

Glossario tecnico sull'Intelligenza Artificiale



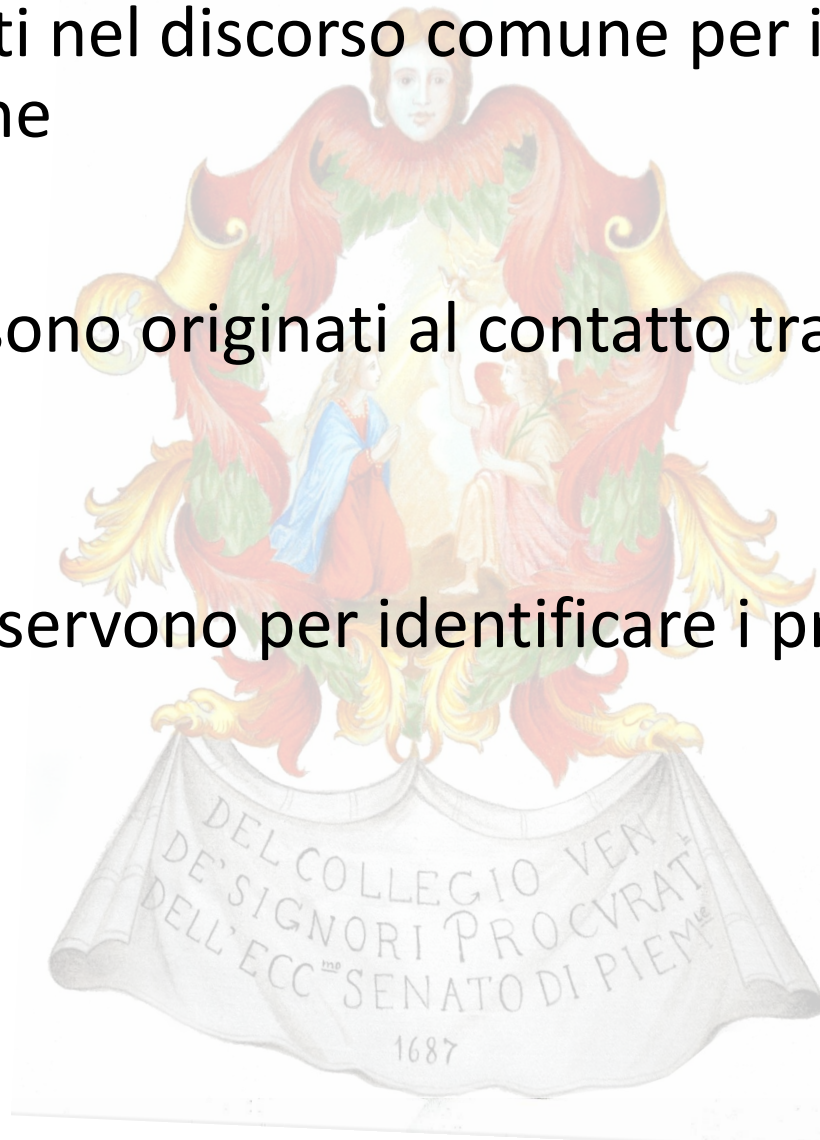
**Intelligenza artificiale e
discriminazione di genere**

Torino 21/06/2023

Germana Bertoli

germana.bertoli@studiolegalebertoli.it

- Termini tecnici entrati nel discorso comune per indicare la nuova tecnologia e le sue caratteristiche
- Termini ibridi che si sono originati al contatto tra nuove tecnologie e ambiti applicativi
- Termini giuridici che servono per identificare i problemi posti dalle nuove tecnologie



Termini «nuovi» nelle nostre discussioni

- Termini tecnici entrati nel discorso comune per indicare la nuova tecnologia e le sue caratteristiche



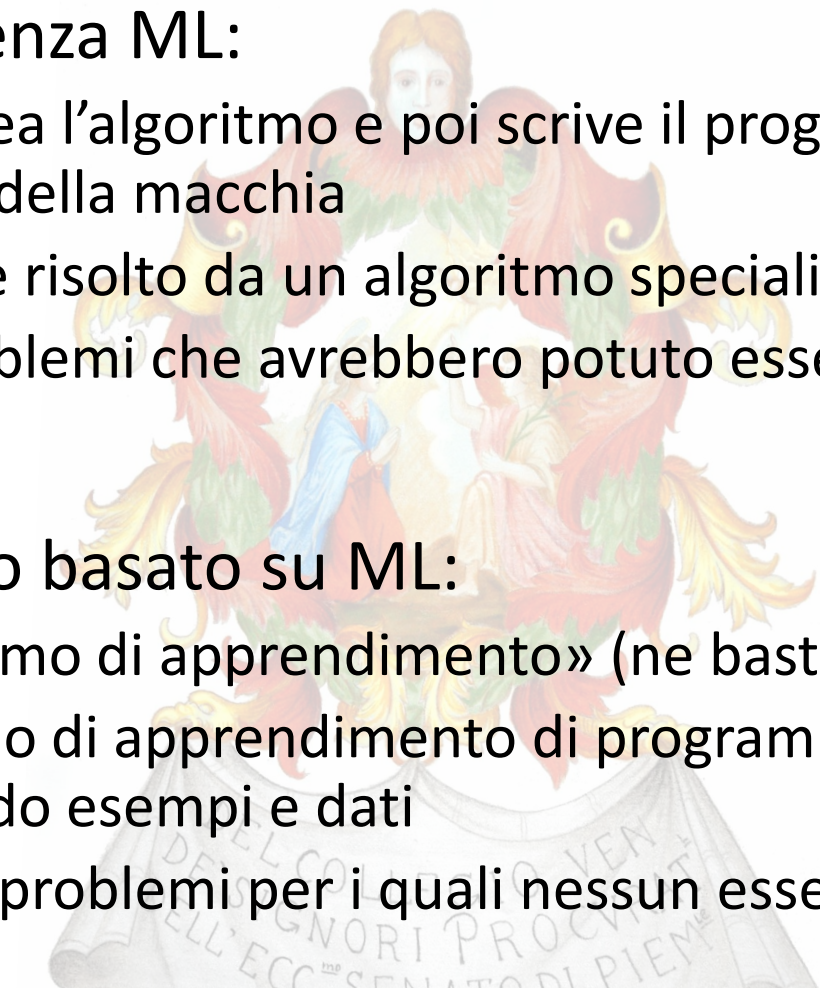
Termini «nuovi» nelle nostre discussioni

Sforzo scientifico/tecnologico per rendere i computer capaci di svolgere compiti per i quali si ritiene necessaria intelligenza (umana)

11 marzo 1997
Deep Blue (IBM) vince la 6ª e decisiva partita contro l'allora campione mondiale Garry Kasparov

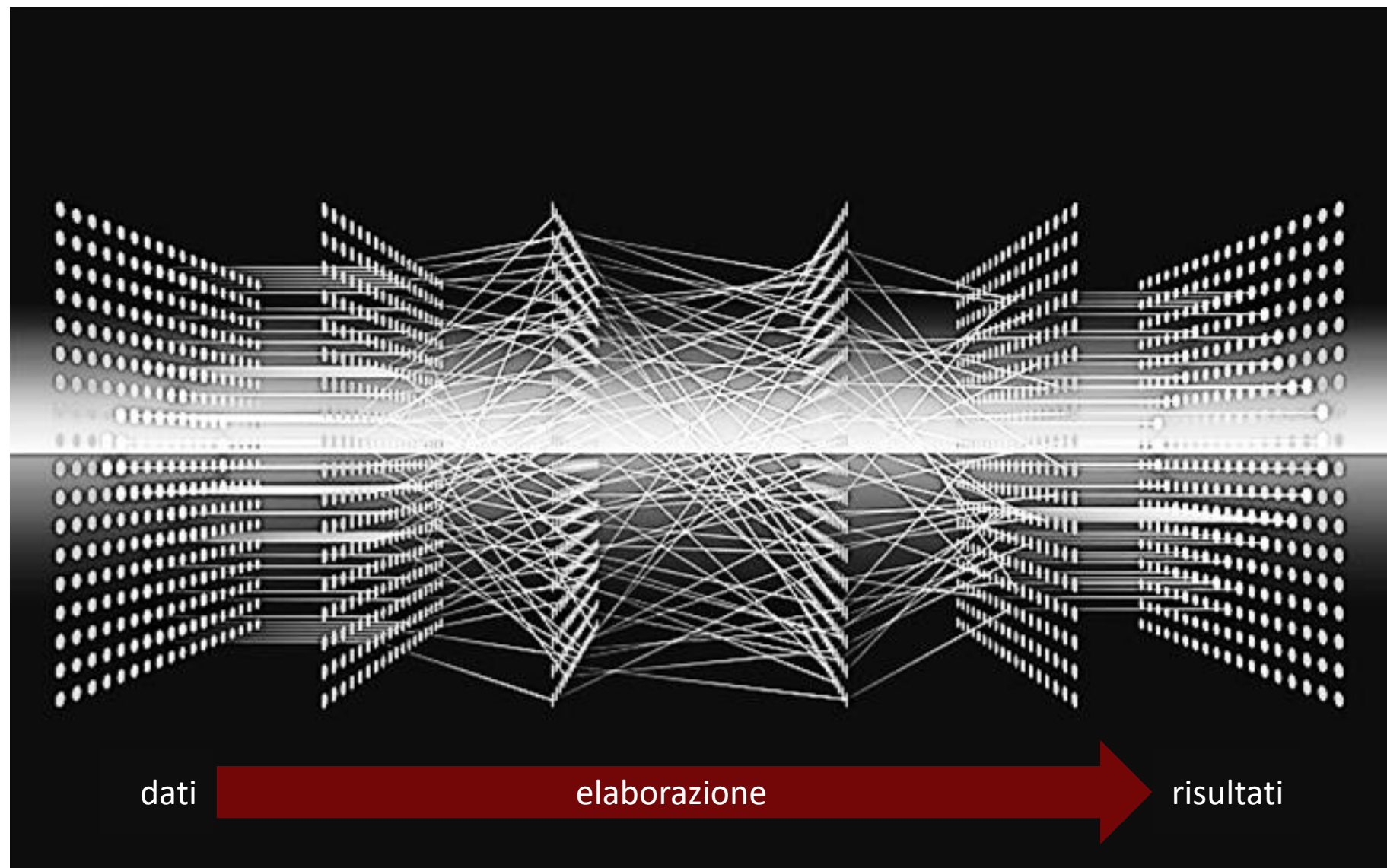


Intelligenza artificiale (AI – Artificial Intelligence)

- 
- Approccio classico senza ML:
 - Un essere umano crea l'algoritmo e poi scrive il programma che rende eseguibile l'algoritmo da parte della macchina
 - Ogni problema viene risolto da un algoritmo specializzato
 - Si risolvono solo problemi che avrebbero potuto essere risolti da un essere umano
 - Approccio alternativo basato su ML:
 - Si sceglie un «algoritmo di apprendimento» (ne bastano un numero limitato)
 - Si chiede all'algoritmo di apprendimento di programmare un computer per risolvere il problema esaminando esempi e dati
 - Si possono risolvere problemi per i quali nessun essere umano sa/saprebbe trovare un algoritmo

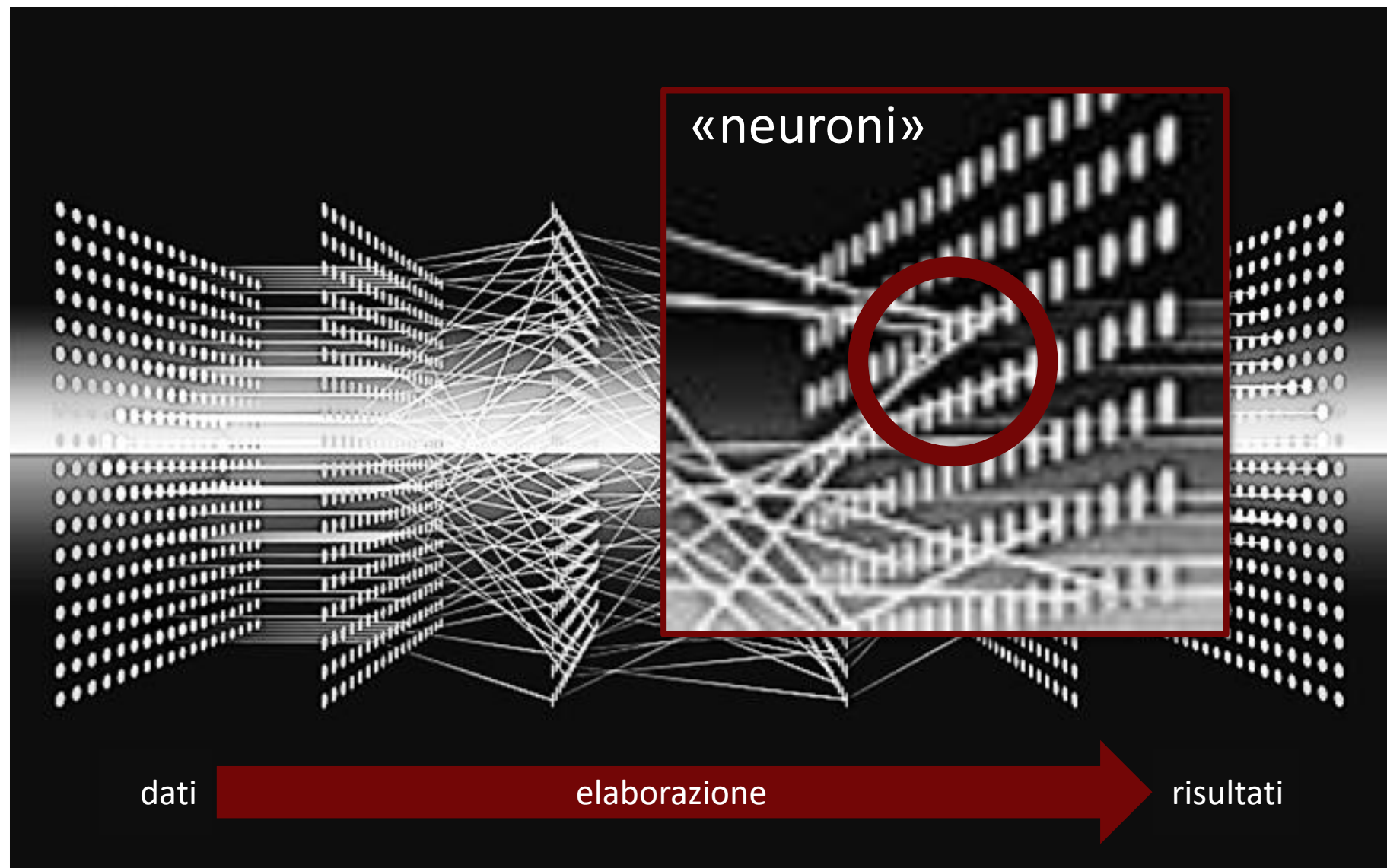
Approcci per risolvere i problemi: classico e basato su ML

ML tramite reti neurali profonde, sistemi di complessità che si avvicina a quella del cervello umano



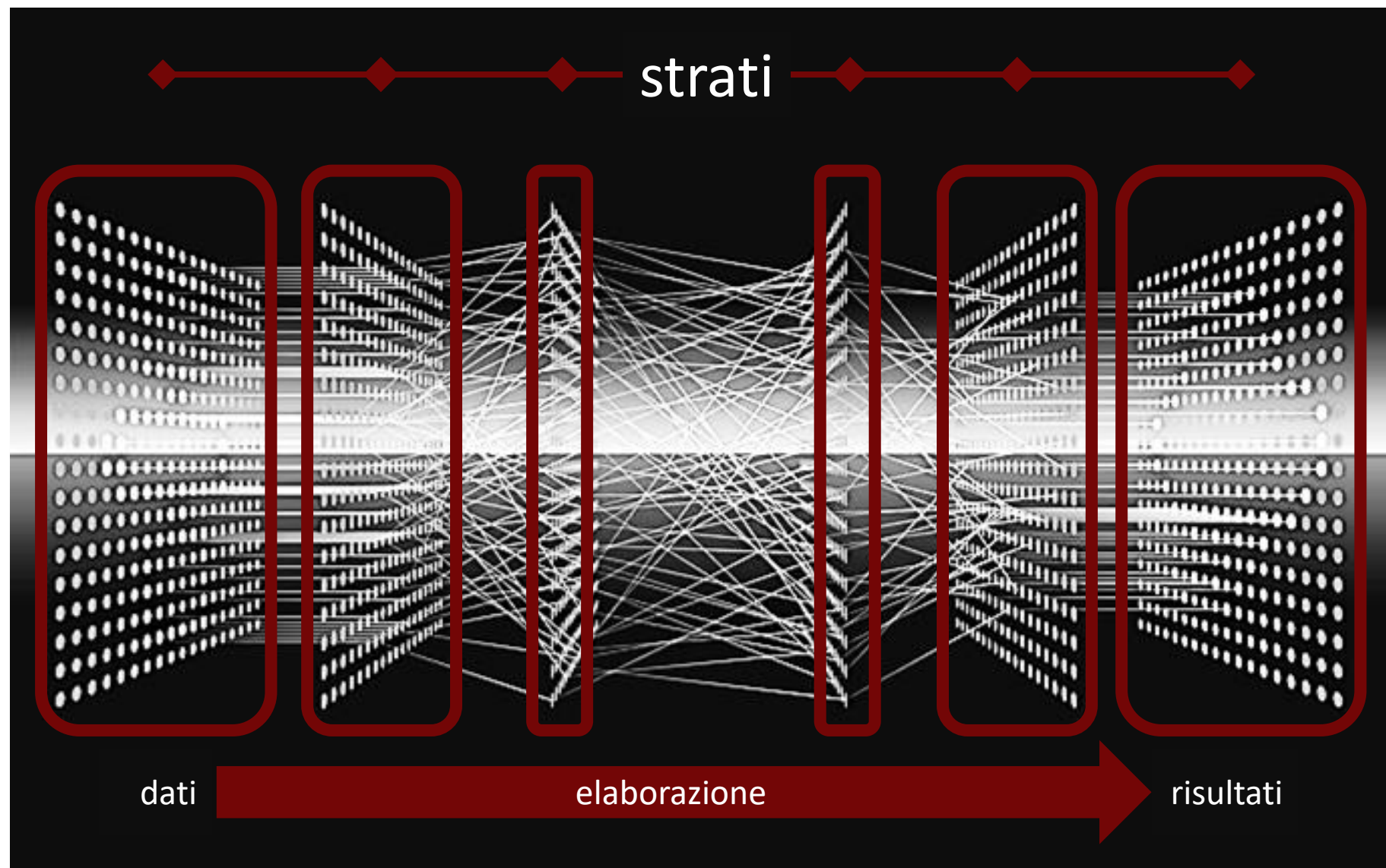
Apprendimento profondo (DL – deep learning)
Reti neurali profonde (DNN – deep neural network)

ML tramite reti neurali profonde, sistemi di complessità che si avvicina a quella del cervello umano



Apprendimento profondo (DL – deep learning)
Reti neurali profonde (DNN – deep neural network)

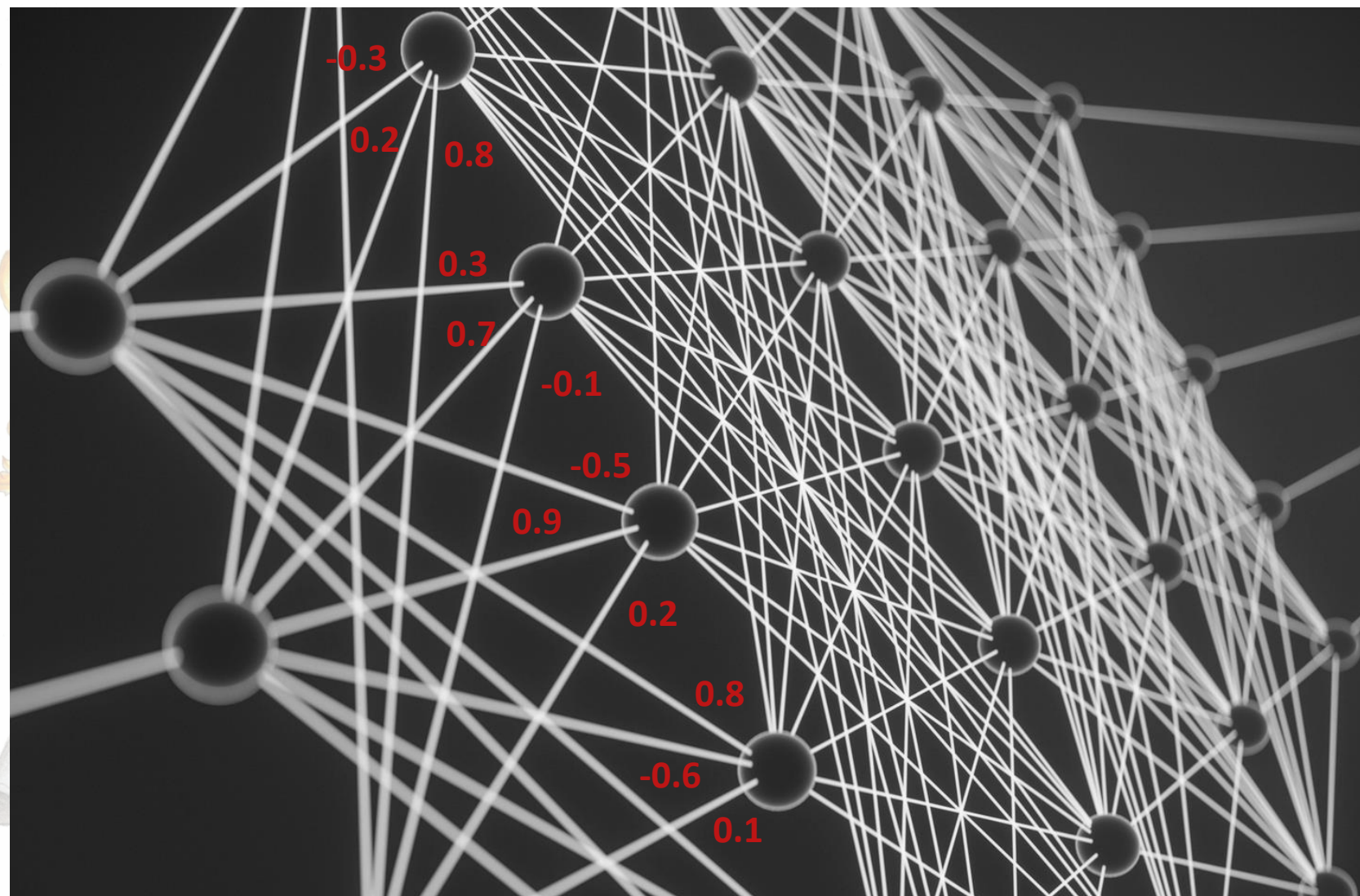
ML tramite reti neurali profonde, sistemi di complessità che si avvicina a quella del cervello umano



Apprendimento profondo (DL – deep learning)

Reti neurali profonde (DNN – deep neural network)

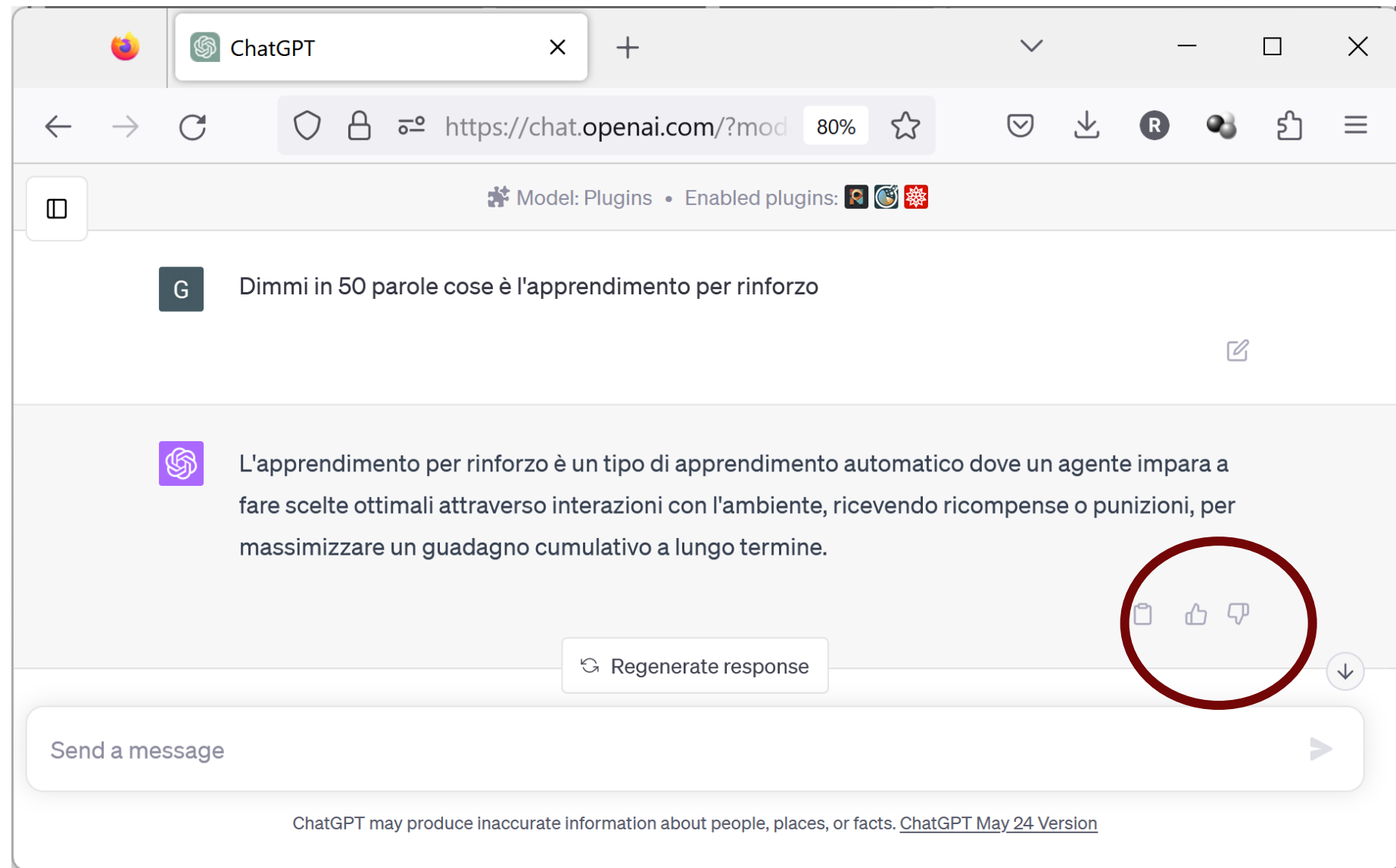
una volta istruita
una DNN, non è
possibile capirne i
meccanismi interni
di funzionamento



Apprendimento profondo (DL – deep learning)

Reti neurali profonde (DNN – deep neural network)

Un tipo di ML con il quale l'AI impara a fare scelte ottimali attraverso interazioni con l'ambiente, ricevendo ricompense o punizioni



Apprendimento per rinforzo (RL – reinforcement learning)

Enormi quantità di dati sia personali che non personali che possono essere usati per l'addestramento di AI tramite ML



Big Data (BD)

AI in grado di soddisfare richieste formulate con il linguaggio naturale per generare testo (LLM), immagini, disegni, musica.

Foto di un girasole con gli occhiali da sole nel mezzo di un campo di girasoli in una giornata di sole



G

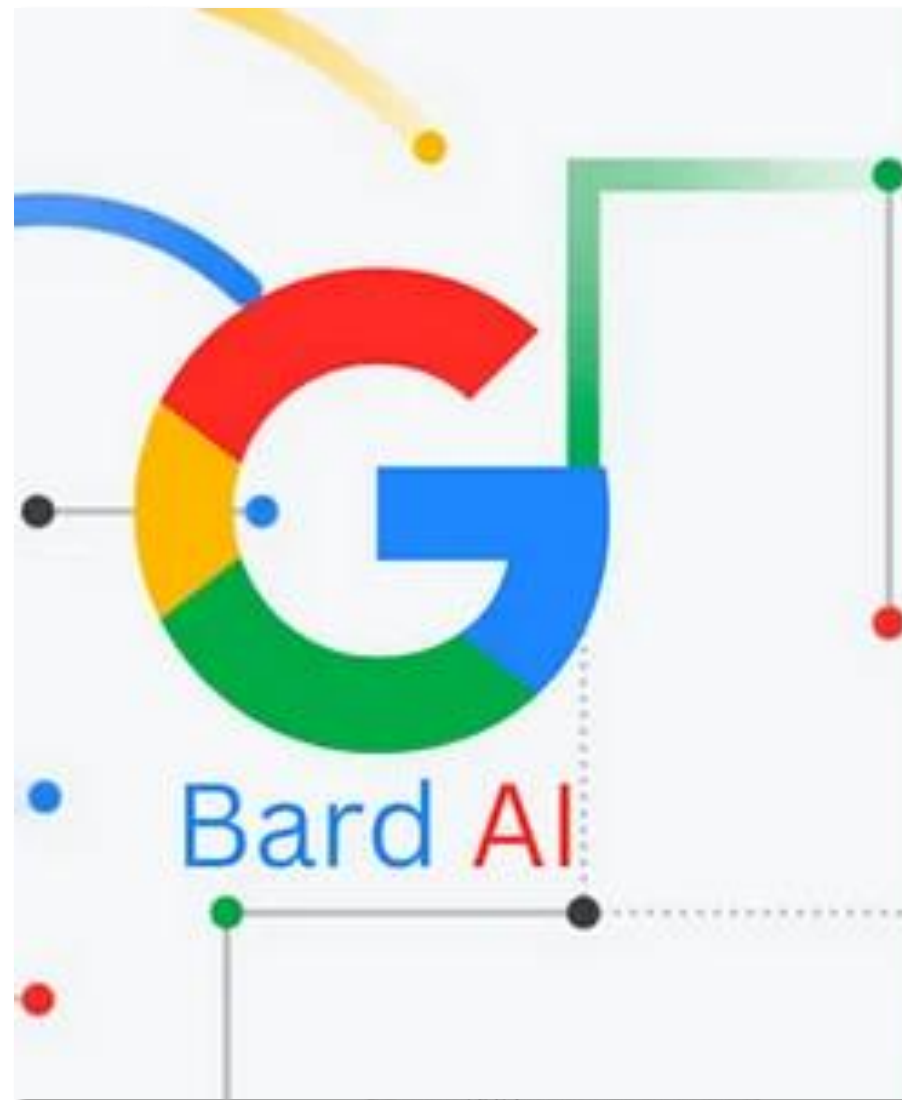
Scrivi un haiku sull'estate



Sole alto splende,
onde danzano al caldo,
l'estate sorride.

AI generativa

Il NLP è una tecnica per realizzare LLM che sono AI in grado di accettare in ingresso, elaborare e produrre in uscita delle informazioni espresse in linguaggio umano senza necessità dei linguaggi di programmazione specializzati per parlare con le macchine



Elaborazione di linguaggio naturale (NLP - Natural Language Processing)
Grandi modelli linguistici (LLM - Large Language Models)

L'arte di comunicare al meglio in linguaggio naturale con le AI di tipo generativo in modo che i contenuti generati siano i migliori possibili



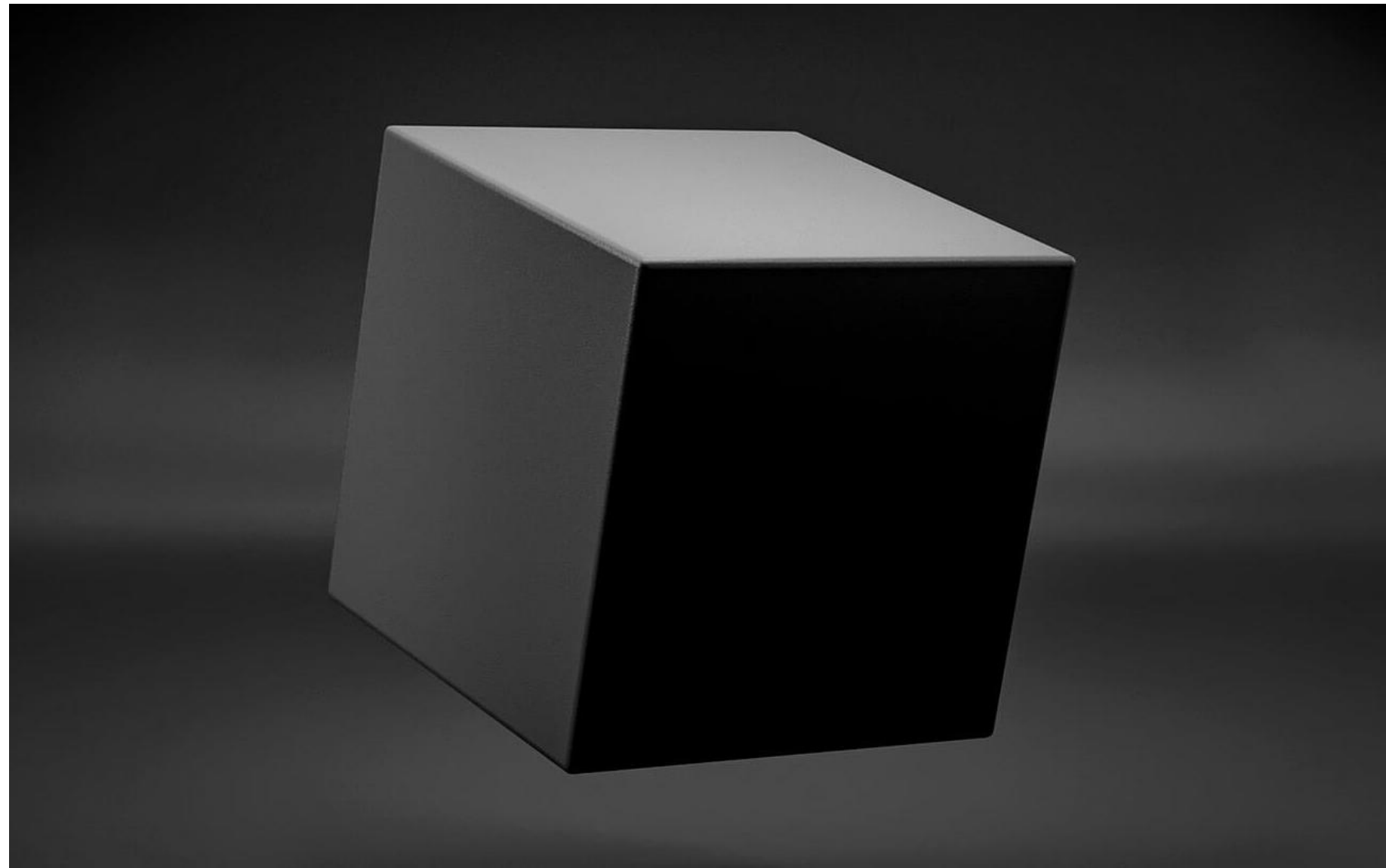
Prompt Engineering

- Termini ibridi che si sono originati al contatto tra nuove tecnologie e ambiti applicativi



Termini «nuovi» nelle nostre discussioni

L'effetto di totale opacità delle AI basate su Deep Neural Networks: se ne sperimenta il funzionamento corretto ma non si riesce a ricostruire la sequenza di passi che lo producono



Effetto «Black Box»

Assenza/presenza di pregiudizi ingiusti nei dati di addestramento che un algoritmo di IA usa per apprendere. Se i dati di addestramento sono distorti o pregiudicati in qualche modo, l'algoritmo di IA può "imparare" e perpetuare questi pregiudizi.



Fairness/Bias

- Termini giuridici che servono per identificare i problemi posti dalle nuove tecnologie



Termini «nuovi» nelle nostre discussioni

Per essere “affidabile”
un’AI deve

- rispettare leggi e regolamenti
- rispettare principi e valori etici
- essere robusta dal punto di vista tecnico (no “allucinazioni”)



AI affidabile (Trustworthy AI)

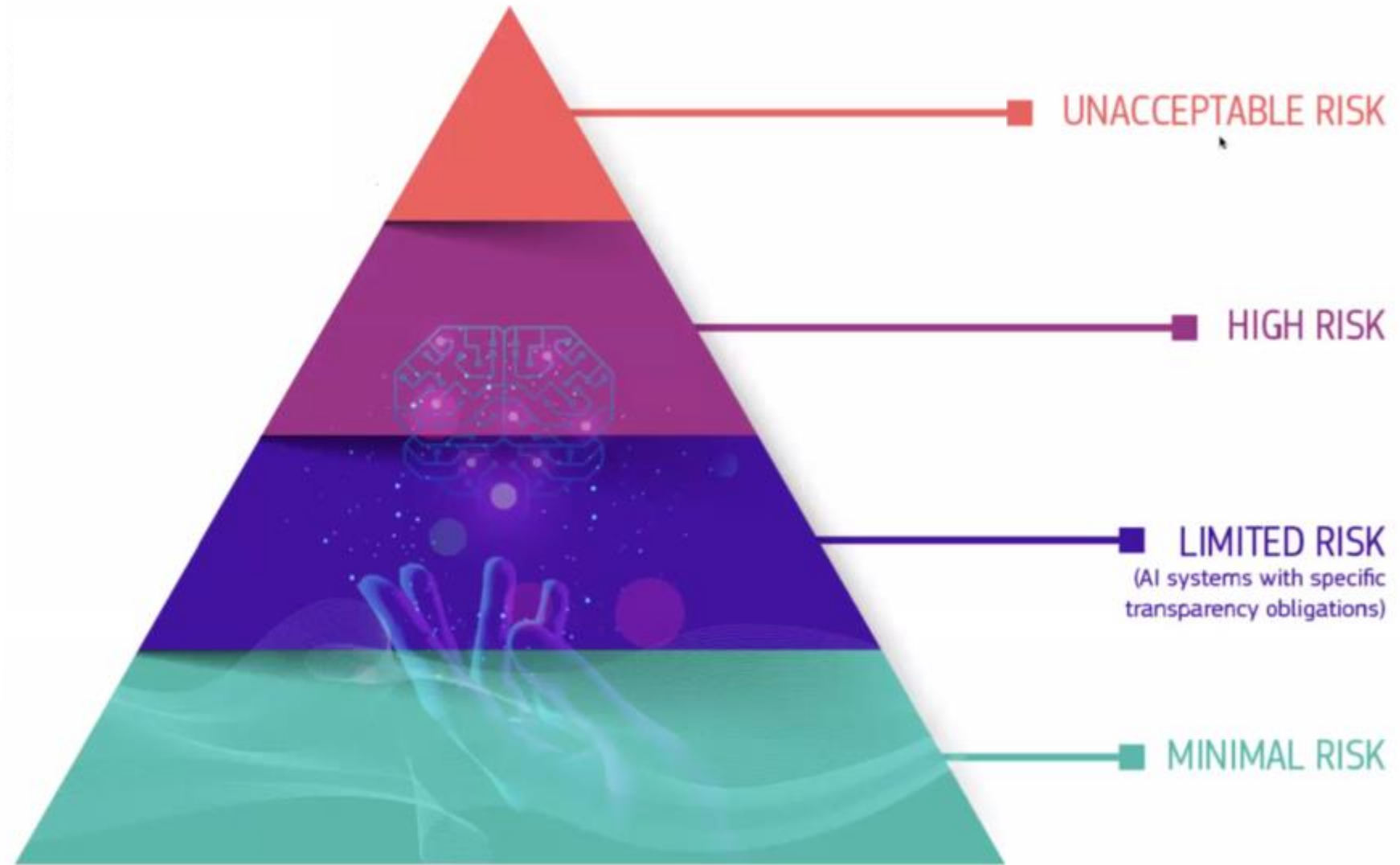
REGOLAMENTO DEL PARLAMENTO EUROPEO E DEL CONSIGLIO

CHE STABILISCE REGOLE ARMONIZZATE SULL'INTELLIGENZA ARTIFICIALE (LEGGE SULL'INTELLIGENZA ARTIFICIALE)

21/04/2021 prima bozza
Commissione Europea

14/06/2023 approvazione
Parlamento Europeo

Possibile entrata in vigore
2024



AI Act



Germana Bertoli

Ringraziamenti