

Le infrastrutture tecnologiche

Fabrizio Pirri

Politecnico di Torino e Istituto Italiano di Tecnologia



ISTITUTO ITALIANO
DI TECNOLOGIA
CENTRE FOR SUSTAINABLE
FUTURE TECHNOLOGIES



Politecnico
di Torino

INTELLIGENZA ARTIFICIALE DAL DATO ALLA SOLUZIONE DI UN PROBLEMA



Attraverso la tecnologia si raccolgono informazioni da sistemi online e fisici su ogni aspetto della realtà che ci circonda.

Disponiamo di dati di testo, audio, video e immagini in grandi quantità.

La data science combina strumenti, metodi e tecnologie statistici per generare un significato dai dati.

L'intelligenza artificiale utilizza i dati per risolvere i problemi cognitivi (apprendimento, riconoscimento di schemi e espressione simile a quella umana).

È fatta di algoritmi complessi che "apprendono" man mano che procedono, migliorando nel tempo la risoluzione dei problemi.

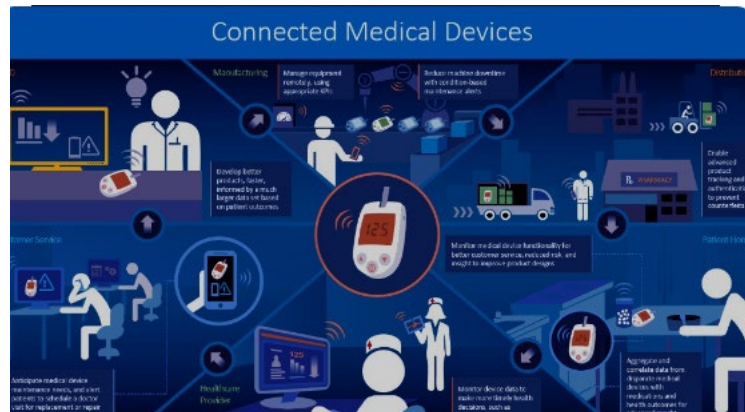
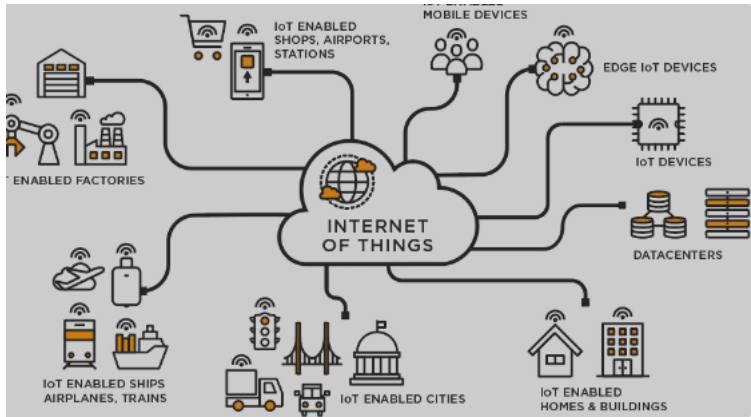
INTELLIGENZA ARTIFICIALE NON E' SOLO SOFTWARE, INTELLIGENZA ARTIFICIALE E' INFRASTRUTTURE

L'intelligenza artificiale poggia su infrastrutture hardware di produzione dei dati, di stoccaggio dei dati, di elaborazione dei dati.

Le infrastrutture hardware sono prodotte in fab di grandi dimensioni che necessitano di infrastrutture di produzione e di ricerca.

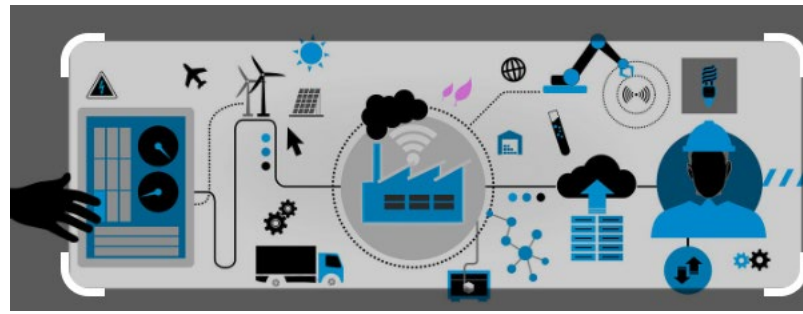


IL DATO NECESSITA DI INFRASTRUTTURE SENSORIALI SEMPRE PIU' COMPLESSE: IoT





ISTITUTO ITALIANO
DI TECNOLOGIA
CENTRE FOR SUSTAINABLE
FUTURE TECHNOLOGIES



Politecnico
di Torino

Quando Artificial Intelligence e IoT si incontrano

Internet of Things (IoT) e Intelligenza Artificiale (AI) sono due importanti tecnologie protagoniste in diversi ambiti di applicazione.

La combinazione di IoT e AI consente di parlare di “AIoT”.

L'intelligenza artificiale sfrutta i Big Data, ma è necessario che queste risorse siano affidabili e sempre disponibili.

Le tecnologie IoT sono in grado di raccogliere, aggregare, analizzare e fornire modelli predittivi sfruttando un sistema di piattaforme e device complessi.

L'utilizzo congiunto di IoT e AI consente di incrementare il valore reciproco delle due soluzioni: da una parte l'AI aumenta le potenzialità dell'IoT implementando l'apprendimento automatico delle macchine e dei dispositivi; dall'altra l'IoT aumenta il valore dell'AI fornendo risorse in termini di connettività e scambio di dati.

Intelligenza artificiale, quali sono le architetture hardware dedicate

Intelligenza artificiale (IA) si occupa della formalizzazione di problemi e compiti simili a quelli eseguiti da una persona.

IA si occupa di studiare tutte le tecniche e le metodologie che permettono di progettare i sistemi hardware e software in grado di elaborare dati e prendere decisioni che simulano l'intelligenza umana.

Idealmente il risultato del lavoro dell'intelligenza artificiale non dovrebbe essere distinguibile da quello svolto da un umano con delle specifiche competenze.

L'hardware per la AI deve accelerare il training e le prestazioni delle reti neurali, limitando nel contempo i problemi relativi a dimensioni, consumi energetici e costi.

L'hardware per i sistemi di Intelligenza Artificiale si basa su **CPU** (Central Processing Unit), **GPU** (Graphics Processing Unit), **TPU** (Tensor Processing Unit) e **FPGA** (Field Programmable Gate Array).

La CPU esegue i programmi che non richiedono un'intensa capacità di calcolo parallelo.

Le GPU gestiscono molti compiti contemporaneamente, rendendole ideali per grandi quantità di calcoli matematici simultanei.

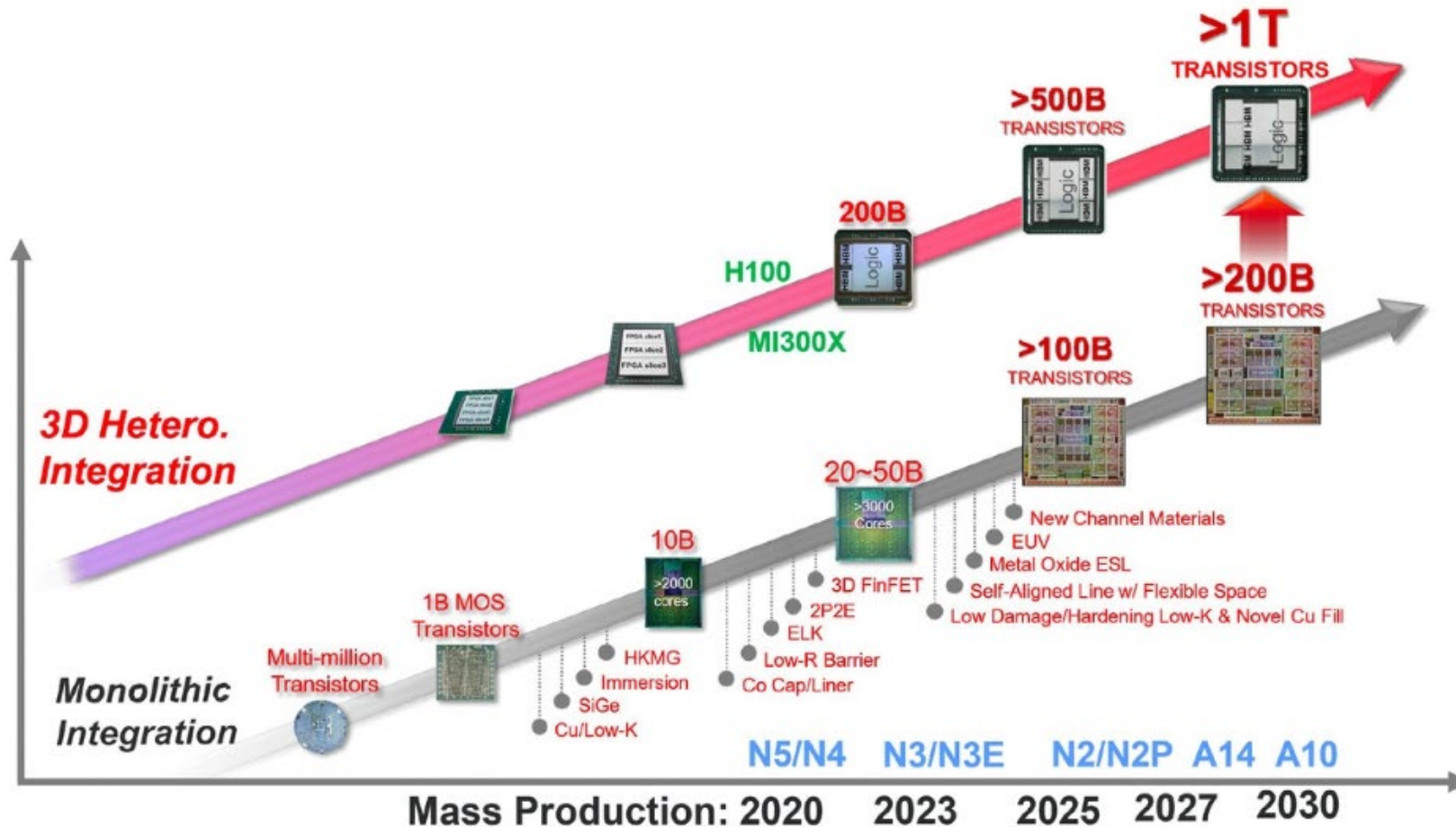
Le TPU sono state sviluppate specificamente per il machine learning.

Le FPGA sono chip che possono essere configurati per svolgere compiti specifici.

La base di tutto l'hardware per IA è il **TRANSISTOR**.

Attualmente, l'unità di calcolo monolitica più complessa sul mercato è la GPU NVIDIA GH100 utilizzata nell'ambito dell'intelligenza artificiale. **Il chip integra 80 miliardi di transistor.**

All'International Electron Devices Meeting (IEDM) si è discusso di un percorso che **nel 2030 dovrebbe portare a soluzioni con design multi-chiplet che integrano fino a 1 trilione di transistor (mille miliardi) e chip monolitici da 200 miliardi di transistor.**



I processori più avanzati per la realizzazione dei transistor hanno nodi tecnologici da 14-10-7 nm.

Si prevede una discesa a 2 nanometri nel 2025 e a 1,4 nanometri nel 2027



Uno scanner litografico di ASML





Politecnico
di Torino

IA e gemello digitale in sanità:

PNRR progetto D3 4H

L'IA amplifica le capacità cognitive del medico con l'analisi di una mole enorme di dati. Il gemello digitale umano consente la creazione di un modello virtuale del paziente che integra dati medici, comportamentali e genetici, per una più diagnosi con la personalizzazione delle cure.



IA e gemello digitale nel settore distribuzione e consumo dell'energia.



Nuovi materiali e molecole: laboratori autonomi guidati da IA ci aiuteranno a trovarli





ISTITUTO ITALIANO
DI TECNOLOGIA
CENTRE FOR SUSTAINABLE
FUTURE TECHNOLOGIES



Politecnico
di Torino

Ma c'è un problema. L'IA consuma enormi quantità di elettricità.

Secondo uno scienziato olandese, Alex De Vries, che ha calcolato il potenziale consumo di elettricità della tecnologia AI a livello globale, potrebbe raggiungere l'incredibile cifra di 85 terrawattora. Ogni anno. E questo è l'estremo inferiore della gamma. Il limite superiore è di 134 terrawattora o 134 miliardi di kilowattora. "Si tratterebbe delle dimensioni di un Paese come i Paesi Bassi in termini di consumo di elettricità.

"L'impronta di CO₂ e il consumo di energia andranno di pari passo con la potenza di calcolo, perché fondamentalmente questi data center vengono alimentati in modo proporzionale alla quantità di calcoli effettuati", afferma [Junchen Jiang](#), ricercatore dell'Università di Chicago.

Oltre alla grande quantità di energia, i centri dati che addestrano e gestiscono i modelli di AI generativa **consumano milioni di litri d'acqua.**



ISTITUTO ITALIANO
DI TECNOLOGIA
CENTRE FOR SUSTAINABLE
FUTURE TECHNOLOGIES



Politecnico
di Torino

Inoltre i costi attuali della tecnologia.....

Per creare un'intelligenza artificiale da zero è necessario seguire un processo complesso e costoso che richiede l'utilizzo di tecnologie avanzate come il machine learning e il deep learning.

Il procedimento è il seguente, ed ognuno dei punti presenta dei costi:

1. Raccolta di grandi quantità di dati
2. Classificazione dei dati per renderli utilizzabili dall'algoritmo di apprendimento
3. Progettazione e sviluppo del modello di machine learning che verrà poi addestrato sui dati raccolti e classificati
4. Addestramento e gestione del modello con dati finché non è soddisfacente nei suoi risultati
5. Implementazione ed esposizione del modello con le sue funzionalità per poterlo utilizzare
6. Implementazione della connessione con nuovi programmi o con programmi già esistenti

I costi da mettere a budget sono:

- Hardware di altissimo livello in quanto le operazioni da fare sfruttano calcolo parallelo e sono moltissime
- Data Analyst per tutto ciò che riguarda i dati
- Esperto di Machine Learning per la realizzazione del modello
- Software house per l'implementazione delle interfacce utente e l'integrazione con gli altri software



ISTITUTO ITALIANO
DI TECNOLOGIA
CENTRE FOR SUSTAINABLE
FUTURE TECHNOLOGIES



Politecnico
di Torino

La continua ricerca e lo sviluppo nel campo delle nanotecnologie e della scienza dei materiali hanno aperto la strada a transistor microscopici e a CPU più potenti. Il calcolo quantistico e altre tecnologie emergenti hanno il potenziale per sviluppare ulteriormente la tecnologia dei processori per computer.

“Calcolo quantistico”: un concetto che si basa sul mondo della fisica delle particelle e a principi come quello di indeterminazione di Heisenberg.

Quando parliamo di utilizzo di sistemi quantistici non parliamo di fisica delle particelle ma siamo nell’ambito della computabilità, ossia della potenza di calcolo di un computer. Il **“quantum computing” è una tecnologia emergente e trasformativa che utilizza i principi della meccanica quantistica per risolvere problemi troppo complessi per essere risolti dai computer classici.**

Il calcolo super accelerato quantistico integra supercomputer tradizionali con computer quantistici, sfruttando l'intelligenza artificiale per superare limiti fisici e algoritmici. L'AI migliora il controllo dei qubit, essenziali nelle unità di elaborazione quantistica (QPU), e sviluppa algoritmi che massimizzano l'efficacia dei computer quantistici. La combinazione di AI e calcolo quantistico potenzia la ricerca e le applicazioni in settori critici come medicina e ottimizzazione industriale, prospettando un futuro in cui supercomputer e computer quantistici collaborano per risolvere problemi complessi con una precisione e velocità senza precedenti. Questo sviluppo richiede progressi nella gestione dei qubit e nella precisione delle operazioni quantistiche, con un focus su nuovi linguaggi di programmazione e strumenti di ottimizzazione.