Vibe coding, agenti IA e cloud: stiamo costruendo un sistema troppo veloce per essere sicuro?

Il mondo di Internet sta cambiando rapidamente. Le quarantene pandemiche del 2020 hanno **accelerato in modo significativo il processo di digitalizzazione** di aziende, governi e istituzioni. In questo contesto di fertilità nei confronti delle nuove tecnologie sono esplose le **intelligenze artificiali generative**, le quali hanno spinto le imprese a sostenere investimenti ingenti di cui oggi, però, gli investitori chiedono conto. La rapida e talvolta goffa assimilazione delle nuove tecnologie informatiche, le aspettative ambiziose del mercato e la pressione a ottenere risultati nel breve periodo stanno alimentando quello che Cisco, multinazionale specializzata in tecnologie della comunicazione e sicurezza informatica, definisce nel suo ultimo report sulla preparazione all'IA come **"debito infrastrutturale"**: la tendenza a privilegiare compromessi immediati a scapito del consolidamento di elementi chiave come l'ammodernamento di software e hardware. Questo approccio non solo riduce le probabilità che le applicazioni sviluppate producano risultati soddisfacenti, ma mina anche in modo significativo la **sicurezza informatica** delle organizzazioni che le adottano.

Secondo quanto dichiarato dall'impresa, **solo il 35% delle organizzazioni italiane si dichiara pienamente consapevole delle minacce informatiche** legate alle intelligenze artificiali, e appena il 32% dispone di strumenti e competenze adeguati per rendere sicuri gli agenti di IA, ovvero quegli strumenti che promettono di rivoluzionare la filiera produttiva, semplificando e automatizzando funzioni articolate. Le criticità della sfera digitale sono inoltre amplificate da equilibri geopolitici in rapido mutamento che trasformano alleati in potenziali avversari e rendono le grandi aziende tecnologiche protagoniste capricciose di un ecosistema sempre più ancorato ai loro servizi. Tutti questi fattori convergono in una fase storica che si apre su un **futuro altamente incerto**, in cui esperti di diversa estrazione delineano scenari ritenuti plausibili ma spesso divergenti tra loro.

Per comprendere davvero le complessità tecnologiche che le infrastrutture digitali devono affrontare oggi, non basta osservare applicazioni e servizi visibili all'utente finale. È necessario andare a monte, dove reti, sicurezza e architetture IT sostengono l'intero sistema finanziario e istituzionale. Per questo abbiamo scelto di confrontarci con **Fabio Panada**, Security Architect della sopracitata Cisco, con l'obiettivo di analizzare, con spirito critico, lo stato di salute digitale di imprese e istituzioni in un contesto in cui sicurezza, resilienza e stabilità delle reti non sono più solo temi tecnici, ma **fattori che incidono anche sugli equilibri economici e geopolitici**.

La definizione di cybercrimine è soggetta a visioni politiche e interessi nazionali. Ci sono governi che classificano l'hacktivismo come cyberterrorismo, mentre altri

preferiscono attribuire a gruppi indipendenti operazioni che in realtà hanno una matrice statale. A questo si aggiungono le attività governative che violano apertamente norme internazionali o legislazioni interne. In un panorama così variegato, quale quadra viene adottata da Cisco per stabilire cosa rientri nella categoria di "minaccia informatica"?

Fabio Panada: Oggi il mondo è veramente complesso: gli attacchi sono variegati e diversi. Quello che vediamo noi è spesso cyberterrorismo, attivisti, minacce provenienti da gruppi spesso sponsorizzati da governi statali. Il mondo va sicuramente in una direzione di complessità multidisciplinare, è spesso difficile capire chi c'è dietro, da questo punto di vista.

Come Talos [la divisione dedicata alla sicurezza informatica di Cisco, n.d.r.], noi teniamo traccia di vari tipi di attacchi; creiamo dei report trimestrali dove riportiamo ciò che vediamo. Più che per quanto riguarda la tipologia di attacchi, ci focalizziamo molto sulle tecnologie e sui settori che sono attaccati. Una cosa che vediamo ultimamente, oltre ai classici attacchi DDoS [il sovraccarico di sistemi informatici, n.d.r.] e ransomware [il sequestro di dati e archivi, n.d.r.], è la focalizzazione su attacchi legati all'identità. Il furto dell'identità è un po' il nuovo trend.

Qual è la proporzione degli attacchi legati a vulnerabilità digitali rispetto alle campagne di social engineering, ovvero all'ottenimento di informazioni personali attraverso stratagemmi e inganni?

Oggi la maggior parte [degli attacchi] sono quelli che potremmo chiamare di *social engineering*, anche perché il termine descrive tante tipologie di attività. Stimiamo che **il 60-65% degli attacchi** sono legati al furto di identità. Se non altro perché, purtroppo, questi attacchi sono molto più semplici. C'è da dire che le aziende, anche grazie alle leggi e alle direttive che abbiamo, si sono strutturate per essere un po' più protette dal punto di vista tecnologico. È difficile trovare delle vulnerabilità, le vulnerabilità tecnologiche sono spesso corrette in tempo, gli aggiornamenti vengono fatti impiegando dei processi… quindi per un attaccante **è più facile sfruttare la debolezza umana** e capire l'identità sfruttando la fiducia ottenuta dall'utente finale, che magari trovare una falla in un firewall. Da questo punto di vista, è ancora l'umano, oggi, l'anello un po' più debole.

Gli attacchi di questo tipo sono d'altronde molto sofisticati ed è quindi comprensibile cascarci quando dall'altra parte ti arriva, magari, una mail che sembra del tutto legittima… anche perché, in realtà, lo è! Perché un altro dei problemi che noi vediamo è che, una volta che una identità è stata carpita, questa diventa un'identità lecita. Vuol dire che anche gli

strumenti che sono in atto per cercare di rilevare qualcosa di pericoloso non riconoscono quell'insidia, perché, di fatto, riconoscono il recapito come lecito. Quindi, sono necessari altri strumenti, anche dal punto di vista difensivo, più adattativi, che analizzino non solo l'utente, che è la cosa importante – perché l'utente oggi è la cosa più importante –, ma il comportamento dello stesso. Quindi, il capire se una mail, per esempio, è diversa da tutte le mail che ha inviato in precedenza.

Però mi pare che, nonostante questo problema, si sia consolidata una progressiva desensibilizzazione sul tema, che molte persone partano ormai dal presupposto che le fughe di dati siano normali e che non ci si possa fare niente. Che sia un naturale sottoprodotto del sistema.

Il furto di dati succede molto spesso, ma ciò è vero proprio perché il dato è oggi fondamentale. È una ricchezza. Noi che lavoriamo con varie tipologie di aziende, ci rendiamo conto di come le stesse siano consapevoli del suo valore e che oggi avere un furto di informazioni o di dati è un qualcosa di impattante anche sul business dell'azienda. Da un lato, è vero che alcune persone iniziano ad abituarsi e danno quasi per scontato che possano verificarsi furti di dati. Dall'altro, però, cresce la consapevolezza dell'importanza della protezione dei dati personali. Questo passaggio evidenzia come, pur aumentando l'assuefazione agli incidenti, si rafforzi anche la sensibilizzazione verso il valore e la tutela dei dati.

In realtà, è un problema che vediamo e che vedremo sempre più frequentemente. Le modalità attraverso cui oggi utilizziamo le tecnologie si sono evolute nel tempo e diventano sempre più complesse. Perché oggi tutti lavoriamo da casa… una volta c'era il perimetro [aziendale], oggi è invece importante proteggere il "mio" laptop. Inoltre, oggi molte applicazioni e molti dati sono in cloud, o in cloud ibrido, e magari non sono più in Italia. Anche in questo senso, la sensibilizzazione è in aumento.

A proposito di cloud: molte aziende stanno esternalizzando la gestione dei server puntando su questa tecnologia con l'idea che affidarsi alle Big Tech sia più sicuro. Altri sostengono invece che non ci si possa fidare dei fornitori esteri e che le infrastrutture critiche dovrebbero essere staccate del tutto dalla rete, così da impedire a priori ogni forma di hacking in remoto. Sono due prospettive completamente opposte, quindi come ci si orienta?

Io non sono, purtroppo, giovanissimo e vedo queste tendenze andare e tornare. Ho iniziato a lavorare quando c'erano i *mainframe* [i vecchi computer ad alte prestazioni che occupavano intere stanze, n.d.r.], quindi era tutto centralizzato. Poi è diventato tutto distribuito, poi è

ritornato tutto centralizzato; non è detto che questa tendenza non venga replicata ancora in futuro. Da una parte abbiamo sempre di più la diffusione di servizi cloud di varia tipologia e, spesso, questi sono gestiti da aziende americane, quindi con un approccio più americano – noi stessi siamo un'azienda americana. Per mille ragioni, molte aziende hanno adottato già da tempo il trend di esternalizzare i propri servizi in un ambiente cloud, il che può essere un vantaggio, anche economico… anche se su questo punto se ne potrebbe discutere.

Dall'altra, quello che sta succedendo negli ultimi tempi – la gestione dei dati non condivisa tra vari Paesi, i vari casi di gestione di dati non propriamente conformi alle politiche di sicurezza delle varie regioni, ecc. – ha portato e sta portando alcune aziende critiche a rivalutare la possibilità di internalizzare quelli che sono i servizi. È in effetti ovvio che se noi custodiamo una serie di dati o le nostre applicazioni in un server, in un bunker, magari non collegato alla rete, l'esposizione al rischio è minore… Però non è neanche uno scenario troppo fattibile, no? Perché ormai siamo tutti connessi, i servizi devono essere accessibili da ovunque e quindi si rende necessario trovare un compromesso.

[Come Cisco] abbiamo creato un'offerta per indirizzare queste problematiche, per permettere alle aziende di mantenere "in casa" – in Europa o in Italia – alcuni tipi di servizi critici, proprio in relazione a quello che sta succedendo negli ultimi tempi e che spinge qualcuno a guardare in maniera diversa la gestione dei dati e delle applicazioni. A questo proposito, soprattutto per quanto riguarda le infrastrutture critiche, non possiamo dimenticarci della direttiva europea NIS2, che impone dei paletti specifici di sicurezza in termini di tecnologie, processi e prodotti per gestire in maniera sicura dati, applicazioni, connettività utenti e via dicendo. Questo ha aiutato molto a livello europeo a indirizzare delle problematiche di sicurezza che prima erano altrimenti un po' più esposte.

Mentre la diffusione dell'intelligenza artificiale ha aperto nuove vulnerabilità, a livello di cybersicurezza? Microsoft afferma che ormai il 30% del proprio codice è generato da sistemi di IA – e l'aumento dei bug in Windows non sta aiutando a mettere in buona luce questa scelta –, ma in generale assistiamo alla nascita di un numero crescente di aziende che puntano sul cosiddetto vibe coding, delegando alle intelligenze artificiali una parte significativa del processo di programmazione.

Eh, parecchie. Utilizzare il *vibe coding* è senz'altro un vantaggio dal punto di vista di programmazione, soprattutto in termini di *time-to-market*, quindi di velocità. Però, come tutti noi sappiamo, nel fare le cose in fretta non sempre le si fanno bene. Soprattutto dal punto di vista della *security*. Quindi utilizzare il *vibe coding* vuol dire tante volte sorvolare alcuni tipi di controlli di sicurezza, vuol dire introdurre alcune vulnerabilità che prima non c'erano perché, magari, prendo il codice esistente da una libreria *open* [codici pre-scritti e

distribuiti gratuitamente, ma non sempre affidabili, n.d.r.]. È chiaro che il mercato chiede sempre più velocità, di fare di più con meno risorse, ma non sempre la *security* va a pari passo con la velocità.

Ciò che noi vediamo oggi è l'adozione crescente di queste modalità di fare sviluppo, perché ci semplifica la capacità di fare *coding*, ma questo non deve andare a discapito di controlli di sicurezza che noi davamo per scontati fino all'anno scorso. Noi stessi in Cisco utilizziamo le AI, però abbiamo un rigido controllo di sviluppo del software con dei passaggi rigidi per quanto riguarda la sicurezza. Noi abbiamo un progetto di Secure by Design, tutte le attività di sviluppo, software o hardware, hanno dei passaggi ristretti di sicurezza dove noi controlliamo il codice più volte. […] Lo controlliamo sia come sorgente che come esecuzione, ottimizziamo e "hardenizziamo" [tempriamo, n.d.r.] tutte le tecnologie che mettiamo all'interno delle nostre soluzioni.

L'AI è qua per rimanere, però è importante sicuramente cercare di gestirla. Lo vediamo anche nell'utilizzo delle tecnologie che proponiamo ai nostri clienti: hanno l'esigenza di controllare sia l'utilizzo delle app AI da parte degli utenti, sia lo sviluppo di applicazioni di AI che utilizzano dei modelli già facilmente utilizzabili e distribuiti, ma che hanno notoriamente delle vulnerabilità che prima non esistevano. E quindi è importante utilizzare processi e tecnologie per cercare di rendere controllato anche l'utilizzo delle AI.

Tutto ciò richiede però una consapevolezza generalizzata… a che punto è l'alfabetizzazione digitale in Italia?

Bella domanda, diciamo che, per come la vedo, c'è spazio per migliorare, ma sono stati fatti negli anni dei passi avanti importanti. È difficile generalizzare, ma diciamo che il livello è medio: si parla di IT da tanto tempo, si parla di *security* da tanto tempo e, secondo me, le direttive europee che dicevo prima hanno aiutato molto a sviluppare consapevolezza nell'adozione e nell'utilizzo delle tecnologie e dei processi. In ambito istituzionale, per esempio, c'è sempre più attenzione e ci sono attività di *awareness* fatte internamente per cercare di migliorare quello che è l'approccio all'*information technology* e alla *security*. È vero però che anche questo è un percorso.

Nel parlare di AI: oggi tutti parlano di *agentic AI*, cioè di agenti che lavorano in maniera autonoma. Ci siamo appena lasciati alle spalle l'utilizzo, la capacità o la facilità di uso dei chatbot e oggi il fulcro del discorso sono gli agenti AI che aiutano le aziende ad automatizzare tutta una serie di attività e che potrebbero aiutare le attività quotidiane. Forse avrete letto che c'è già un primo social network fatto esclusivamente da agenti AI che ha già 1 milione e 500.000 iscritti. È molto curioso vedere cosa questi agenti si dicono tra

loro su questo portale. È forse un esempio meno produttivo e più goliardico, ma la tendenza va in questa direzione. Noi abbiamo già agenti AI che aiutano nella gestione delle reti, che in maniera autonoma sono in grado di analizzare il comportamento delle reti e di identificare eventuali guasti anche in maniera preventiva.[…] Tutto questo si sta spostando anche nella cybersecurity. Quindi anche il mondo AI, l'IT, la tecnologia… sta veramente andando a una velocità talmente rapida che è molto difficile poi avere continua consapevolezza di quello che sta succedendo.

Visto che hai citato Moltbook [il social per le IA, n.d.r.]… in quel caso è abbastanza complesso capire quanto è importante il ruolo degli utenti nel gestire le interazioni tra agenti IA. È tutto molto torbido, non è chiaro quanto e come gli interventi delle intelligenze artificiali siano guidati dagli esseri umani. Questa ambiguità crea complessità nel settore della cybersicurezza?

Assolutamente sì. Crea delle complessità perché, come hai detto tu, non è mai facile capire cosa c'è dietro, se c'è l'umano e fin dove c'è l'umano. Lascio stare un attimo Molbook e guardo un po' quello che vediamo dal punto di vista più commerciale: oggi la capacità di gestire dal punto di vista dei processi questi agenti AI è qualcosa di complesso, sia dal punto di vista della legislazione – che non c'è –, che da quello propriamente tecnologico.

Dal punto di vista della *security*, oggi non ci viene più richiesto di fornire o di produrre delle tecnologie che controllino cosa fa un utente, ma di applicare delle tecnologie che vanno a controllare che cosa fa un agente AI e che siano in grado di distinguere l'attività di un'intelligenza artificiale da quella di un essere umano. E che siano in grado, magari, di applicargli *policy* diverse di controllo. Non è banale, da questo punto di vista. Le tecnologie esistono, ma quello che ancora manca è il processo per gestirle. Vediamo una tecnologia che evolve molto rapidamente e – al di là della legislazione, che è ovvio non possa tenere lo stesso passo – anche i processi interni e la complessità che i clienti devono gestire, inclusi quelli più grandi, fanno fatica a stare al ritmo. In questo momento non c'è una velocità univoca uguale per tutti… e forse questa è la cosa più complicata, la più complessa.

Vibe coding, agenti IA e cloud: stiamo costruendo un sistema troppo veloce per essere sicuro?

Walter Ferri

Giornalista milanese, per *L'Indipendente* si occupa di analisi nel campo della tecnologia, dei diritti informatici, della privacy e dei nuovi media, indagando le implicazioni sociali ed etiche delle nuove tecnologie. È coautore e curatore del libro *Sopravvivere nell'era dell'Intelligenza Artificiale*.