Tutelattiva

ASSOCIAZIONE CONSUMATORI ACP
consumo lavoro cittadinanza

Presentazione 2025

INTELLIGENZA ARTIFICIALE

Conoscerla e usarla in sicurezza

Una guida semplice per capire cosa fa l'AI, come funziona e quali rischi comporta per il nostro futuro digitale.

20 Slide 15 Minuti

AI SAFETY

Cos’è l’Intelligenza Artificiale?

Apprendimento da Esempi

È una tecnologia che permette ai computer di imparare osservando dati ed esempi, simulando il processo di apprendimento umano.

Riconoscimento di Schemi

Non "pensa" come noi: analizza enormi quantità di informazioni per trovare pattern e strutture nascoste.

Onnipresenza Invisibile

Oggi la usiamo tutti, ogni giorno: dai social media ai navigatori GPS, spesso senza nemmeno rendercene conto.

Le origini dell’Intelligenza Artificiale

Dai primi concetti teorici alla nascita ufficiale della disciplina

1950

Test di Turing

Alan Turing pubblica "Computing Machinery and Intelligence" e propone il famoso test: "Le macchine possono pensare?"

1956

Conferenza di Dartmouth

John McCarthy e altri pionieri si riuniscono per un workshop estivo. Nasce ufficialmente il termine "Artificial Intelligence".

1974 - 1993

Gli "Inverni dell'AI"

Dopo l'entusiasmo iniziale, i risultati non arrivano velocemente come promesso. Tagli ai fondi e scetticismo bloccano la ricerca per anni.

Le tappe più Importanti

I momenti che hanno definito la storia moderna dell'Intelligenza Artificiale

1997

Deep Blue vs Kasparov

Il supercomputer di IBM batte il campione del mondo di scacchi Garry Kasparov. È il trionfo della potenza di calcolo bruta.

2011

Watson a Jeopardy!

Watson vince al famoso quiz televisivo USA. Dimostra per la prima volta la capacità di comprendere il linguaggio naturale complesso.

OGGI

Deep Learning & GPT-4

L'era dell'AI Generativa. Modelli come GPT-4 non solo analizzano, ma creano contenuti, codice e arte con qualità quasi umana.

Come funziona l’AI moderna

Il meccanismo di base che trasforma i dati in decisioni intelligenti.

1

DATI

Riceve Informazioni

La rete riceve un input (immagini, testo, numeri). Questa è la materia prima che l'AI deve analizzare.

2

NEURAL NETWORK

Elabora

Formata da "neuroni artificiali" collegati tra loro, ispirati al cervello biologico. Analizza le caratteristiche dei dati.

3

RISULTATO

Produce Risposta

La rete restituisce una predizione, una classificazione o un nuovo contenuto basato sull'analisi fatta.

Training Continuo: Più dati la rete "vede" ed elabora, più diventa precisa e brava nel tempo.

Tre modi in cui l’AI può imparare

Apprendimento Supervisionato

Si basa su esempi forniti dall'uomo che fungono da guida. È il metodo più comune per la classificazione.

COME FUNZIONA

Gli si mostrano esempi già etichettati (es. foto di gatti con tag "gatto", foto di cani con tag "cane").

Apprendimento Non Supervisionato

L'AI analizza i dati grezzi senza guida umana per trovare strutture nascoste.

COME FUNZIONA

Trova autonomamente somiglianze e differenze nei dati, raggruppandoli in cluster logici.

Apprendimento per Rinforzo

L'algoritmo impara interagendo con l'ambiente per massimizzare una ricompensa.

COME FUNZIONA

Impara per tentativi (trial & error), ricevendo premi e penalità (come AlphaGo).

L’AI NON capisce davvero

REALITY CHECK

Nessuna coscienza, solo matematica.

Anche i modelli più avanzati non hanno una reale comprensione del mondo. Non sanno cosa stanno dicendo; calcolano probabilità.

Non hanno intenzioni o sentimenti.

Non distinguono il vero dal falso.

"Pappagalli Stocastici"

Definizione coniata dai ricercatori per descrivere i grandi modelli linguistici (LLM).

Come un pappagallo che ripete suoni senza capirne il significato, l'AI ripete pattern statistici trovati nei dati di addestramento, assemblando parole in modo plausibile ma non consapevole.

ESEMPIO CONCRETO: GPT-4

Può generare una poesia commovente o un codice complesso, ma per il modello sono solo sequenze di token (numeri). Non "sa" cos'è una poesia o cosa fa il codice, sa solo che quelle parole statisticamente stanno bene insieme.

CLASSIFICAZIONE

Tipi di Intelligenza Artificiale

Cosa esiste oggi e cosa è ancora fantascienza.

REALTÀ ATTUALE

AI Debole (Narrow AI)

Sistemi progettati per svolgere un singolo compito specifico. Eccellono nel loro dominio ma non hanno coscienza o flessibilità generale.

ESEMPI CHE USI OGNI GIORNO

- Assistenti vocali (Siri, Alexa)
- Traduttori automatici
- Raccomandazioni (Netflix/Spotify)

NON ESISTE

AI Generale (AGI)

Un'intelligenza ipotetica pari a quella umana. Sarebbe capace di comprendere, imparare e risolvere problemi in qualsiasi dominio, non solo in uno specifico.

È il "Santo Graal" della ricerca, ma siamo ancora lontani dal raggiungerla.

TEORIA

Superintelligenza (ASI)

Un intelletto che supera di gran lunga le migliori menti umane in ogni campo, dalla creatività scientifica alla saggezza sociale.

Tema ricorrente nella fantascienza e nella filosofia etica, ma puramente speculativo oggi.

Il problema della “scatola nera”

BLACK BOX SYNDROME

Complessità inspiegabile

Le moderne reti neurali (Deep Learning) sono così complesse, con miliardi di parametri, che spesso nemmeno i creatori sanno spiegare il percorso logico esatto di una decisione.

Mancanza di "Explainability": sappiamo cosa ha deciso, ma non perché.

Perché è pericoloso?

In settori critici, "fidarsi e basta" non è accettabile:

Medicina
Perché è stata diagnosticata questa malattia e non un'altra?

Tribunali
Perché l'imputato è stato valutato "ad alto rischio"?

Banche & Finanza
Su quali basi esatte è stato negato il mutuo?

Senza trasparenza, è impossibile correggere errori sistematici o garantire l'assenza di discriminazioni nascoste.

ETICA & RISCHI

Quando l’AI diventa ingiusta

Caso Studio: Algoritmo COMPAS (USA)

Il Software

Utilizzato nei tribunali americani per prevedere il "rischio di recidiva" (probabilità di commettere nuovi crimini) degli imputati, aiutando i giudici a decidere le pene.

Obiettivo: Rendere le sentenze più oggettive.

I Dati di Training

L'algoritmo ha imparato dai dati storici sugli arresti. Questi dati riflettevano decenni di disuguaglianze e controlli di polizia sproporzionati verso le minoranze.

BIAS RAZZIALE

Risultato: Discriminazione

Un'indagine di ProPublica ha rivelato che l'algoritmo sbagliava in modi diversi per gruppi diversi:

TASSO DI "FALSI POSITIVI" (ERRONEAMENTE CLASSIFICATI AD ALTO RISCHIO)

**Gli imputati neri venivano etichettati come "alto rischio" il doppio delle volte rispetto ai bianchi, anche quando non commettevano più reati.*

Il paradosso

Anche togliendo la variabile "razza", il problema rimaneva. L'AI trovava "proxy" (es. codice postale, reddito) correlati alla razza.

AI e Cybersecurity: una forza doppia

Arma a Doppio Taglio

L'Intelligenza Artificiale è una tecnologia neutrale che può essere utilizzata sia per proteggere le infrastrutture che per violarle.

L'Uso degli Attaccanti

I criminali informatici sfruttano l'AI per creare attacchi più sofisticati, veloci e difficili da rilevare.

La Risposta dei Difensori

Aziende e governi integrano l'AI nei sistemi di sicurezza per anticipare le minacce e reagire in tempo reale.

Come gli ATTACCANTI usano l'AI

Le nuove tecnologie rendono le minacce più sofisticate, veloci e difficili da riconoscere.

Phishing Perfetto

Email e messaggi senza errori grammaticali, altamente personalizzati e persuasivi, generati in pochi secondi.

Deepfake Audio & Video

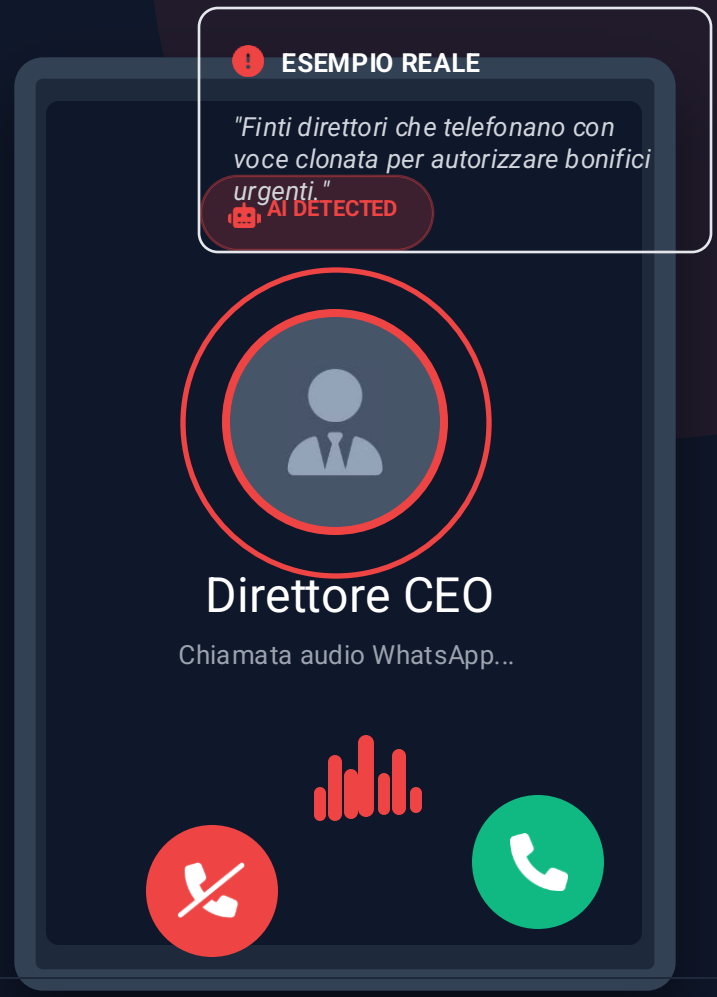
Creazione di cloni vocali o video sintetici indistinguibili dalla realtà per truffare dipendenti e famiglie.

Manipolazione di Massa

Generazione massiva di fake news e contenuti polarizzanti per influenzare l'opinione pubblica.

Attacchi Mirati e Veloci

L'AI automatizza la ricerca di vulnerabilità, rendendo gli attacchi più rapidi e precisi.



CYBERSECURITY

Come i DIFENSORI usano l’AI

Le organizzazioni sfruttano l'AI per anticipare le minacce, analizzare i dati e proteggere le infrastrutture in tempo reale.

Rilevamento Preventivo

Individuare schemi anomali nel traffico di rete per bloccare gli attacchi prima che causino danni reali.

Analisi Rapida dei Log

Elaborare milioni di eventi di sistema in pochi secondi per identificare incidenti invisibili all'operatore umano.

Addestramento e Simulazione

Generare scenari di attacco realistici (es. phishing simulato) per formare il personale e testare la resilienza.

Risposta Automatizzata

Isolare automaticamente i sistemi compromessi e applicare contromisure istantanee (SOAR) senza intervento manuale.

VANTAGGIO CHIAVE

"Riduzione dei tempi di risposta da giorni a pochi minuti grazie all'automazione intelligente."

SCENARIO GLOBALE

Il contesto geopolitico

COMPETIZIONE TECNOLOGICA

USA

Leader attuale trainato dal settore privato e dalle Big Tech.

- Supremazia HardwareInnovazione su chip e design.
- Modelli GenerativiPatria di GPT, Claude, Gemini.

RISCHIO GAP TECNOLOGICO

EUROPA

Cerca una "Terza Via" basata sui diritti e sulla regolamentazione.

- RegolamentazioneAI Act come standard globale.
- Approccio EticoProtezione dei cittadini.

CINA

Lo sfidante che punta all'autosufficienza e al primato entro il 2030.

- Pianificazione StataleInvestimenti massicci governativi.
- Controllo e DatiEnorme disponibilità di dati per training.

LA SFIDA CRUCIALE: SOVRANITÀ DIGITALE

Per non diventare "colonie digitali", nazioni e aziende devono mantenere il controllo sui propri dati, sulle infrastrutture critiche e sulle tecnologie che utilizzano

SCENARIO ATTUALE

Le minacce CONCRETE oggi

Tecniche sempre più sofisticate rendono difficile distinguere la realtà dalla finzione.

Truffe "Pig Butchering"

Manipolazione emotiva a lungo termine (spesso romantica) per indurre le vittime a investimenti fraudolenti.

Deepfake Iper-realistici

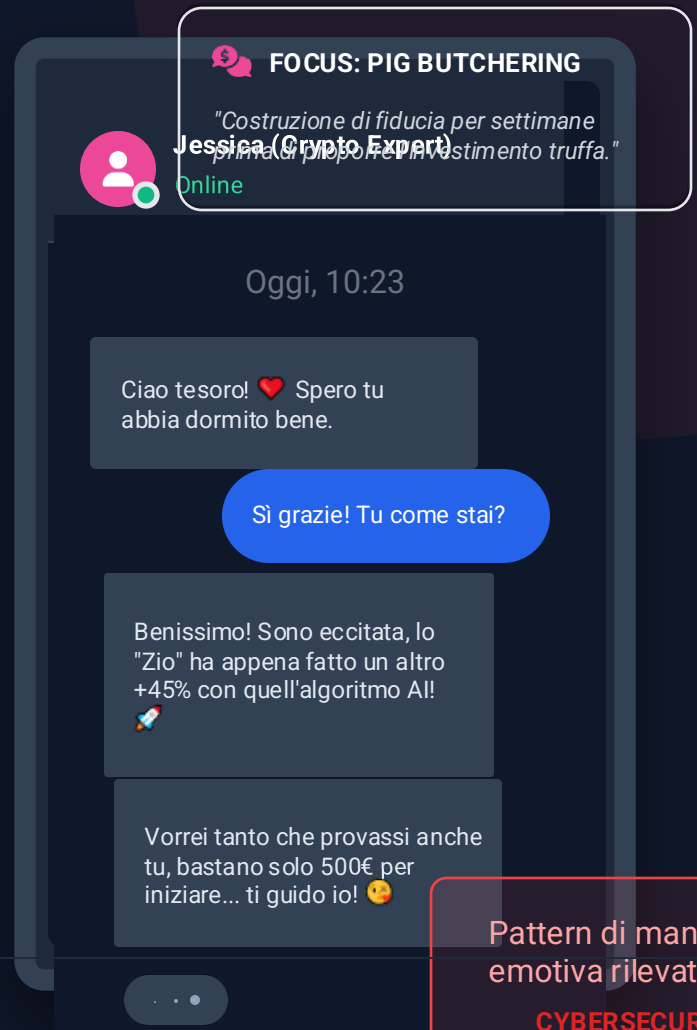
Clonazione perfetta di volti e voci, usata per truffe al CEO, ricatti e furti d'identità digitale.

Realtà vs Sintetico

Diventa sempre più difficile per l'occhio umano distinguere ciò che è vero da ciò che è generato dall'AI.

Influenza sul Voto

Creazione massiva di contenuti polarizzanti e fake news mirate per alterare l'opinione pubblica.



Pattern di manipolazione emotiva rilevato dall'AI.

CYBERSECURITY

Norme europee: un’OPPORTUNITÀ

Non si tratta di limiti, ma di regole che abilitano una crescita sicura e sostenibile.

AI Act

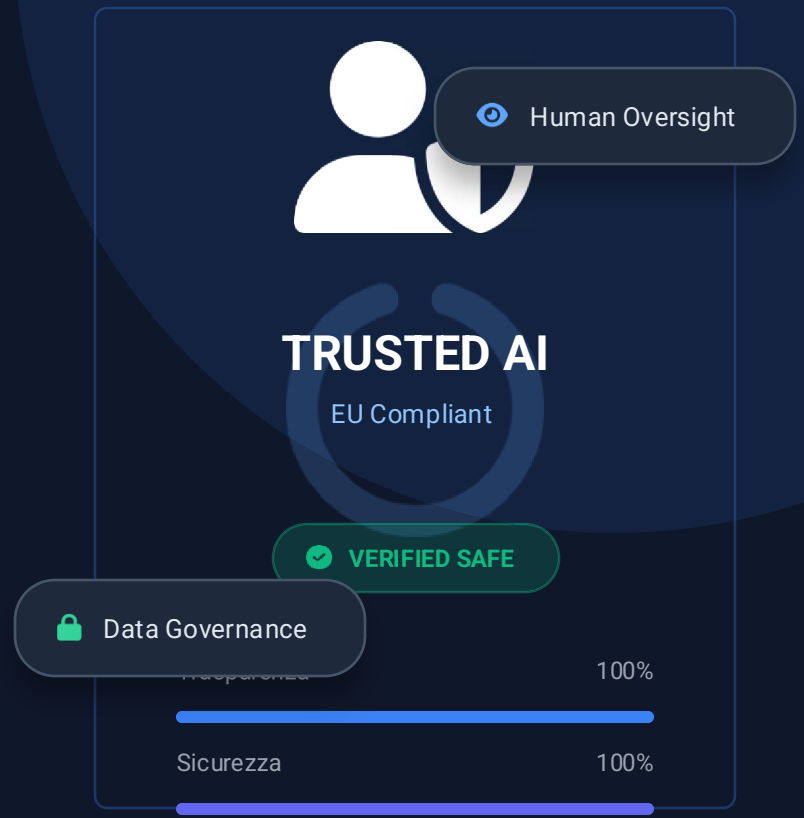
Il primo regolamento mondiale sull'IA. Definisce regole chiare per un uso sicuro, trasparente e rispettoso dei diritti fondamentali.

NIS2 Directive

Rafforza la sicurezza informatica delle infrastrutture critiche e delle aziende essenziali, proteggendo l'intera supply chain.

Vantaggio Competitivo

La conformità normativa genera fiducia nei consumatori e negli investitori, differenziando le aziende europee nel mercato globale.



La chiave: Collaborazione

Nessuno può vincere da solo: accademia, aziende e istituzioni devono fare sistema.

1

ECOSISTEMA

Linguaggio Comune

Accademia, imprese e istituzioni devono superare le barriere comunicative e allinearsi su obiettivi condivisi per una difesa efficace.

2

CRESCITA

Strumenti per PMI

Le Piccole e Medie Imprese sono vulnerabili. Servono risorse accessibili e competenze pratiche per permettere loro di difendersi.

3

VANTAGGIO

Abilitazione AI

L'Intelligenza Artificiale non è solo per i giganti: diventa un "moltiplicatore di forza" che aiuta anche chi ha risorse limitate.

Resilienza Collettiva: Un sistema interconnesso dove la sicurezza di uno rafforza la sicurezza di tutti.

CYBERSECURITY

La prima difesa: le PERSONE

La tecnologia da sola non basta. La consapevolezza degli utenti è l'arma più importante contro le minacce informatiche.

Consapevolezza

Capire che chiunque può essere un bersaglio. L'errore umano è la causa principale della maggior parte degli incidenti.

Formazione Continua

Le minacce evolvono rapidamente. La formazione costante permette di riconoscere anche gli attacchi più nuovi.

Attenzione e Vigilanza

Prestare attenzione ai dettagli: un indirizzo email sospetto o un link strano possono rivelare una truffa.

Sano Dubbio

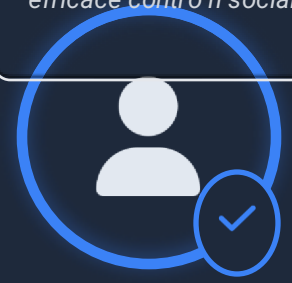
Verificare sempre prima di cliccare o autorizzare pagamenti, specialmente se la richiesta sembra urgente.

CONCETTO CHIAVE

HUMAN FIREWALL

ID: CFC-USER-2024

"L'utente consapevole diventa un 'Human Firewall', la barriera più efficace contro il social engineering."



Utente Verificato

PROTECTED

Rilevamento Phishing 98%

Attenzione ai Dettagli High

Resilienza Social Eng. Robust

Tecnologie emergenti per un'AI più sicura

Nuovi strumenti e architetture per mitigare i rischi, garantire l'affidabilità e proteggere i dati sensibili.

Sistemi RAG (Retrieval-Augmented)

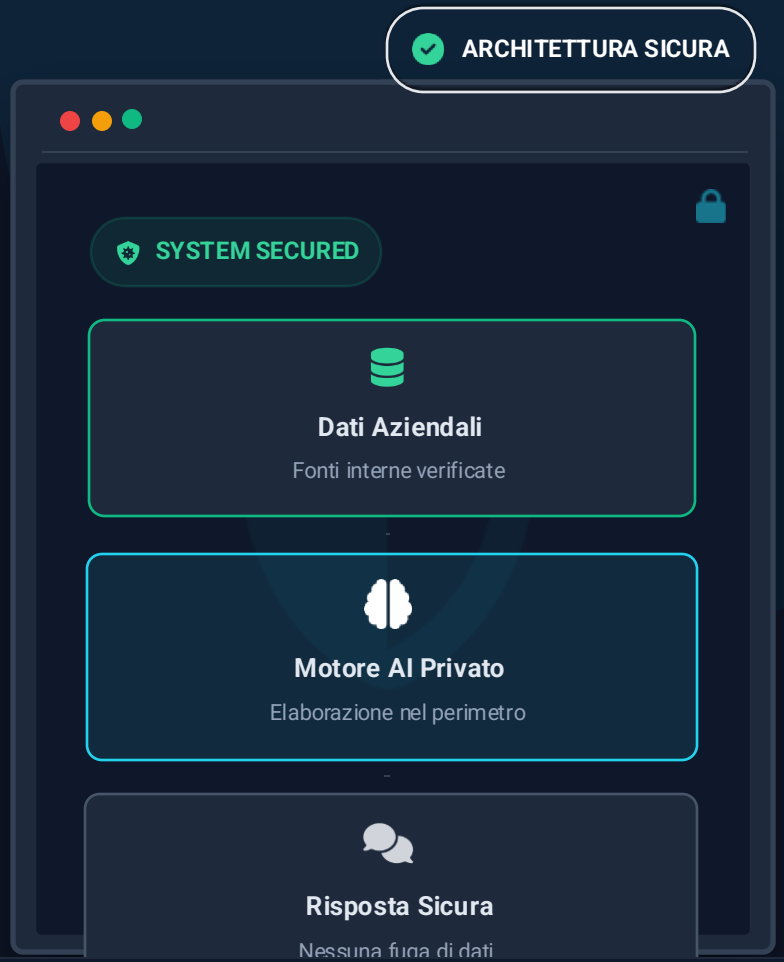
L'AI non inventa ma consulta una "biblioteca" di fonti aziendali affidabili prima di rispondere, riducendo le allucinazioni.

Agenti AI Aziendali

Assistenti virtuali specializzati, con permessi limitati e supervisionati, progettati per compiti specifici e confinati.

AI nel Perimetro Aziendale

Esecuzione di modelli (Enterprise LLM) all'interno dell'infrastruttura aziendale per garantire che i dati non escano mai.



CONCLUSIONI

L’AI: una potente alleata

"L'Intelligenza Artificiale è affascinante e potente. Può aiutare enormemente la società, contribuendo a trovare soluzioni a problemi complessi in medicina, ambiente ed energia."

Tuttavia, come ogni strumento rivoluzionario, il suo impatto positivo non è automatico.

I 4 PILASTRI PER UN'AI SICURA:

Consapevolezza

Capire cosa l'AI può fare e dove fallisce. Mantenere uno spirito critico.

Regole

Normative come l'AI Act per garantire trasparenza e sicurezza.

Collaborazione

Accademia, imprese e istituzioni devono lavorare insieme.

Responsabilità

Usare la tecnologia eticamente per il bene comune, non per nuocere.

Pronti a guidare il cambiamento?

IL FUTURO DIPENDE DA NOI

Tutelattiva

Per informazioni:
Via Matteo Pescatore 15, Torino (TO)
Tel 011 4346654
sportello@tutelattiva.it

ASSOCIAZIONE CONSUMATORI ACP
consumo lavoro cittadinanza

Per informazioni:
Via San Francesco D'Assisi 17, Torino (TO)
Tel 011 4366566
sportello@consumatoripiemonte.it