CONNUBIO CRITICO TRA SICUREZZA DELLE INFORMAZIONI
ED INTELLIGENZA ARTIFICIALE

AI per la Cybersicurezza

Pierluca Ferraro

23 Ottobre 2026

Di cosa parliamo

Perché

Che cosa vuol dire “sicuro”, e perché non riguarda solo i computer

01

02

Dagli attacchi...

03

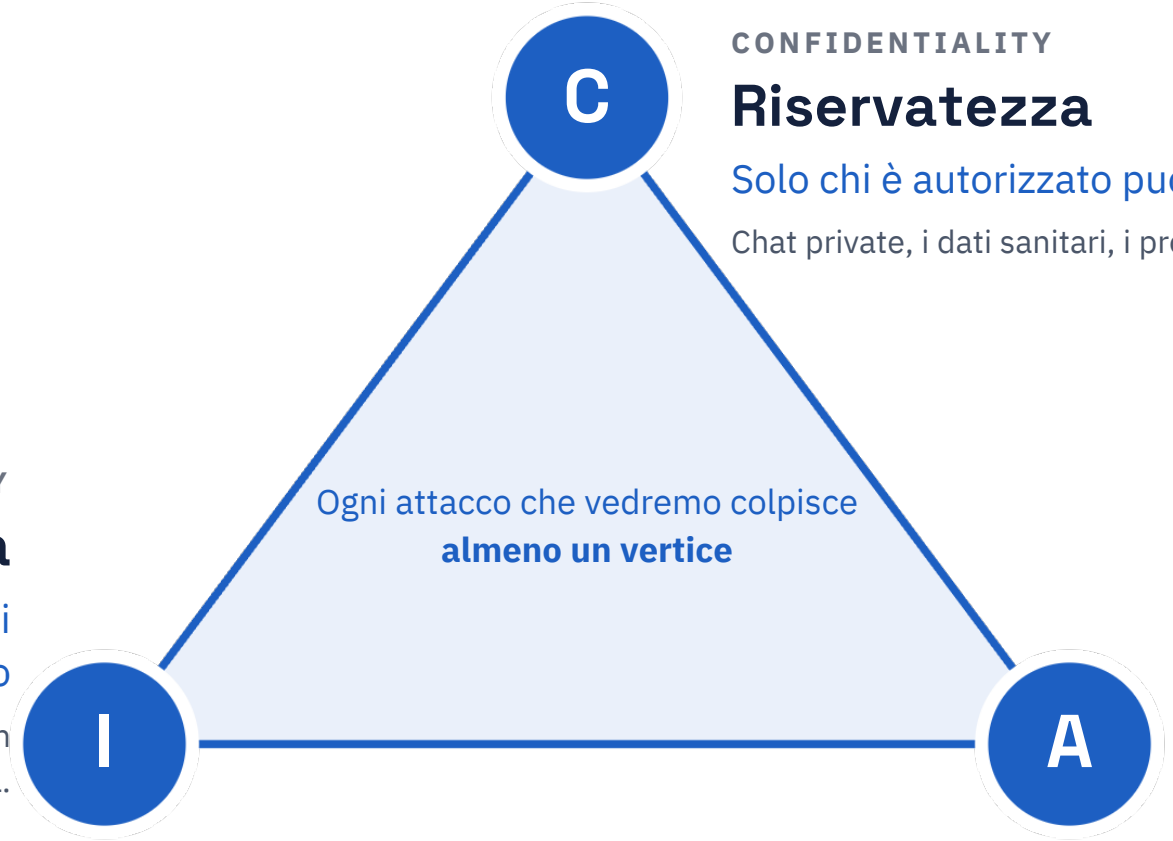
...alla difesa: come può aiutarci l’IA?

04

Sfide aperte

INTRODUZIONE

Che cosa vuol dire “sicuro”?



CONFIDENTIALITY

Riservatezza

Solo chi è autorizzato può leggere

Chat private, i dati sanitari, i progetti di un’azienda.

INTEGRITY

Integrità

Nessuno può modificare di nascosto

Un voto sul registro elettronico, un bonifico, la dose di cloro nell’acqua.

AVAILABILITY

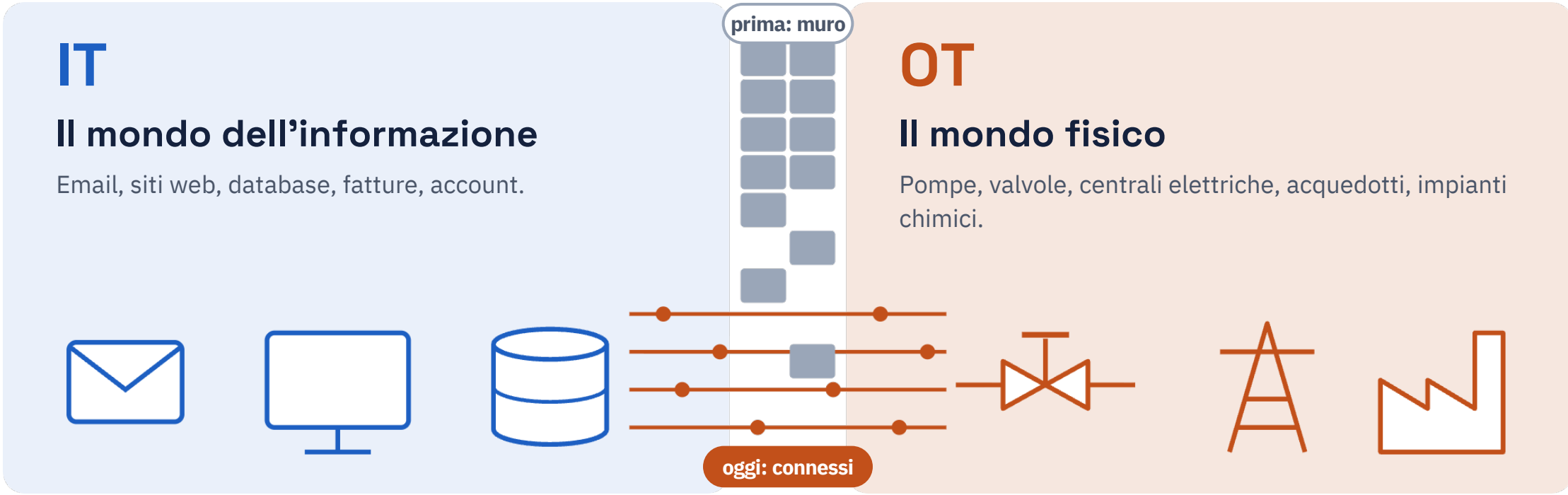
Disponibilità

Il servizio c’è quando serve

Il sito della banca, la rete elettrica, il pronto soccorso.

INTRODUZIONE

Non solo computer



Quando un attacco esce dallo schermo, le conseguenze diventano **fisiche**: blackout, carburante che manca, acqua contaminata.

02

Dagli attacchi... alla difesa

Il casinò e l’acquario

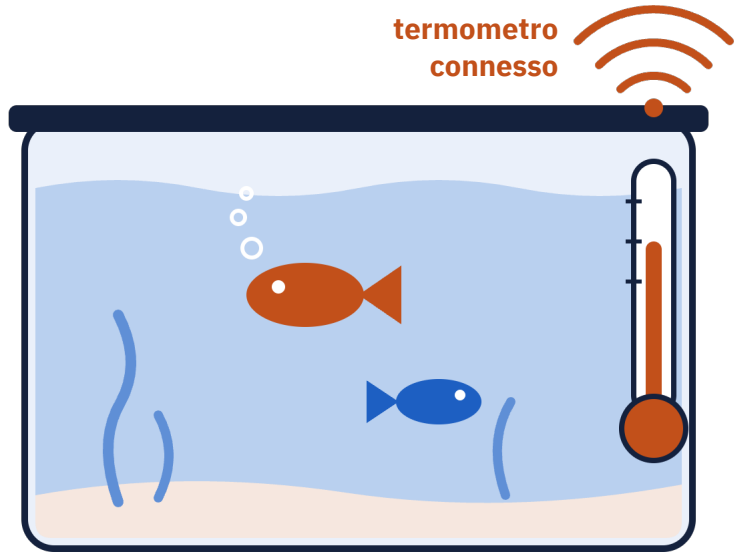
Nord America, 2017

IL CASINÒ, NORD AMERICA, 2017

Il termometro dell’acquario

Nella hall di un casinò non meglio identificato c’è un grande acquario, e per controllarne la temperatura a distanza qualcuno installa un **termometro connesso a Internet**.

Gli attaccanti entrano da lì, si muovono nella rete interna e arrivano al **database dei clienti più facoltosi**, che poi portano fuori passando di nuovo dal termometro.

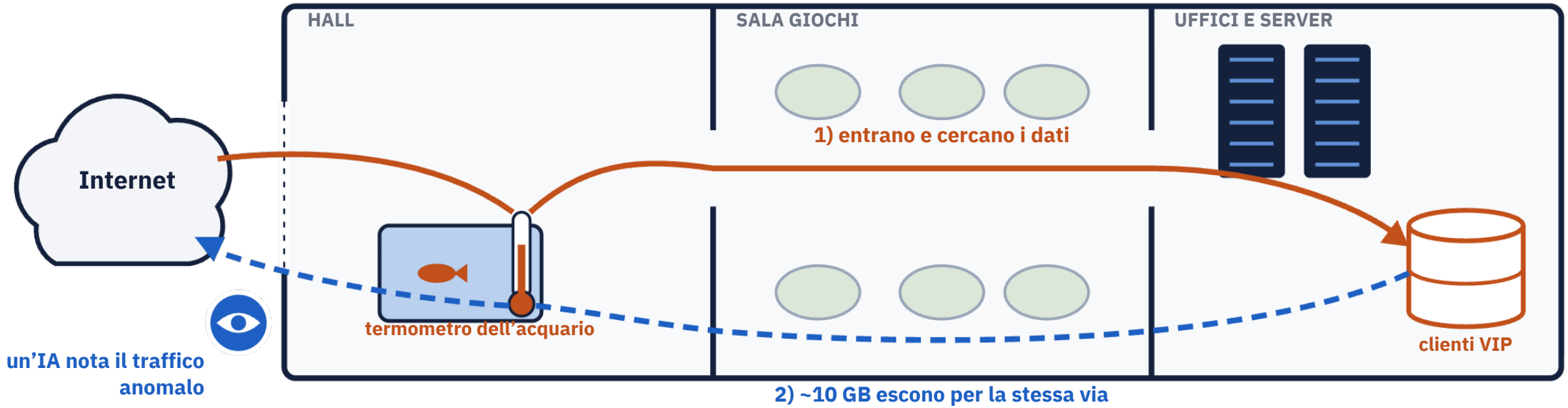


Un sensore da pochi euro

è bastato per aprire la porta a tutta la rete

IL CASINÒ

L’anello più debole



Tutto fa parte del perimetro

Ogni oggetto connesso fa parte del perimetro, anche quello a cui nessuno pensa.

Triton

Arabia Saudita, 2017

TRITON, ARABIA SAUDITA, 2017

Il guasto che non era un guasto



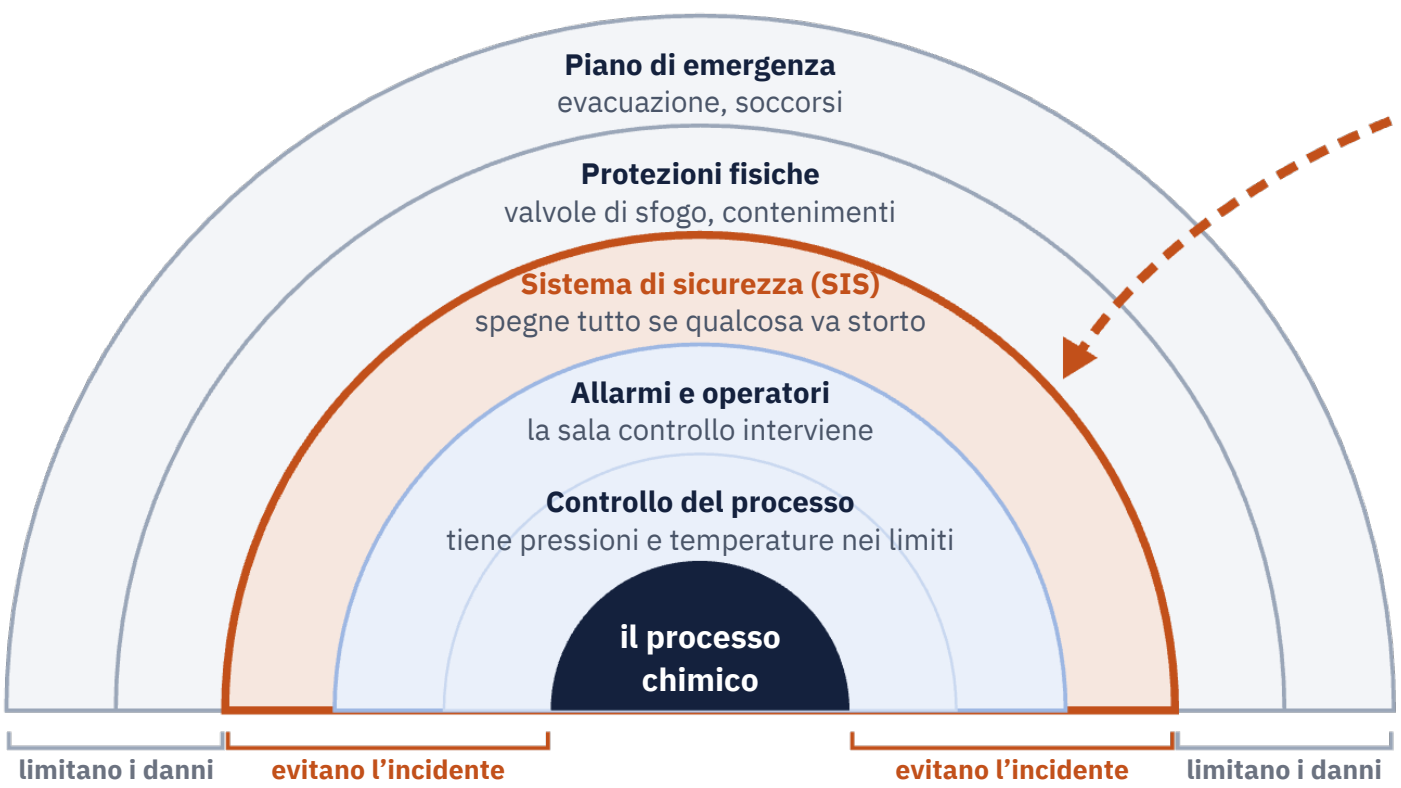
Petro Rabigh: raffineria e impianto petrolchimico sul Mar Rosso, esposto a rischi come fughe di gas tossico ed esplosioni.

A **giugno 2017** un controllore di sicurezza spegne all'improvviso una sezione dell'impianto, ma l'ispezione non trova alcuna anomalia.

Il caso viene archiviato come guasto meccanico.

TRITON

4 agosto 2017, ore 19:43



Triton mira all'ultima barriera prima dell'incidente

Due sistemi di sicurezza scattano insieme, mentre in sala controllo gli schermi mostrano che è **tutto normale**.

Gli esperti trovano **Triton**, un malware costruito per manomettere **il SIS**, il sistema che spegne tutto prima che un guasto diventi un'esplosione.

NotPetya

Ucraina, e poi il resto del mondo, 2017

NOTPETYA

Che cosa fa un ransomware

1 **Entra**

attraverso un’email di phishing, una password rubata o un servizio non aggiornato.

2 **Ruba e cifra**

copia i dati più preziosi e poi rende illeggibili i file, spesso anche i backup.

3 **Ricatta**

chiede un riscatto in criptovaluta per la chiave e minaccia di pubblicare i dati rubati.



NOTPETYA, GIUGNO 2017

Un aggiornamento avvelenato



Sembrava un ransomware, perché chiedeva un riscatto, ma **non esisteva alcun modo di recuperare i dati**, nemmeno pagando.

NOTPETYA

Danni collaterali

Maersk è una delle più grandi compagnie di navigazione al mondo per il trasporto di merci in container. In pochi minuti i suoi terminal si bloccano, le navi non vengono scaricate e i camion restano in coda ai cancelli dei porti.

Sono serviti 10 giorni per ripartire, lavorando su carta e WhatsApp



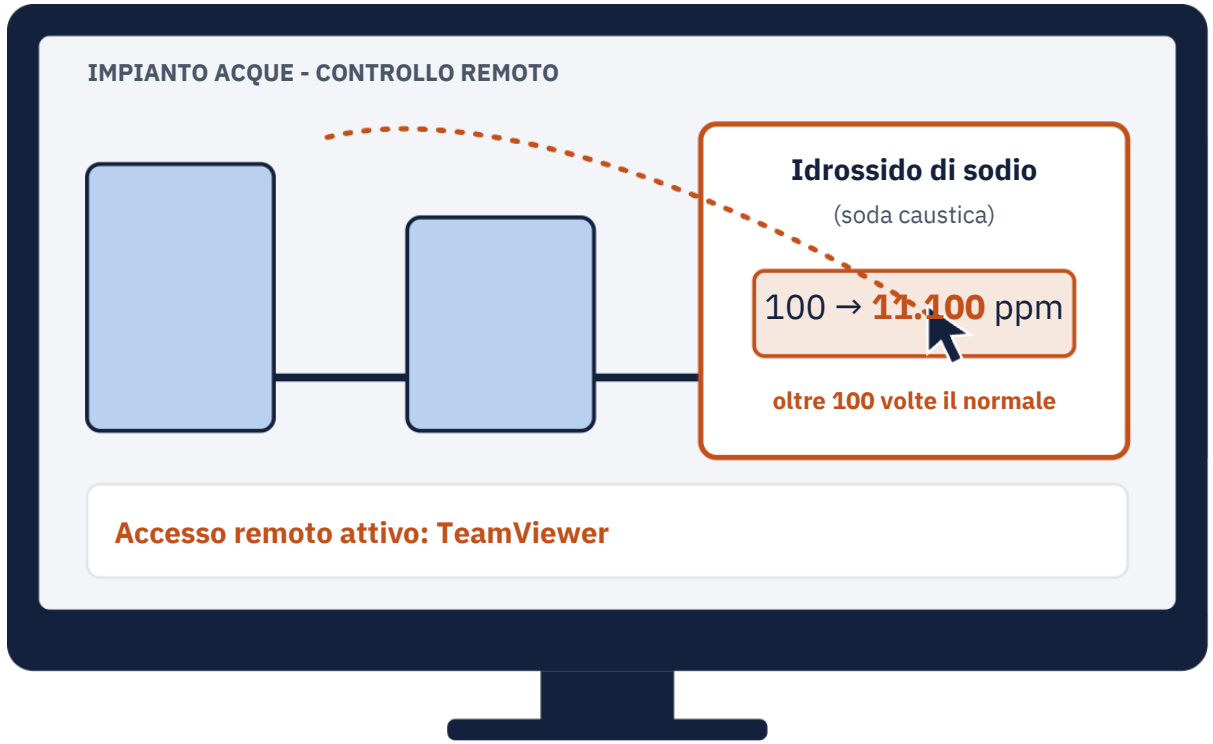
Oldsmar

Florida, 2021

OLDSMAR (FLORIDA), 5 FEBBRAIO 2021

Una tranquilla giornata davanti allo schermo

Un piccolo impianto di **trattamento delle acque** che serve 15.000 abitanti. Il sistema di controllo si può usare anche da remoto.



8:00 circa
Un accesso remoto breve: l'operatore pensa sia il suo capo, che lo fa spesso.

13:30
Il cursore torna a muoversi da solo e apre il pannello della soda caustica.

3-5 minuti
Il valore viene portato a 11.100 ppm, poi la sessione si chiude.

Subito dopo
L'operatore riporta il valore a 100 e dà l'allarme: nessuna conseguenza.

OLDSMAR

Il colpo di scena

Cosa hanno trovato gli investigatori

- Windows 7, senza aggiornamenti di sicurezza
- Software di controllo remoto esposto su Internet
- Una sola password, condivisa da tutti i PC

Nel 2023 un ex manager dichiara:
“**forse non c’è stato nessun hacker**, ma l’errore di un dipendente”.



Colonial Pipeline

Stati Uniti, 2021

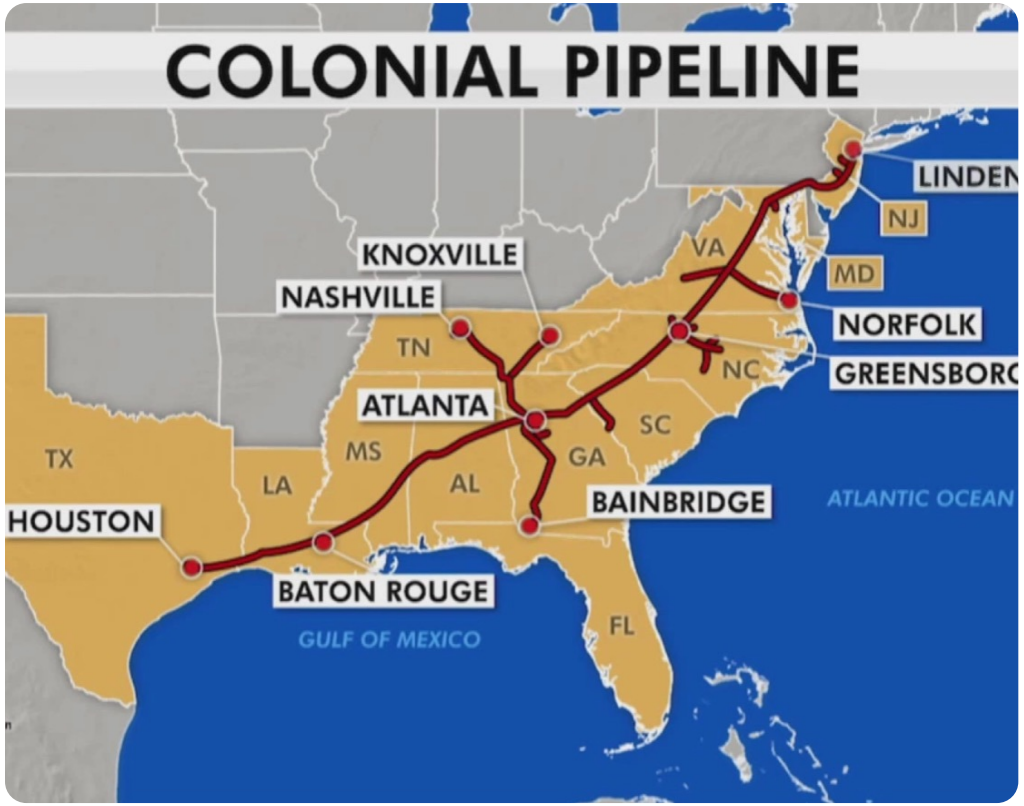
COLONIAL PIPELINE

Una sola password per l’intero oleodotto

Colonial Pipeline: il più grande oleodotto di carburanti degli USA, dal Texas al New Jersey. Circa il **45%** del carburante della costa est.

La porta d’ingresso: un account VPN **non più usato**, senza autenticazione a due fattori, **con una password già trapelata altrove**.

Il gruppo DarkSide ruba circa 100 GB di dati, poi cifra i sistemi con un **ransomware**.



COLONIAL PIPELINE

Stop al sistema di fatture



Il ransomware **non ha mai toccato** pompe, valvole e sensori, quindi l'oleodotto avrebbe potuto continuare a funzionare.

I sistemi di **fatturazione** però erano cifrati e l'azienda non sapeva più quanto carburante consegnava, né a chi, così ha deciso di fermare tutto.

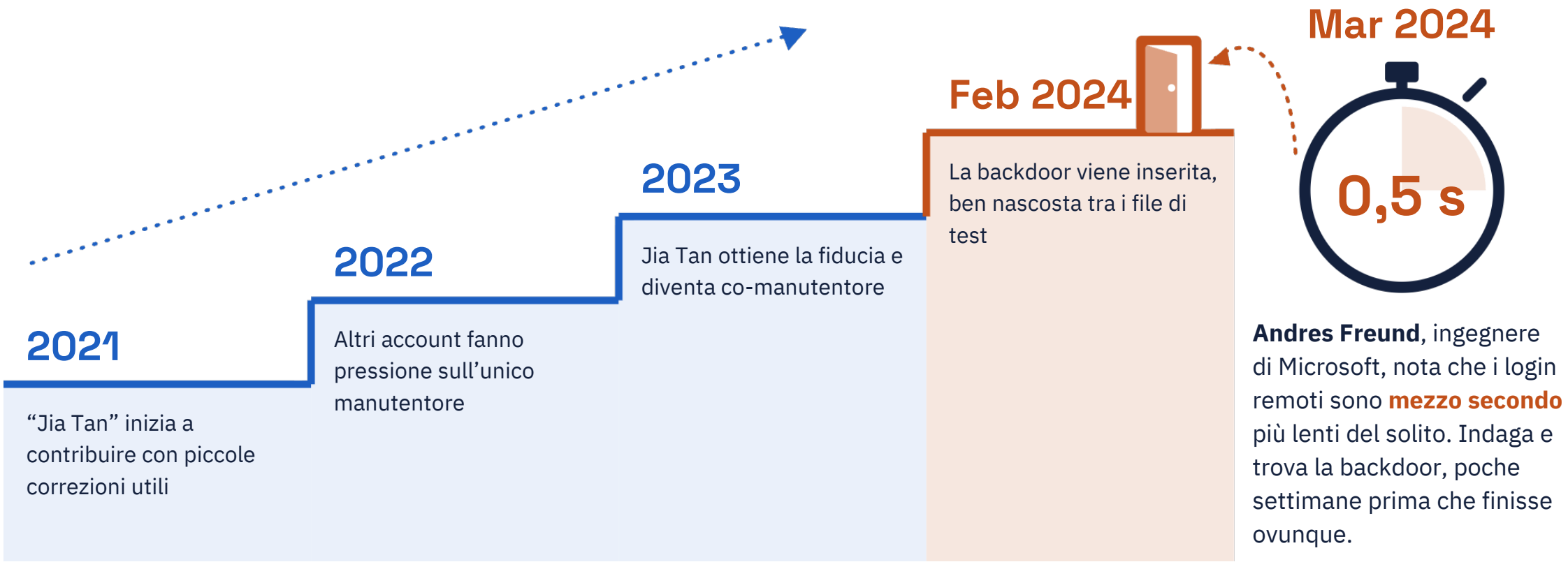
Ne sono seguiti lo stato di emergenza e code ai distributori.

XZ Utils

Il mondo open source, 2024

XZ UTILS

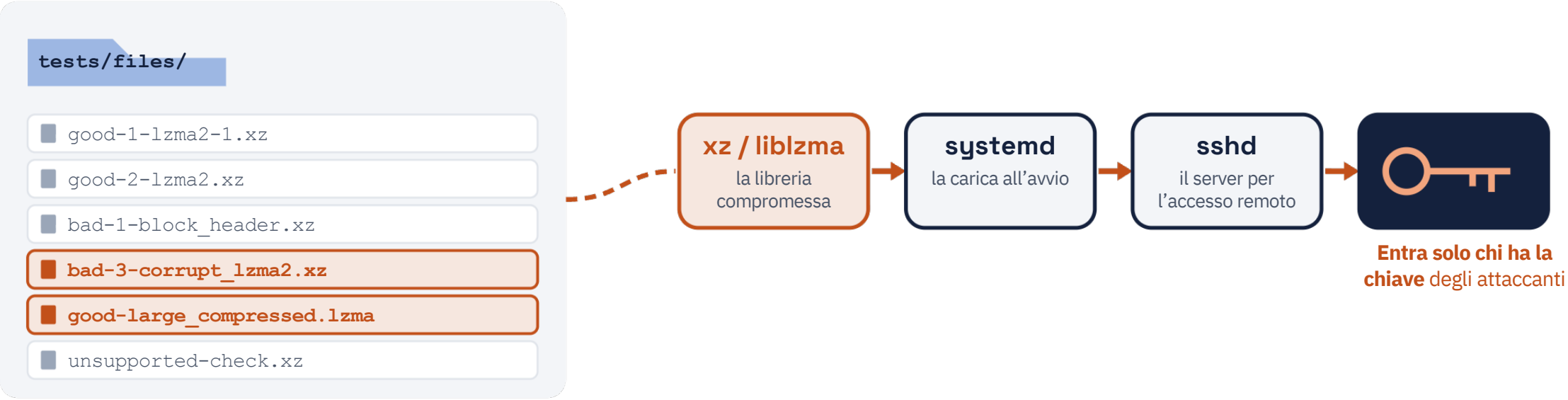
Mezzo secondo di troppo



Andres Freund, ingegnere di Microsoft, nota che i login remoti sono **mezzo secondo** più lenti del solito. Indaga e trova la backdoor, poche settimane prima che finisse ovunque.

XZ UTILS: LA BACKDOOR

Nascosta in piena vista



Sembrano normali file di prova "rotti". In realtà contengono, cifrato, il codice della backdoor.

Una libreria di compressione, che nessuno associa alla sicurezza, è diventata la **chiave dell'accesso remoto** a milioni di server.

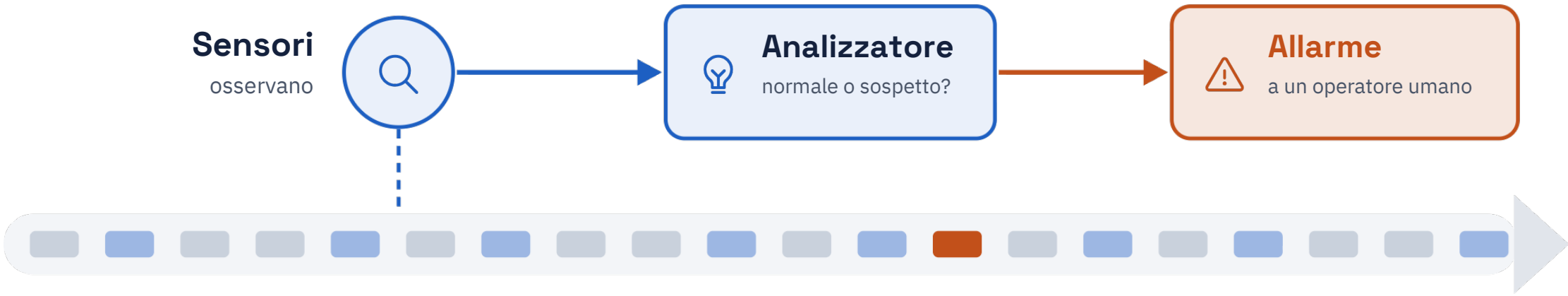
03

Come può aiutarci l'IA?

Gli Intrusion Detection System e le macchine che imparano dagli esempi

INTRUSION DETECTION SYSTEM

L’antifurto della rete



Traffico: tutto ciò che passa nella rete

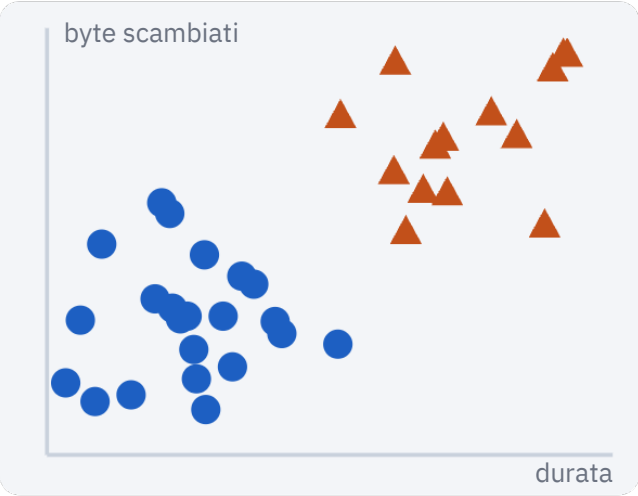
...che **continua a scorrere**, anche quello sospetto

Un IDS non blocca il traffico, ma **osserva e avvisa**, come un sistema di videosorveglianza con qualcuno che guarda i monitor.

MACHINE LEARNING

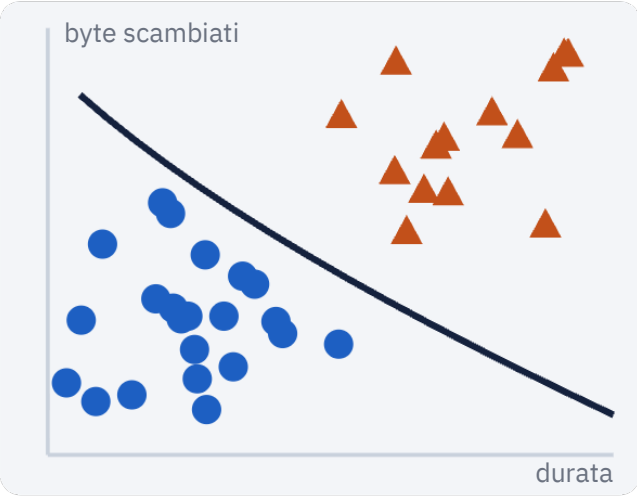
Come impara una macchina

1 Esempi



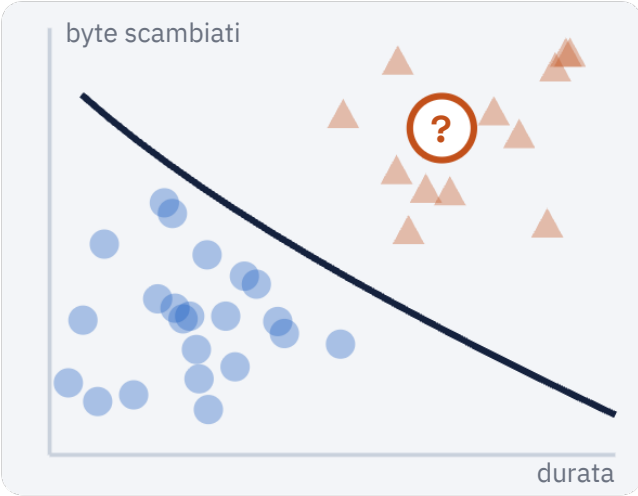
Connessioni etichettate:
normale o **attacco**

2 Addestramento



Il modello traccia una frontiera

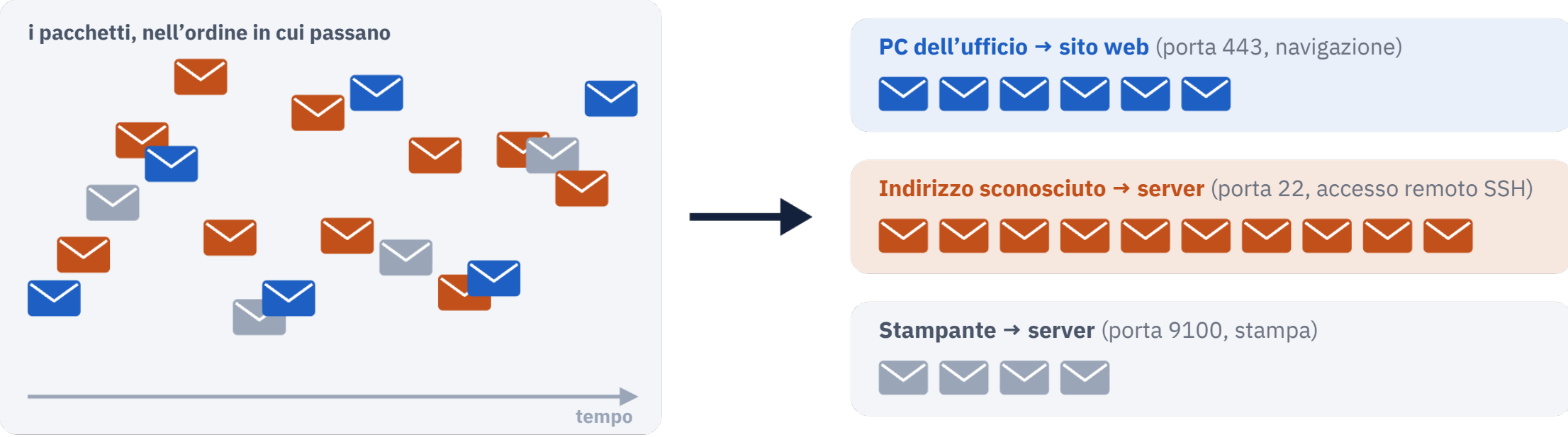
3 Decisione



Il nuovo punto cade tra gli attacchi

DAI DATI ALLE CARATTERISTICHE

Dai pacchetti alle conversazioni



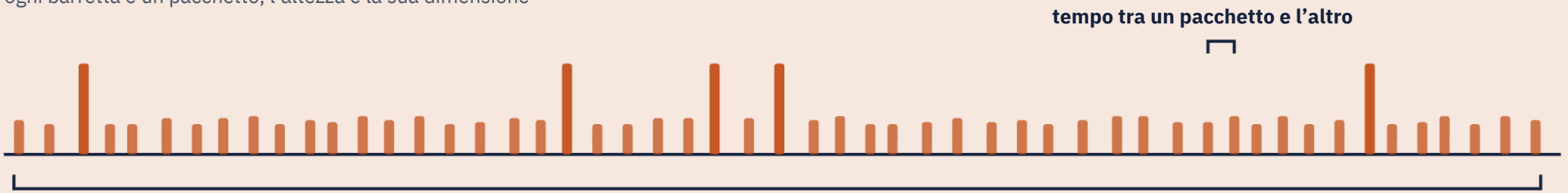
Come per la posta, si raggruppano le buste con lo stesso **mittente**, lo stesso **destinatario** e lo stesso **servizio**. Ogni gruppo è un **flusso**, una "conversazione".

DAI DATI ALLE CARATTERISTICHE

Dalla conversazione ai numeri

un flusso: la conversazione dell'indirizzo sconosciuto con il server SSH

ogni barretta è un pacchetto, l'altezza è la sua dimensione

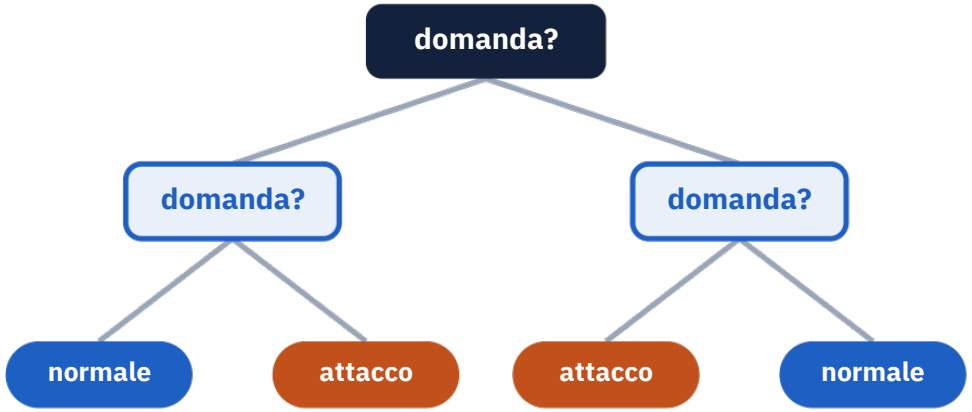


si riassume tutto in una riga di numeri: le caratteristiche

35	97%	2 MB	SSH	03:12
tentativi al minuto	login falliti	dati inviati	servizio (porta 22)	ora

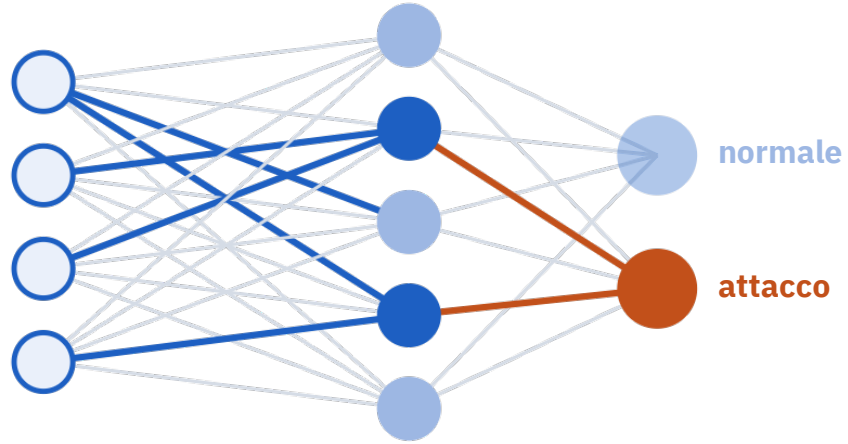
MACHINE LEARNING

Due modi diversi di ragionare



Albero decisionale

Una sequenza di domande: a ogni risposta si scarta una parte delle possibilità.



Rete neurale

Tante piccole unità di calcolo collegate tra loro, ispirate ai neuroni del cervello.

ESEMPIO

Una connessione da giudicare

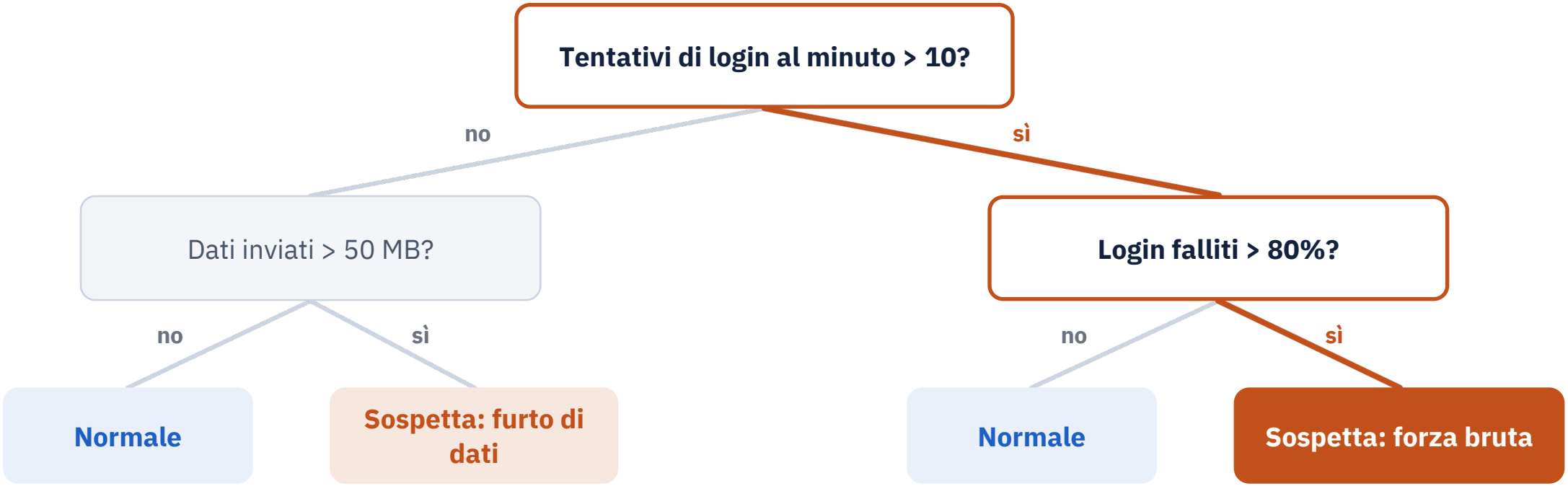
Tentativi di login al minuto	**35**
Login falliti	**97%**
Dati inviati	**2 MB**
Servizio	**accesso remoto (SSH)**
Ora	**03:12**

Un esperto, guardandola, direbbe che qualcuno sta provando a **indovinare una password**.

Come ci arrivano i due modelli?

ALBERO DECISIONALE

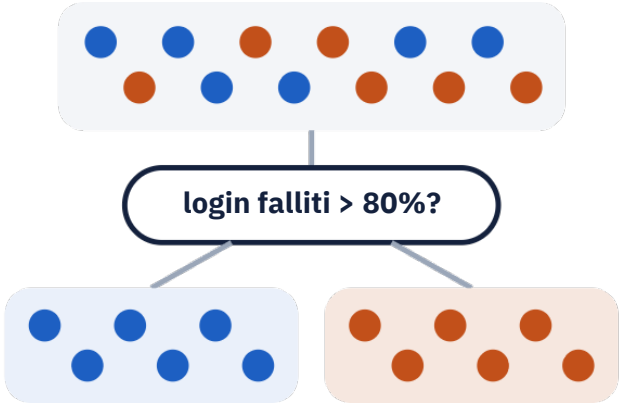
Una domanda dopo l’altra



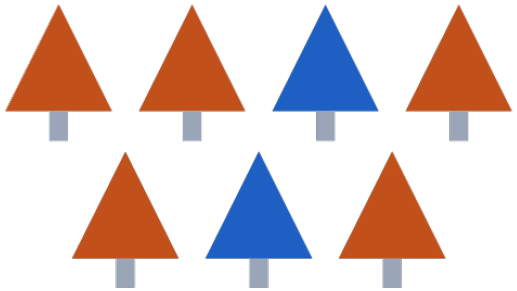
La nostra connessione fa **35 tentativi** al minuto e il **97%** fallisce: l’albero segue il percorso evidenziato e conclude che è un attacco a forza bruta.

ALBERO DECISIONALE

Da dove vengono le domande?




```
SE tentativi/min > 10
E  login falliti > 80%
ALLORA
  → forza bruta
```



5 voti su 7: attacco

Come impara

Tra tutte quelle possibili, l'algoritmo sceglie la domanda che separa meglio gli esempi normali da quelli malevoli, e ripete su ciascun ramo.

Punto di forza

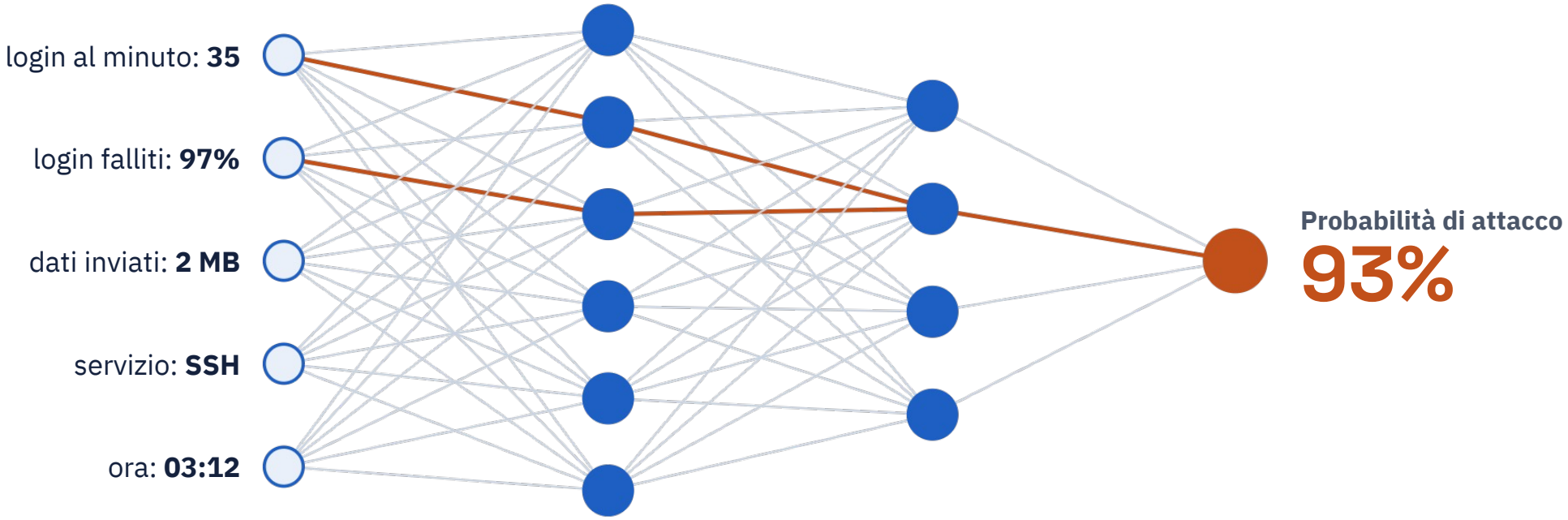
Si legge come un elenco di regole "se... allora...", quindi possiamo sempre ricostruire perché il modello ha preso una certa decisione.

Limite e rimedio

Un singolo albero è fragile. Per questo si usano le **foreste**: centinaia di alberi diversi che votano insieme, più accurati ma meno facili da leggere.

RETE NEURALE

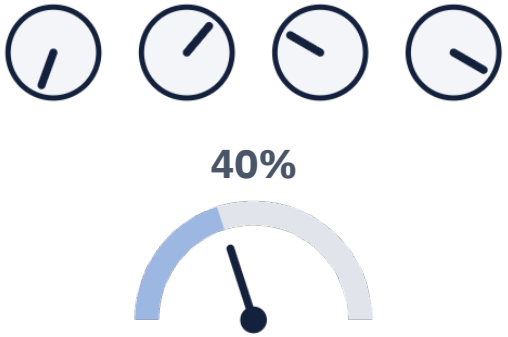
Tanti indizi combinati insieme

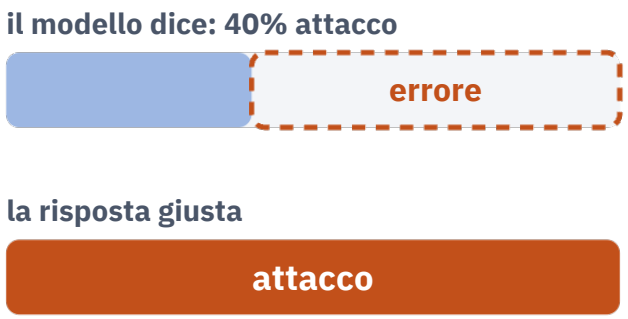


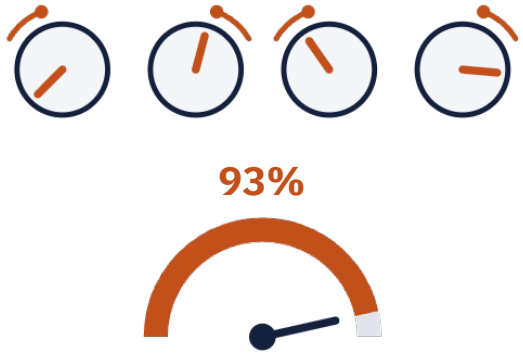
Ogni cerchio è un “neurone” che pesa i numeri ricevuti, li somma e passa il risultato allo strato successivo.
Imparare significa **regolare i pesi** finché le risposte sugli esempi diventano corrette.

RETE NEURALE

Da dove vengono i pesi?







1 All’inizio, a caso

I pesi sono come manopole: all’inizio sono girate a caso, e la rete dà risposte sbagliate.

2 Confronta

Per ogni esempio si confronta la risposta della rete con quella giusta: la differenza è l’errore.

3 Correggi un poco

Ogni manopola viene girata di poco, nella direzione che riduce l’errore, e si ricomincia.

Ripetuti su **milioni di esempi**, questi piccoli aggiustamenti costituiscono l’addestramento.

04

Sfide aperte

Il mondo cambia continuamente... e le decisioni?

CONCEPT DRIFT

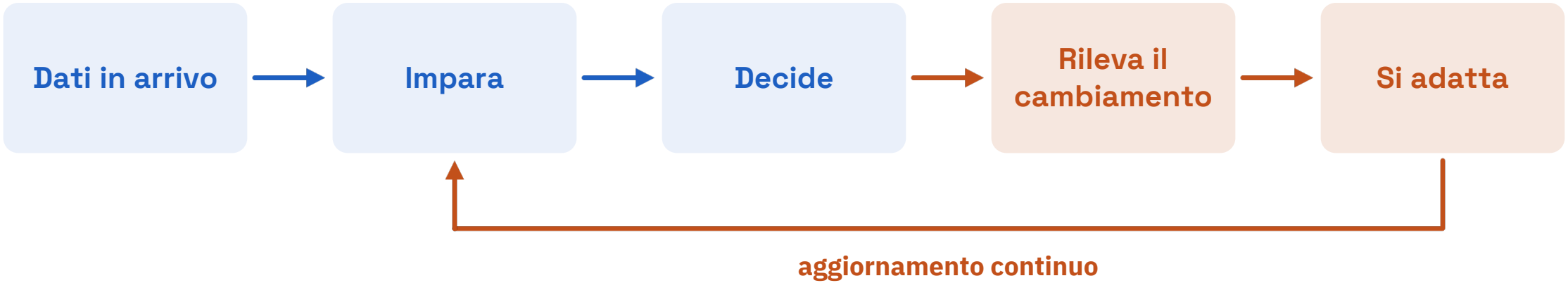
Il mondo cambia, il modello no

Un modello **impara dal passato**, ma il mondo va avanti: arrivano nuovi servizi e nuove abitudini, e gli attaccanti **inventano tecniche nuove**. Ciò che ieri era “normale” oggi può essere sospetto, e viceversa.



CONCEPT DRIFT

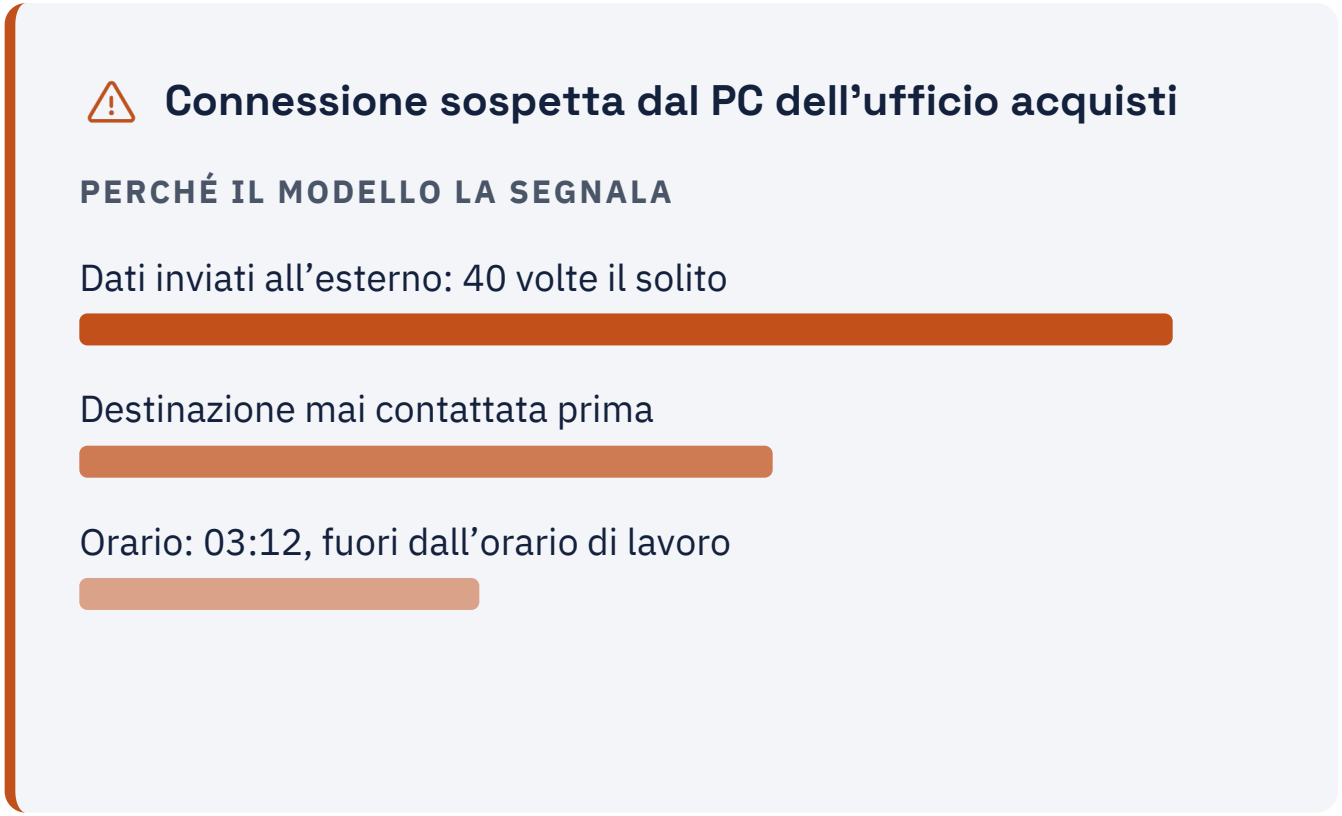
Un modello che si aggiorna



Non si addestra una volta per tutte: il modello **sorveglia se stesso** e, quando i dati cambiano, si aggiorna.

SPIEGABILITÀ

Un allarme che spiega il motivo



Con una spiegazione l’operatore può **verificare**, **decidere** in fretta e **correggere** il modello quando sbaglia.