



Daitek

**Intelligenza Artificiale
e Trasparenza: come
rendere l'IA uno
strumento accessibile
a tutti**

GenOA week, 5 Novembre 2024

Aitek

- ▶ Dal 1986, Aitek si occupa di soluzioni tecnologiche innovative. Progettiamo sistemi intelligenti per il trasporto e la sicurezza e sviluppiamo applicazioni software basate su tecnologie digitali all'avanguardia.
- ▶ Il nostro portafoglio di prodotti supporta un alto numero di soluzioni ed è il risultato di oltre 35 anni di grandi commesse e di un impegno costante verso una tecnologia all'avanguardia.





1986
Data di nascita



60 dipendenti
60% laureati



2 Sedi
Genova e Savona



5,0 M€
Indotto 2023



1700+
Progetti



10% di ricavo
In ricerca e sviluppo

VIDEO SORVEGLIANZA
& VIDEO ANALISI



SISTEMI DI PEDAGGIO
& AUTOMAZIONE



ICT &
APPLICAZIONI WEB



DIGITAL SIGNAGE



L'**intelligenza artificiale** ha il potenziale per rivoluzionare le industrie, facendo passi da gigante in diverse applicazioni.

L'**analisi video** consente ai sistemi di analizzare, elaborare ed estrarre informazioni preziose dai contenuti video in tempo reale o in post-elaborazione.

Se combinati con l'intelligenza artificiale, i dispositivi **IoT** possono fornire preziose informazioni, automazione e miglioramento del processo decisionale in vari settori.

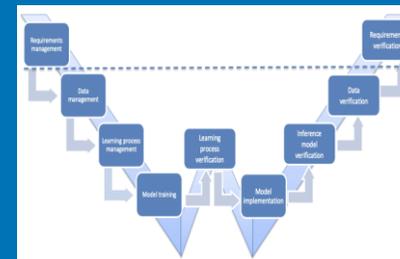
La **logistica** portuale mira a ottimizzare il flusso delle merci, a ridurre al minimo i ritardi e a migliorare l'efficienza complessiva delle operazioni portuali.

Display digitali, come schermi a LED, LCD o di proiezione, per veicolare informazioni personalizzate, **vendita al dettaglio e pubblicità** e una forma flessibile di comunicazione visiva.

AUTOMOTIVE & DRIVER MONITORING



SAFETY ASSESSMENT



MOBILITY & SMART CITY



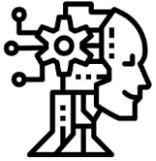
PEOPLE SECURITY ENSURANCE



SMART AGRICULTURE

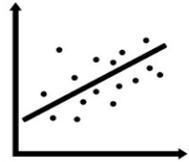


Definizioni



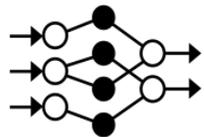
Intelligenza Artificiale

Tutte le tecniche che imitano l'intelligenza umana utilizzando la logica, il linguaggio computazionale e i modelli statistici.



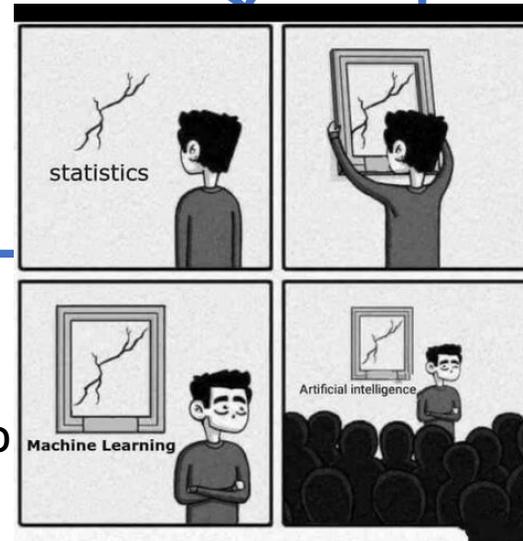
Machine Learning

Algoritmi per fare previsioni e prendere decisioni in condizioni di incertezza sulla base di dati osservati.



Deep Learning

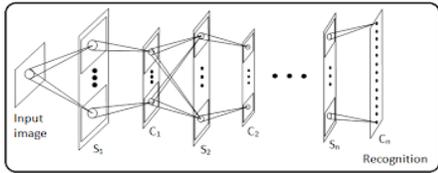
Modelli di ML costituiti da reti neurali profonde e multistrato che apprendono ed estraggono automaticamente modelli e rappresentazioni complesse da grandi quantità di dati.



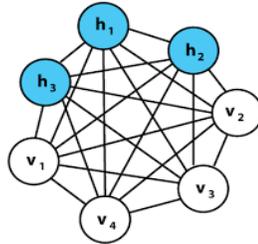
Data Science

Lo studio dei dati per estrarre informazioni utili a prendere decisioni basate sui dati. Si tratta di un approccio multidisciplinare che combina principi e pratiche di matematica, statistica e informatica per studiare grandi quantità di dati.

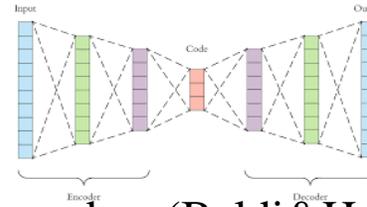
INTELLIGENZA ARTIFICIALE



Neocognitron
(Fukushima, 1980)



Boltzmann Machines
(Ackley et al., 1986)



Autoencoders (Baldi & Hornik, 1989) and
Convolutional NN
(LeCun, 1989)



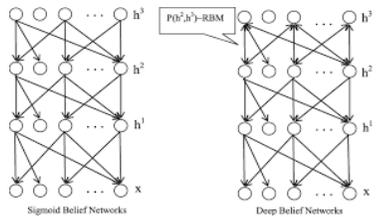
OpenAI
2021



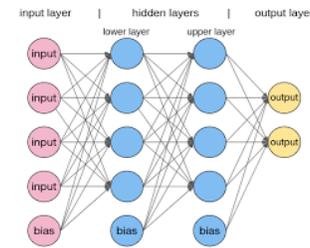
Sora
2024



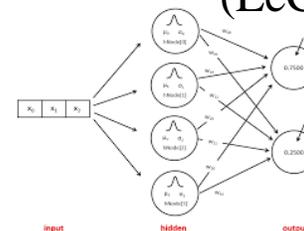
Rosenblatt's
Perceptron



Hopfield Network
(Kohonen, 1982)



Multilayer Perceptron and
Backpropagation
(Rumelhart et al., 1986)



RBF networks
(Broomhead & Lowe, 1988)



Transformers and
Attention Layers, 2017



LLAMA-META
2023



Meta
2023

1958

1980

1982

1986

1988

1992

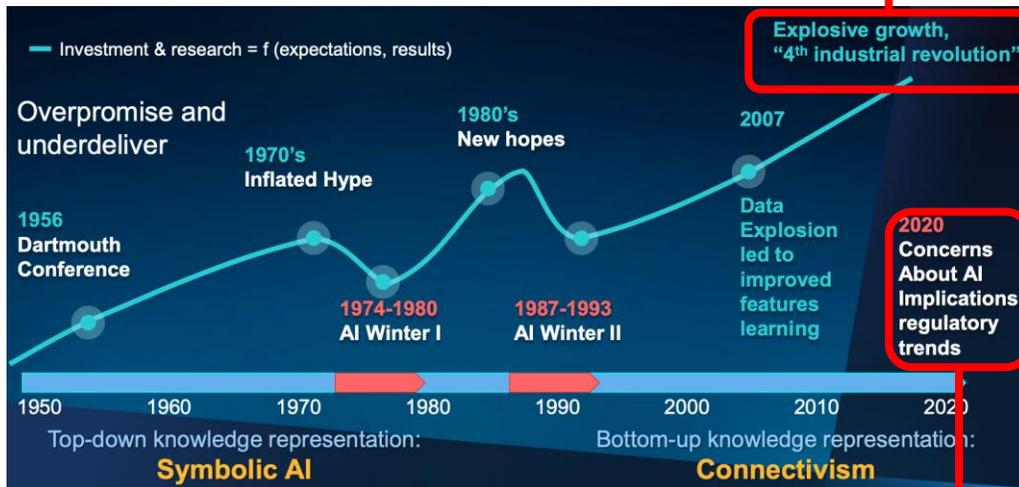
1993

2000

2024

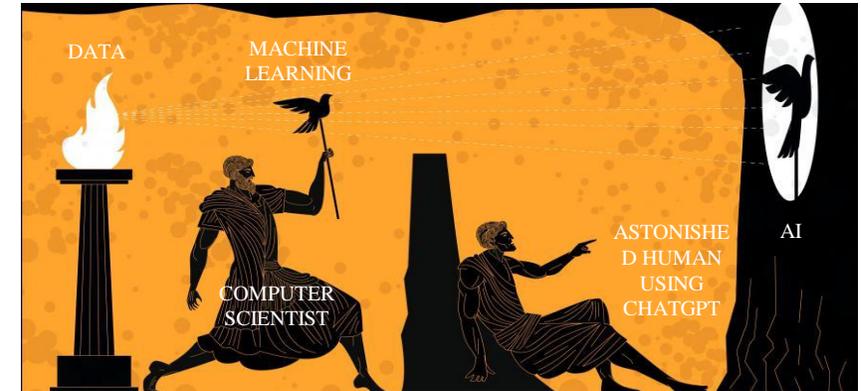
La direzione dell'IA

Gli Inverni dell'IA



Credit: Clive Thompson, Medium

Stiamo vivendo il "Mito della Caverna" ?



"AI Act" e linee guida etiche dell'Europa

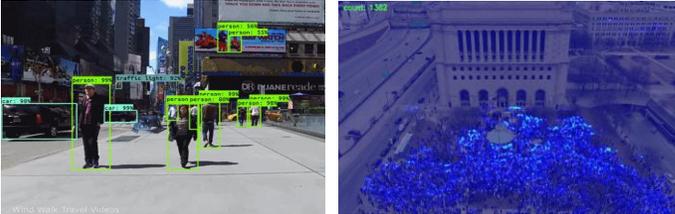


L'IA affidabile (trustworthy) deve essere:

- **Legale** - rispettando tutte le leggi e i regolamenti applicabili.
- **Etica** - rispettando i principi e i valori etici.
- **Robusta** - sia da un punto di vista tecnico che tenendo conto del suo ambiente sociale.

I livelli di rischio dell'IA

EU Artificial Intelligence Act: Risk levels



Sicurezza



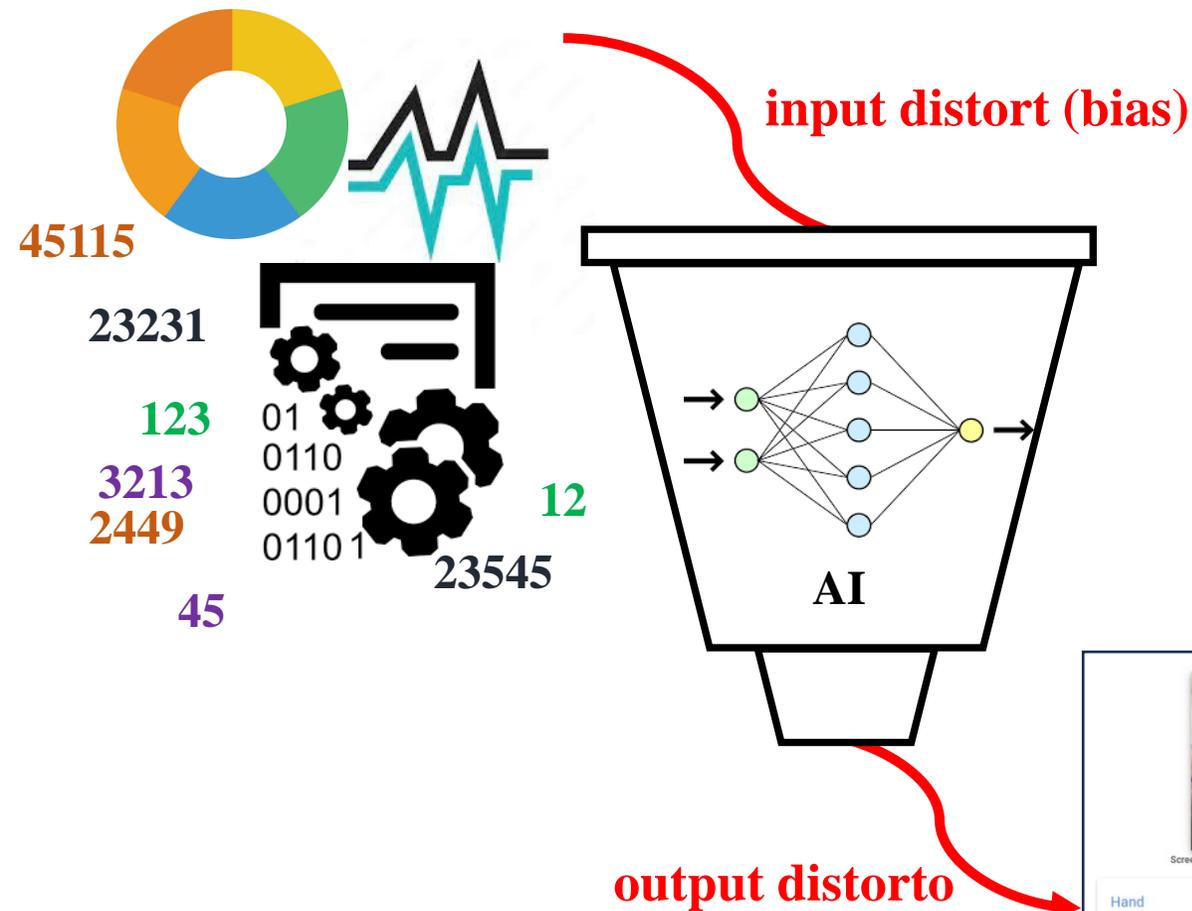
Sistemi Autonomi



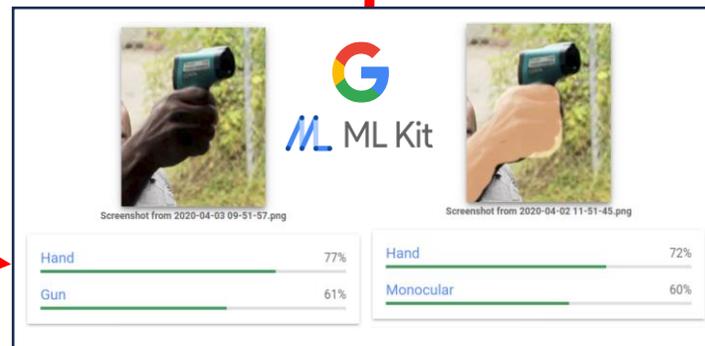
Retail

Credit: RAQUEL CARRETERO JUAREZ

Perchè l'IA non è equa?



Paradosso di Simpson:
Evidenze statistiche mostravano che persone appartenenti alla comunità afroamericana erano più probabili di essere criminali...ma questo solo perchè i dati contenevano più dati di detenuti afroamericani che di uomini bianchi



Le "allucinazioni" dell'IA

Estratto dallo «European Data Protection Supervisor»

LLMs, if trained on biased data, could perpetuate or even amplify biases present in their training data. This might lead to unfair or discriminatory outputs, potentially violating the principle of fair processing of personal data.

«Hallucinations»



Un video generato da SORA in cui il prompt era di generare un uomo che corre su un tapis roulant.

Stanno aumentando le tecniche per gestire e mitigare queste problematiche («Uncertainty Quantification»)

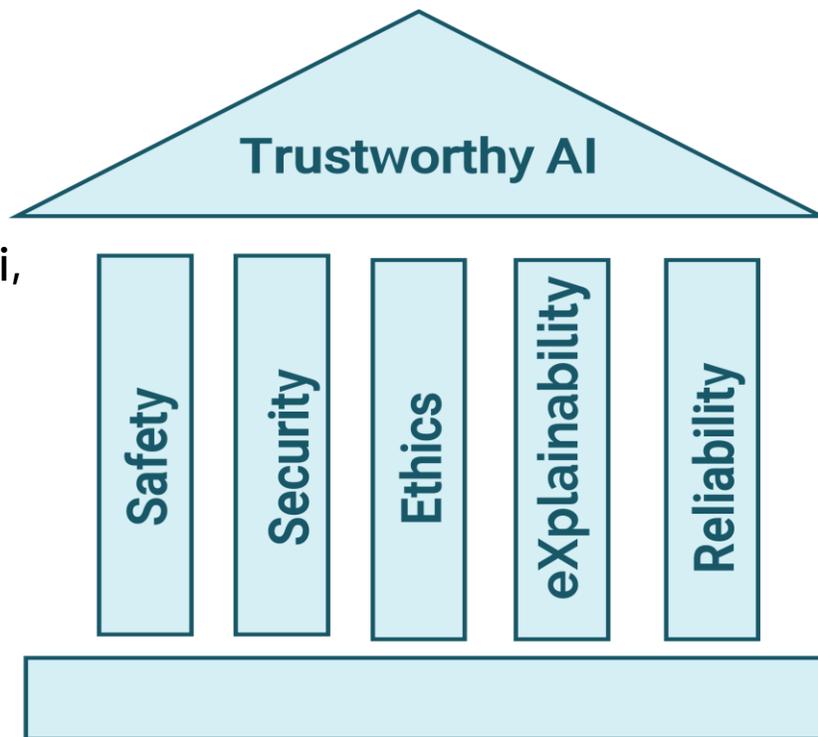
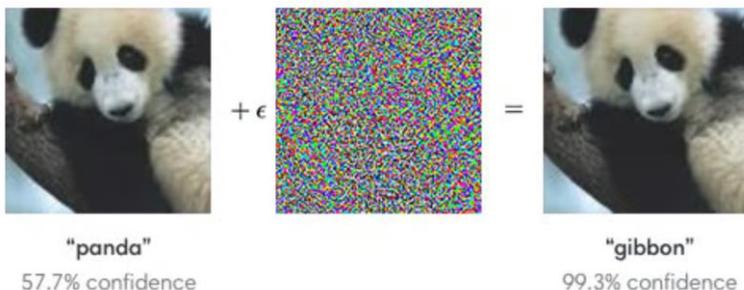


Conformal Prediction



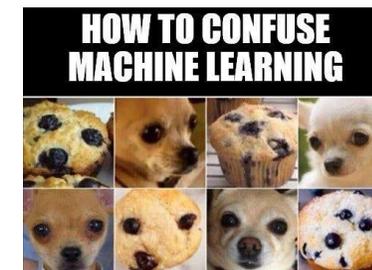
Trustworthy AI – IA affidabile

- I sistemi di intelligenza artificiale devono prevenire le conseguenze indesiderate e i potenziali pregiudizi, **riducendo al minimo i rischi.**

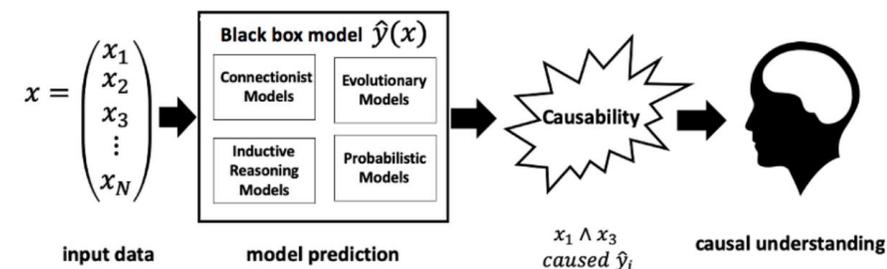


- I sistemi di IA devono essere **equi e coerenti** con le normative.
- I sistemi di IA devono prevedere procedure di **sicurezza dei dati e di protezione della privacy.**

- I sistemi di intelligenza artificiale dovrebbero comportarsi come ci si aspetta, **producendo risultati solidi.**



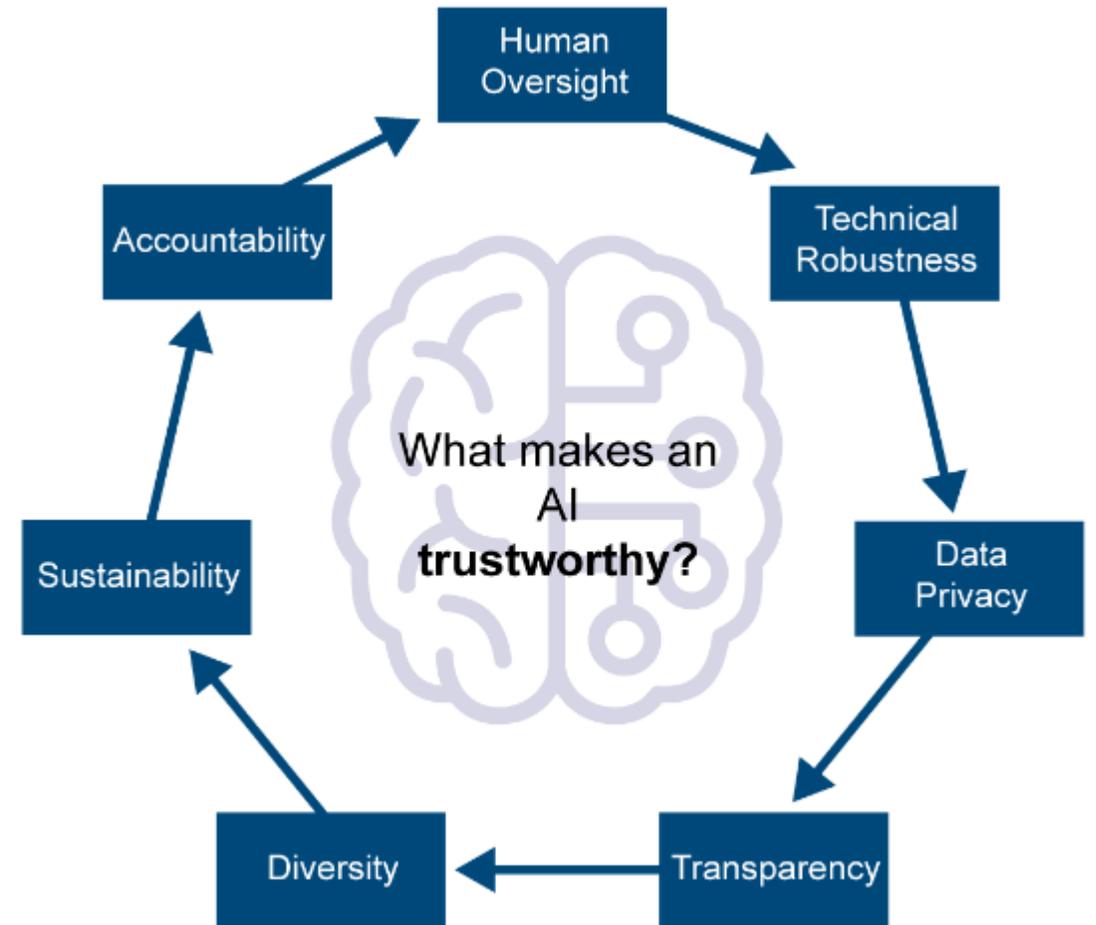
- I sistemi di intelligenza artificiale dovrebbero essere **comprensibili a tutti.**



Trustworthy AI – IA affidabile

Le linee guida sull'etica hanno introdotto il concetto di IA affidabile, basato su sette requisiti chiave:

1. Agenzia e supervisione umana
2. Solidità tecnica e sicurezza
3. Privacy e governance dei dati
4. Trasparenza
5. Diversità, non discriminazione ed equità
6. Benessere ambientale e sociale e
7. Responsabilità



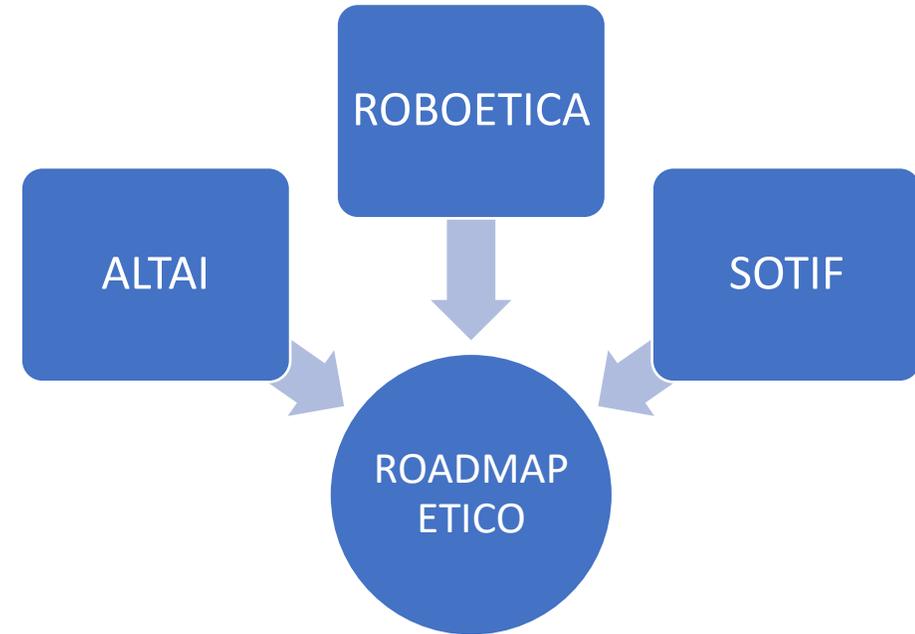
RoadMap etico

Roboetica:

La priorità dei sistemi di intelligenza artificiale deve essere la sicurezza degli utenti finali. In caso di movimenti autonomi, i componenti devono sempre permettere all'utente di riprendere il controllo, passando dalla modalità automatica alla modalità di guida manuale. L'essere umano deve essere "in the loop" e "on the loop".

Safety Of The Intended Functionalities (SOTIF):

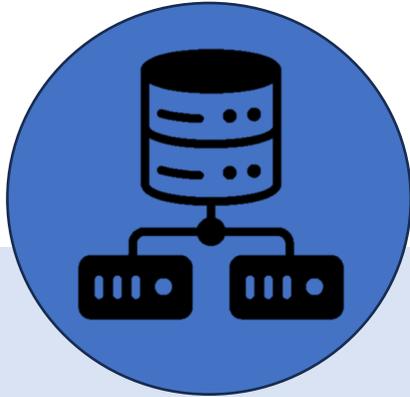
- Per ciascuna funzione intenzionale:
 - Elencare le condizioni di attivazione;
 - Elencare i punti deboli/limitazioni delle prestazioni del componente;
 - Elencare i comportamenti di malfunzionamento;
- Definire i pericoli e il Dominio Operativo Progettuale (ODD);
- Per ciascun pericolo, assegnare l'esposizione, la gravità e la controllabilità;
- Calcolare il Livello di Integrità della Sicurezza (SIL).



ALTAI:

- Garantire che il set di dati di addestramento e il set di dati di test siano privi di modifiche o alterazioni dannose;
- Considerare l'uso di tecnologie di spiegabilità per analizzare il comportamento del modello di output;
- Ricercare possibili dati che influenzano la stabilità;
- Generare modelli multipli di diverse dimensioni e simulazioni.

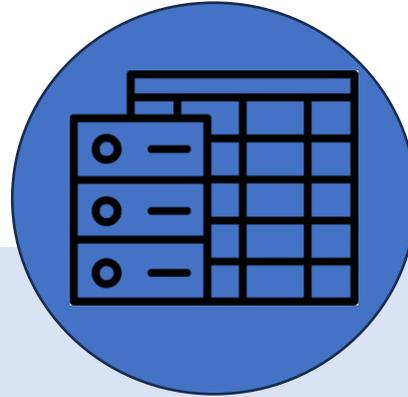
Come spiegare l'IA?



Data explainability

Che tipo di dati vengono utilizzati per addestrare l'IA?

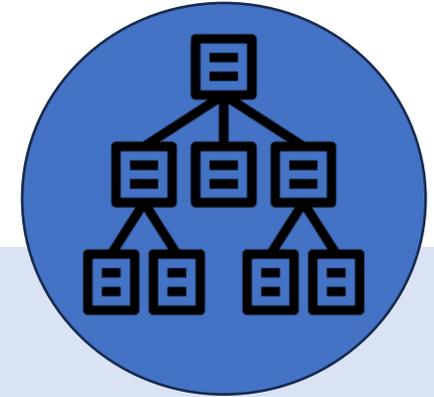
Quali caratteristiche hanno i dati?



Variables explainability

Quali caratteristiche sono importanti per determinare una certa classe?

Quali caratteristiche possono portare a una classificazione errata?



Output explainability

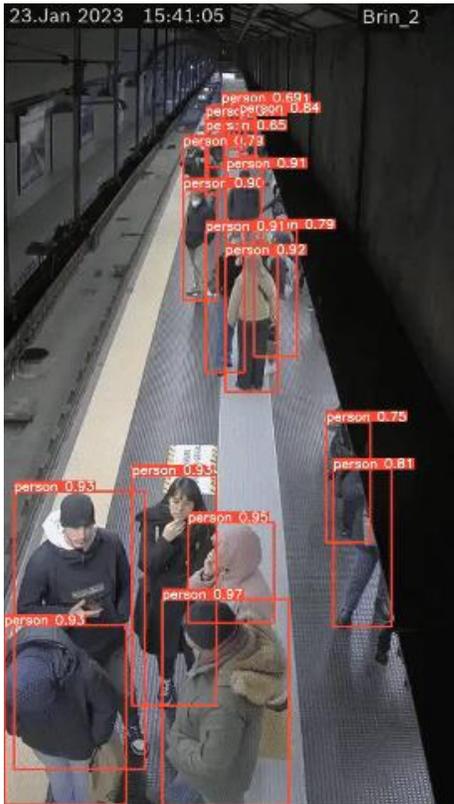
Qual è il ragionamento che porta a una certa decisione?

Come possiamo guidare questo ragionamento?



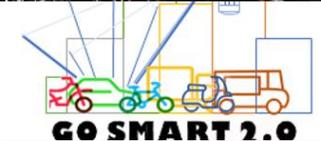
Esempi di IA etica e accessibile

IA per una mobilità più accessibile



Obiettivi principali

- ▶ Migliorare la gestione e l'ottimizzazione del trasporto pubblico
- ▶ Ridurre il traffico
- ▶ Garantire la sicurezza e la privacy degli utenti
- ▶ Segnalare i parcheggi liberi

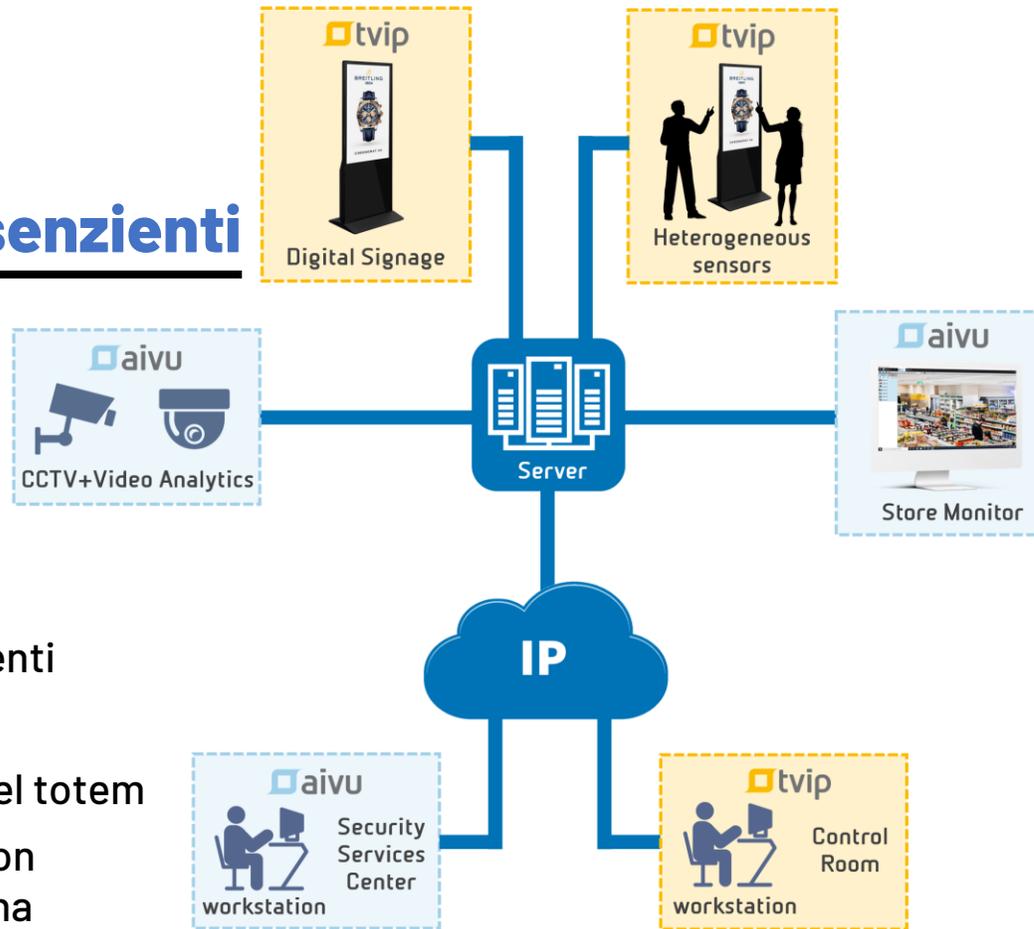


IA al servizio degli utenti

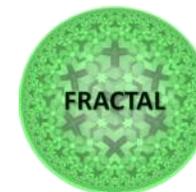
tvip aivu

Soluzione per negozi senziienti

- ▶ Sicurezza video
- ▶ Analisi video per:
 - ▶ Conteggio degli utenti,
 - ▶ Riconoscimento dei comportamenti,
 - ▶ Tracciamento dei movimenti
- ▶ Digital Signage Dinamico:
 - ▶ Marketing in prossimità del totem
 - ▶ Applicazioni «in-Store» con interazioni uomo-macchina (touchscreen / voce / gestualità)



- ▶ Totem intelligente da supporto personalizzato in base agli utenti
- ▶ Sensori eterogenei
- ▶ Tecnologia EDGE-AI distribuita

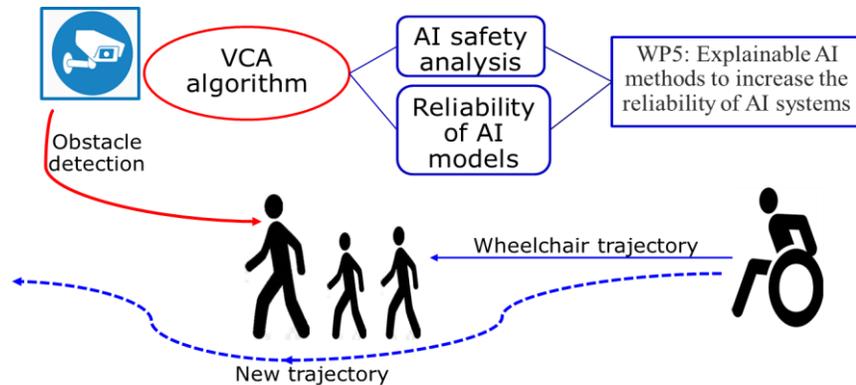


Analisi dei rischi dell'IA

Analisi dei rischi

- ▶ Valutazione della sicurezza del Sistema globale per garantire la safety e la security delle persone.
- ▶ Migliorare l'affidabilità dell'IA usando metodi statistici che siano spiegabili.
- ▶ Cercare di spiegare il ragionamento dell'IA, andando a capire perché possono esserci dei bias o dei FN-FP.

Questo processo iterativo consente di migliorare le prestazioni e l'affidabilità del modello di IA.



Reliable & Explainable Swarm Intelligence for People with Reduced Mobility



DOMANDE ?

 **Aitek S.p.A.**

Via della Crocetta, 15
16122 Genova - Italy

Via Nizza, 28/2
17100 Savona - Italy

 +39 010 846731

 info@aitek.it

 www.aitek.it

