di tentativi di truffa basati sull'intelligenza artificiale, in aumento del 22% rispetto all'anno precedente, con perdite potenziali superiori a 12,5 miliardi di dollari.

Secondo il report annuale di Sumsub (un fornitore di verifica dell'identità), in Europa gli attacchi informatici basati sul deepfake sono cresciuti del 780% nel 2023, a conferma di come l'arrivo dell'AI generativa abbia dato un'accelerata importante a questo particolare settore. In questo scenario, il 6,8% di tentativi di truffa generati con l'AI in Europa lo scorso anno, sono stati localizzati in Italia.

A livello globale, la stessa società ha stimato che nel 2023 gli attacchi deepfake nel solo settore fintech sono aumentati del 700% rispetto all'anno precedente. Un autentico boom.

Banche in allerta

Gli istituti finanziari sono in una situazione di massima allerta, come conferma il comunicato diffuso dalla Banca d'Italia. Del resto, le frodi guidate dall'intelligenza artificiale non solo incidono sui loro bilanci, ma minano anche la fiducia dei clienti e l'integrità del sistema finanziario. E anche se da sempre le banche sono in prima linea nell'utilizzo di tecnologie innovative per combattere le frodi, un recente rapporto del Tesoro americano ha rilevato che «i quadri di gestione del rischio esistenti potrebbero non essere adeguati per coprire le tecnologie emergenti dell'intelligenza artificiale».

E allora qual è l'antidoto alle nuove truffe create con l'intelligenza artificiale? La risposta degli esperti sembra unanime: l'intelligenza artificiale stessa. Per questo alcune banche stanno già incorporando i *large language model (LLM)* per rilevare segnali di frode, come quello utilizzato da JPMorgan per la validazione delle email. Altre banche utilizzano l'intelligenza artificiale per automatizzare i processi che diagnosticano i tentativi di furto. In Europa, HSBC ha implementato strumenti basati sull'intelligenza artificiale per rilevare e prevenire le frodi nei pagamenti. Il suo sistema analizza milioni di transazioni, identificando modelli e segnalando anomalie che potrebbero suggerire attività fraudolente. In questo modo la banca prova a prevenire transazioni non autorizzate prima che avvengano, salvaguardando il patrimonio dei clienti e la sua reputazione.

Il capitolo deepfake, tuttavia, rimane una ferita aperta, che crea molta preoccupazione e che ha già tratto in inganno. I big del settore AI si stanno concentrando nello sviluppo di tecnologie capaci di svelare il trucco in tempo reale. Lo stesso trucco che i loro sistemi hanno aiutato a generare.

I NUMERI

40

Miliardi di dollari di danni
Il Center for Financial Services di Deloitte prevede che i contenuti falsi, creati con l'intelligenza artificiale generativa, daranno un nuovo impulso alle truffe finanziarie, che potrebbero raggiungere i 40 miliardi di dollari solo negli Stati Uniti entro il 2027

880mila

Denunce negli Usa
L'Internet Crime Complaint Center dell'FBI ha ricevuto nel 2023 più di 880mila denunce di tentativi di truffa basati sull'intelligenza artificiale