



Verso un nuovo umanesimo digitale

Carlo Ghezzi
Professore Emerito
DEIB—Politecnico di Milano

Contesto

- Oggi assistiamo a quella che viene considerata la *rivoluzione digitale*
- Una rivoluzione che tocca tutti gli esseri umani, la società intera, il pianeta in cui viviamo
- Una rivoluzione che può sia portarci verso un mondo migliore, ma anche compromettere il nostro futuro
- **Umanesimo digitale:** mettere al centro dello sviluppo della tecnologia digitale gli esseri umani e i loro valori, favorire la democrazia e l'inclusione

Contesto

- Oggi assistiamo a quella che viene considerata la *rivoluzione digitale*
- Una rivoluzione che tocca tutti gli esseri umani, la società intera, il pianeta in cui viviamo
- Una rivoluzione che può sia portarci verso un mondo migliore, ma anche compromettere il nostro futuro
- **Umanesimo digitale**: mettere al centro dello sviluppo della tecnologia digitale gli esseri umani e i loro valori, favorire la democrazia e l'inclusione

si tecnologia per gli umani **no** esseri umani per la tecnologia

Obiettivo della conferenza

- La rivoluzione digitale come tappa della storia umana
- Perché la rivoluzione digitale è particolarmente radicale?
- Perché è necessario fondare un nuovo umanesimo digitale?
- Quali riflessi sull'educazione e la formazione tecnologica
- Quali riflessi sulla ricerca

Una visione storica

Perché partire dalla storia

- «Historia vero testis temporum, lux veritatis, vita memoriae, magistra vitae, nuntia vetustatis»
- *La storia in verità è testimone dei tempi, luce della verità, vita della memoria, maestra di vita, messaggera dell'antichità*

Cicerone, De Oratore, II, 9, 36

La madre di tutte le rivoluzioni dell'era moderna: la rivoluzione scientifica

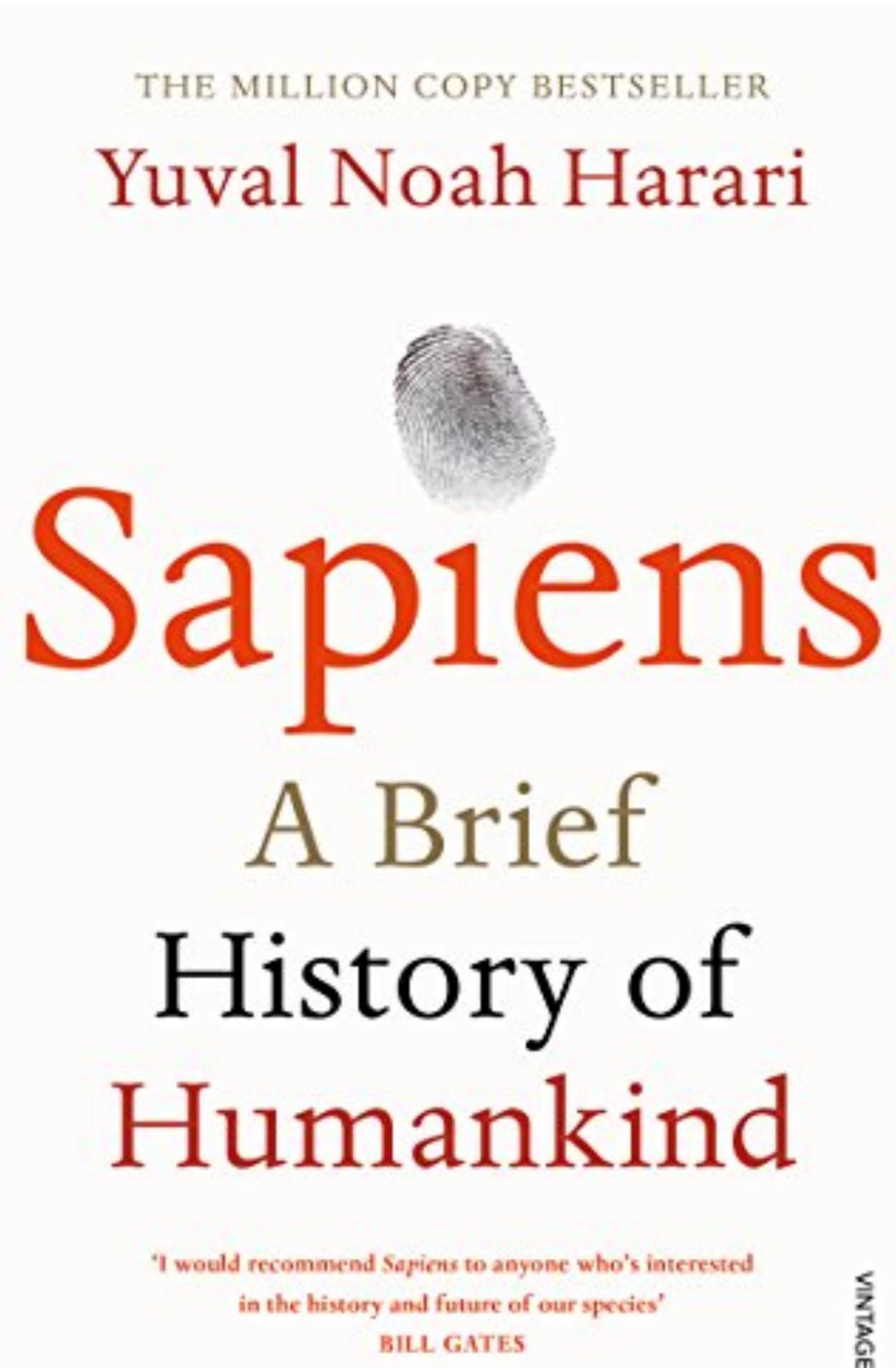
- Fine XVI/inizio XVII secolo
 - Nasce il metodo scientifico moderno con Bacone, Copernico, Galileo, Cartesio, Newton, ...
 - Osservazioni empiriche, formulazione di teoria, convalida della teoria
- Gli ultimi 500 anni hanno assistito a una crescita fenomenale a mai prima sperimentata del potere esercitato dall'umanità, generata dalla rivoluzione scientifica

Sapiens. Da animali a dèi: Breve storia dell'umanità, Yuval Noah Harari, Bompiani 2017
(parte 4, La Rivoluzione Scientifica)

La madre di
la rivoluzione

- Fine XVI/ini
- Nasce il r
Copernic
- Osserv
convali
- Gli ultimi 50
fenomenale
dall'umanità

Sapiens. Da
(parte 4, La F



moderna:

Bacone,

di teoria,

scita
otere esercitato
ntifica

npiani 2017

La rivoluzione scientifica

- Nei millenni precedenti ciò che gli esseri umani chiamavano scienza erano credenze senza giustificazioni, che derivavano dalla religione, dalla superstizione, o da affermazioni che non avevano un supporto empirico
- Pilastri del metodo scientifico
 - Ammissione di ignoranza (sappiamo di non sapere tutto e accettiamo che affermazioni che ci sembrano vere possano dimostrarsi false)
 - Centralità del ricorso a rigorose osservazioni empiriche e della formulazione di teorie matematiche

Rivoluzione scientifica —> rivoluzione industriale

- La rivoluzione scientifica ha generato invenzioni tecnologiche e innovazioni che dalla metà del XVIII secolo hanno generato diverse tappe della rivoluzione industriale
- Da un mondo per millenni basato sull'agricoltura e l'artigianato nasce il nuovo mondo delle macchine e delle fabbriche —> nuovo modo di lavorare e di vivere —> trasformazione della società
- Dalla società feudale nascono la borghesia e il proletariato
- Grandi opportunità e nuove minacce per l'umanità

Caratteristiche salienti

- Uso di (nuovi) materiali per costruire artefatti (ferro e acciaio)
- Uso di nuove sorgenti energetiche e di produzione di energia (carbone, motore a vapore, elettricità, petrolio, motore a scoppio)
- Invenzione di nuove macchine per la produzione di beni, per esempio tessili, con aumento di produttività e riduzione di dispendio di fatica umana
- Nuove forme di organizzazione del lavoro: la fabbrica e la parcellizzazione delle funzioni
- Sviluppo di infrastrutture e servizi: trasporti (locomotiva a vapore, navi a vapore, automobili, aerei), comunicazioni (telegrafo, radio)
 - > utilizzo crescente di risorse naturali e produzione su larga scala di prodotti manifatturieri —> sviluppo *contro* ambiente naturale

La rivoluzione digitale

- Inizia nella seconda metà del XX secolo
- Si realizza in pochi decenni in 3 dimensioni
 - **Ampiezza**
 - tocca ogni ambito della vita delle persone e della società — pervasività
 - **Profondità**
 - incide nella profondità dell'individuo e della società
 - **Velocità**
 - i cambiamenti avvengono a una velocità inedita nella storia

Quali sono le radici tecnologiche della rivoluzione digitale?

Il puzzle

Il puzzle

Le macchine digitali

Il puzzle

Le macchine digitali

Le reti di interconnessione tra macchine

Il puzzle

Le macchine digitali

Le reti di interconnessione tra macchine

L'interazione digitale con esseri umani e mondo fisico

Il puzzle

Le macchine digitali

Le reti di interconnessione tra macchine

L'interazione digitale con esseri umani e mondo fisico

La capacità di agire: il software

Il puzzle

Le macchine digitali

Le reti di interconnessione tra macchine

L'interazione digitale con esseri umani e mondo fisico

La capacità di agire: il software

Dall'*automazione*
All'*autonomia*

Il puzzle

Le macchine digitali

Le reti di interconnessione tra macchine

L'interazione digitale con esseri umani e mondo fisico

La capacità di agire: il software

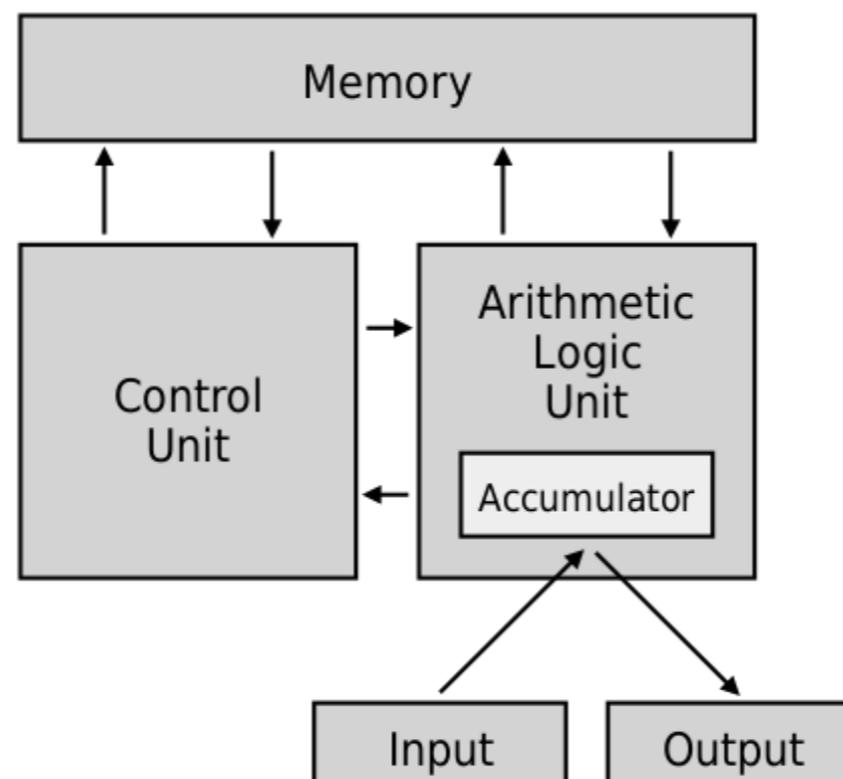
Dall'*automazione*
All'*autonomia*

AI

La rivoluzione (digitale) richiede *macchine (digitali)*

Inizio della storia: il concetto di macchina digitale programmabile

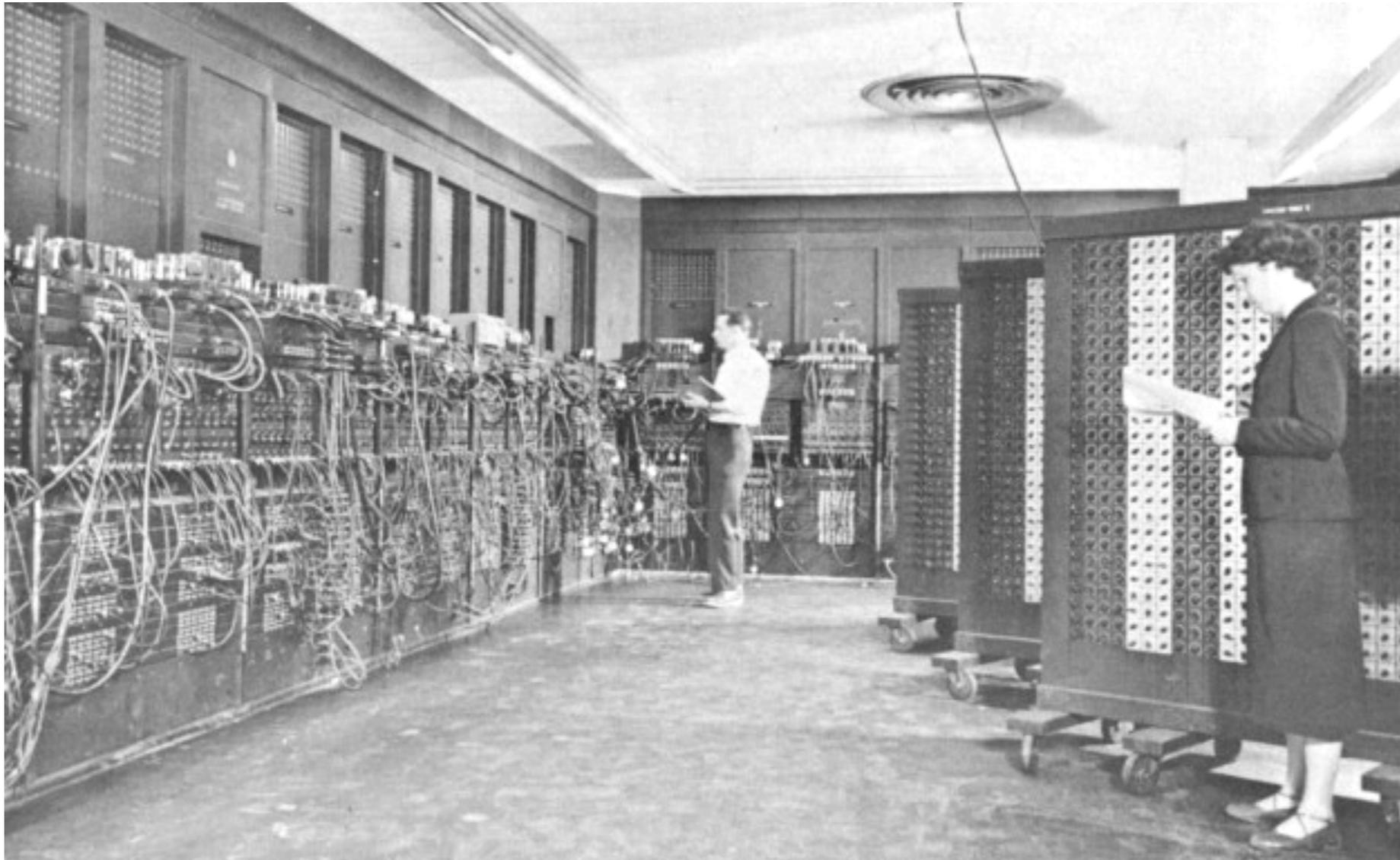
- Rappresentazione binaria delle informazioni e esecuzione di operazioni binarie mediante circuiti (C. Shannon, 1930)
- La "macchina di Turing" come modello di computazione e la "macchina universale di Turing" (A. Turing, 1936, 1945)
- L'architettura di vonNeumann (J. vonNeumann, 1945)



Potere della macchina digitale programmabile

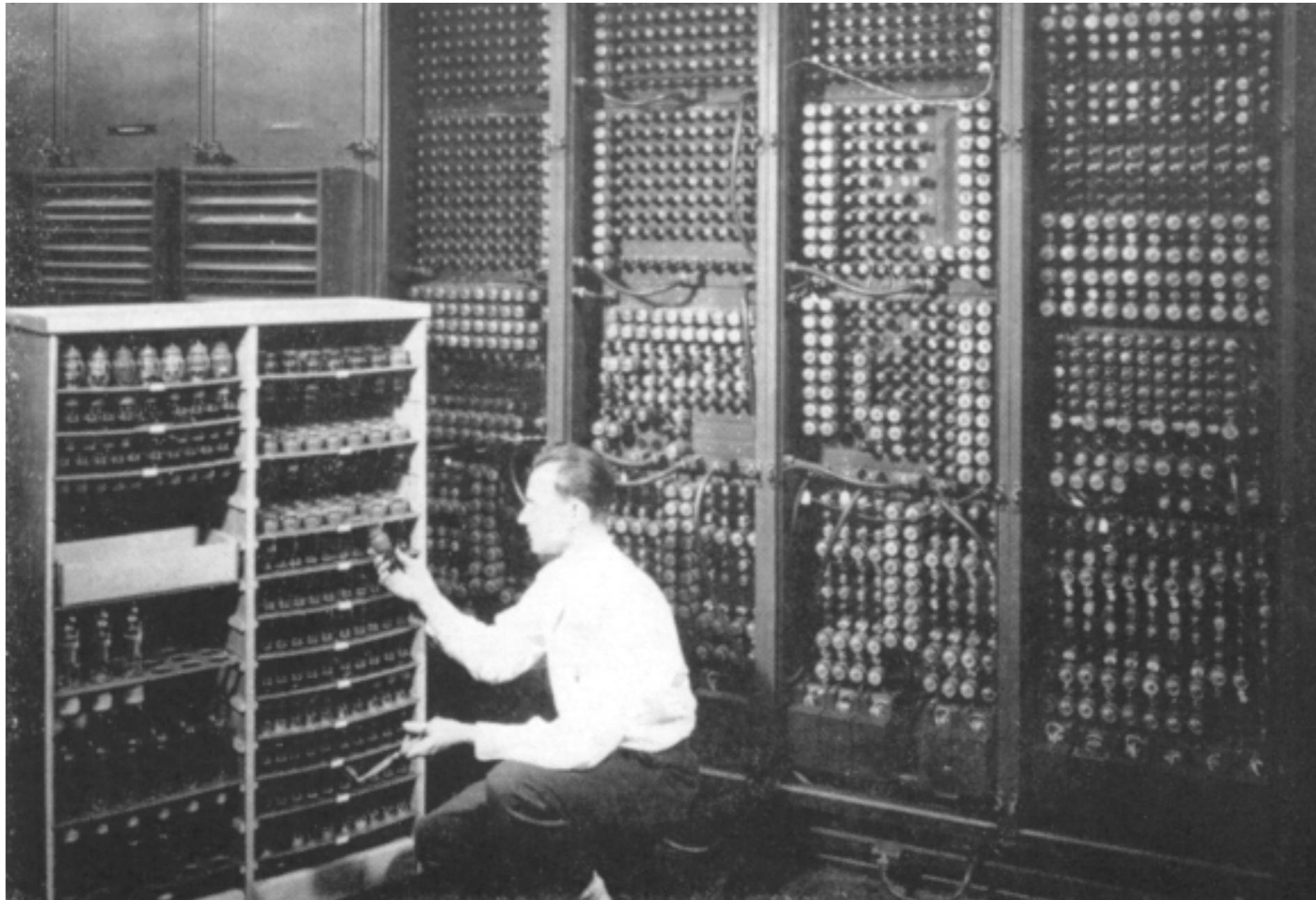
- Ogni informazione può essere rappresentata — codificata — mediante 0 e 1 e trasformata mediante algoritmi da una "macchina digitale" (computer)
- Il mondo reale può essere introiettato in una macchina
- La macchina può modificare il mondo, interagendo con le persone e direttamente attraverso "sensori" e "attuatori" ad essa connessi

Inizi della tecnologia digitale: ENIAC

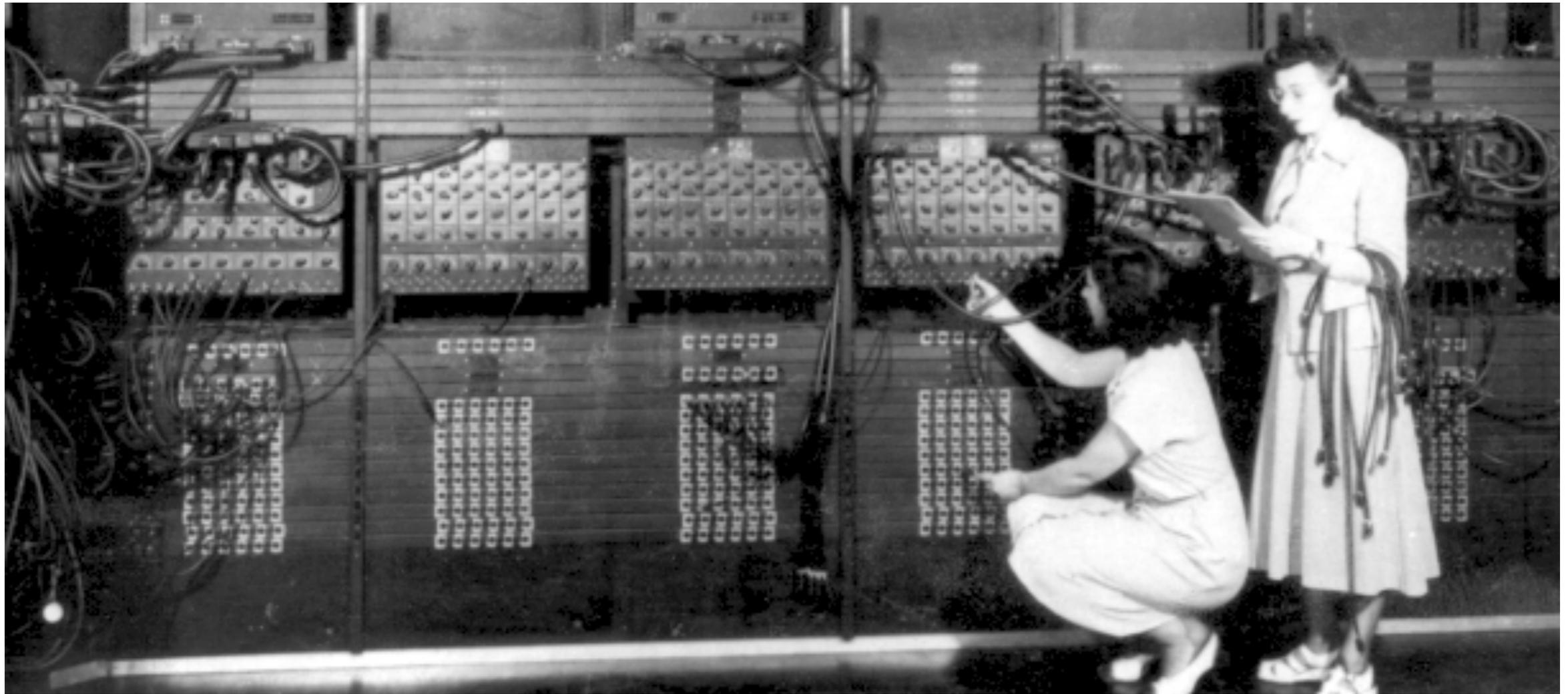


1945 - University of Pennsylvania
spazio: 170 m², peso: 30 t

Attività quotidiane: riparazione



Attività quotidiane: programmazione



I circuiti integrati e la legge di Moore

I circuiti integrati e la legge di Moore

- In un anno raddoppia il numero di componenti di un circuito integrato, aumentando la potenza, diminuendo i costi e l'energia necessaria

I circuiti integrati e la legge di Moore

- In un anno raddoppia il numero di componenti di un circuito integrato, aumentando la potenza, diminuendo i costi e l'energia necessaria



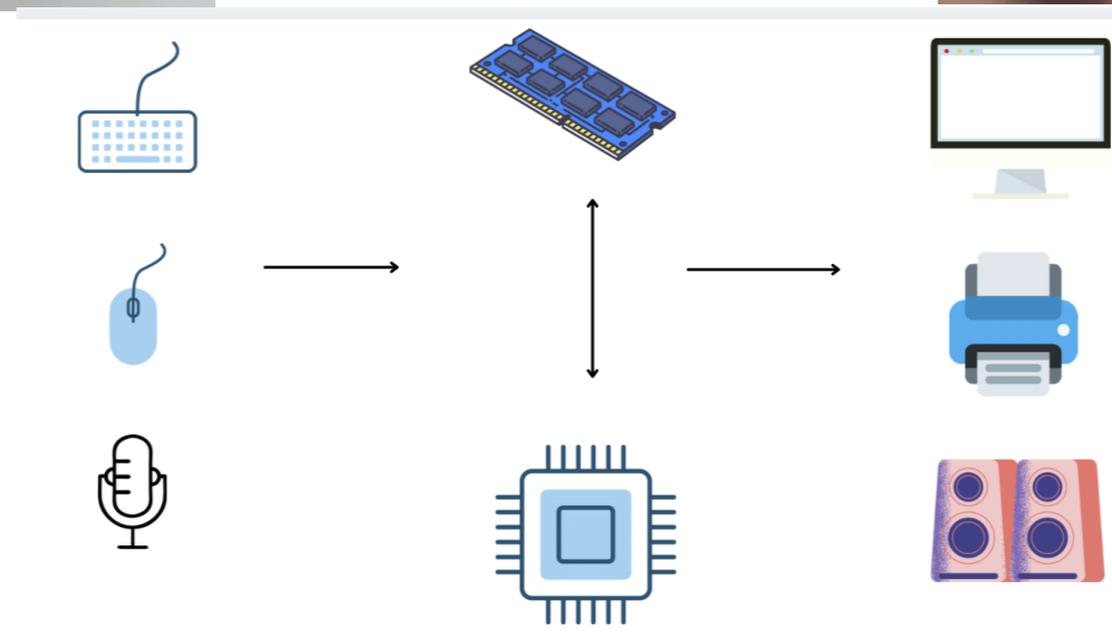
I circuiti integrati e la legge di Moore

- In un anno raddoppia il numero di componenti di un circuito integrato, aumentando la potenza, diminuendo i costi e l'energia necessaria



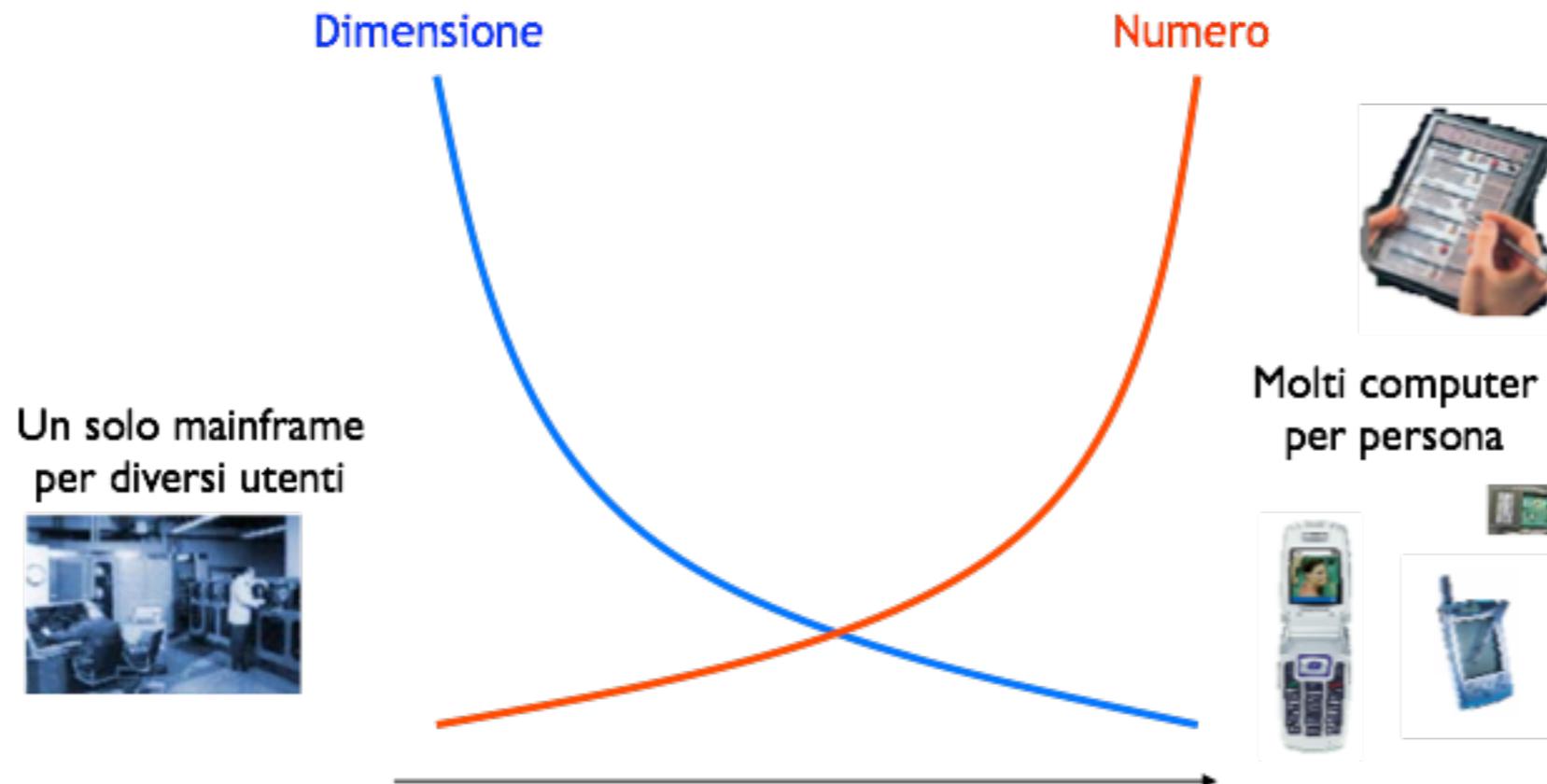
I circuiti integrati e la legge di Moore

- In un anno raddoppia il numero di componenti di un circuito integrato, aumentando la potenza, diminuendo i costi e l'energia necessaria



Computer per le masse

- Dal primo Personal Computer (IBM, 1981)
- ... ai dispositivi



Qualche dato

- microprocessore A15 (cellulari Apple, 2021)
 - 15 Mrd transistori
 - dimensione fisica 107 mm²
 - ~140 Mio transistors/mm²
 - microprocessore M1 Pro (computer Apple, 2021)
 - 34 Mrd transistori
 - dimensione fisica 245 mm²
 - ~140 Mio transistors/mm²
- eseguono Mrd di operazioni al secondo (Eniac ne eseguiva 100k)

La rivoluzione digitale richiede *macchine interconnesse*

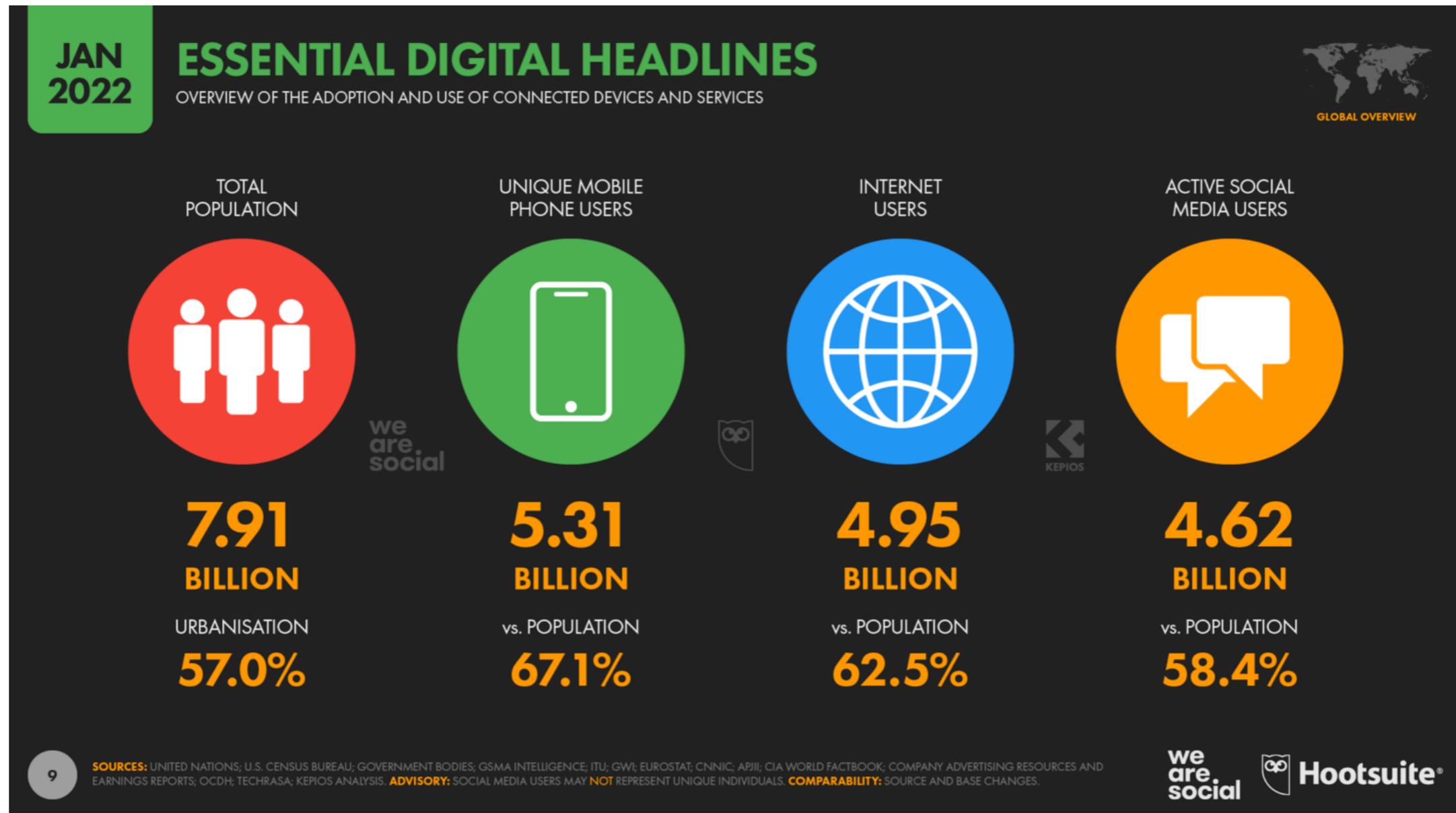
Interconnessione tra macchine digitali

- ARPAnet, l'origine di Internet: nel 1969 collega UCLA, SRI, UCSB, U-Utah
- Internet nasce nel 1983 con la transizione al protocollo TCP/IP; uso prevalentemente accademico
- 12/1995 16 milioni di utenti
- 06/2022 5.473 milioni di utenti, 69% popolazione mondiale

Connessioni di rete per le masse

- 12/1991 GSM
- 1997 wifi rilasciato per uso consumer
- l'utenza di massa accede a Internet con smartphone

Penetrazione della tecnologia



Sorgente: <https://datareportal.com/reports/digital-2022-global-overview-report>

La rivoluzione digitale richiede l'*interazione con il mondo*

Interazione con il mondo fisico

- Sensori e attuatori
- Smartdust
 - nodi autonomi di dimensione di mm che contengono minuscoli sistemi elettromeccanici (MEMS) in grado di rilevare luce, movimento, temperatura, campo magnetico, sostanze chimiche, ... connessi wireless con computers
- Smart tags

Internet of things

- Ogni "cosa" è potenzialmente collegabile a Internet e accessibile ovunque
- Si prevede per il 2025 che ci saranno 125 miliardi di "cose" connesse a Internet, 15 per ciascun essere vivente

15
DEVICES
PER PERSON



Interazione con gli esseri umani

- Inizio della storia: interfacce testuali
- Anni 70: interfacce grafiche — "a picture is worth a thousand words" — Xerox PARC, poi Microsoft Windows e Apple
 - icone, mouse, touchscreen, penne, ...
- Anni 90: Interfacce sonore e riconoscimento vocale
- Dalla metafora interattiva dell'*assistente* virtuale
- ... verso la metafora dell'*agente* virtuale
- Contributi interdisciplinari: psicologia, sociologia, antropologia, design, ...

Interazione con gli esseri umani: dati fisiologici

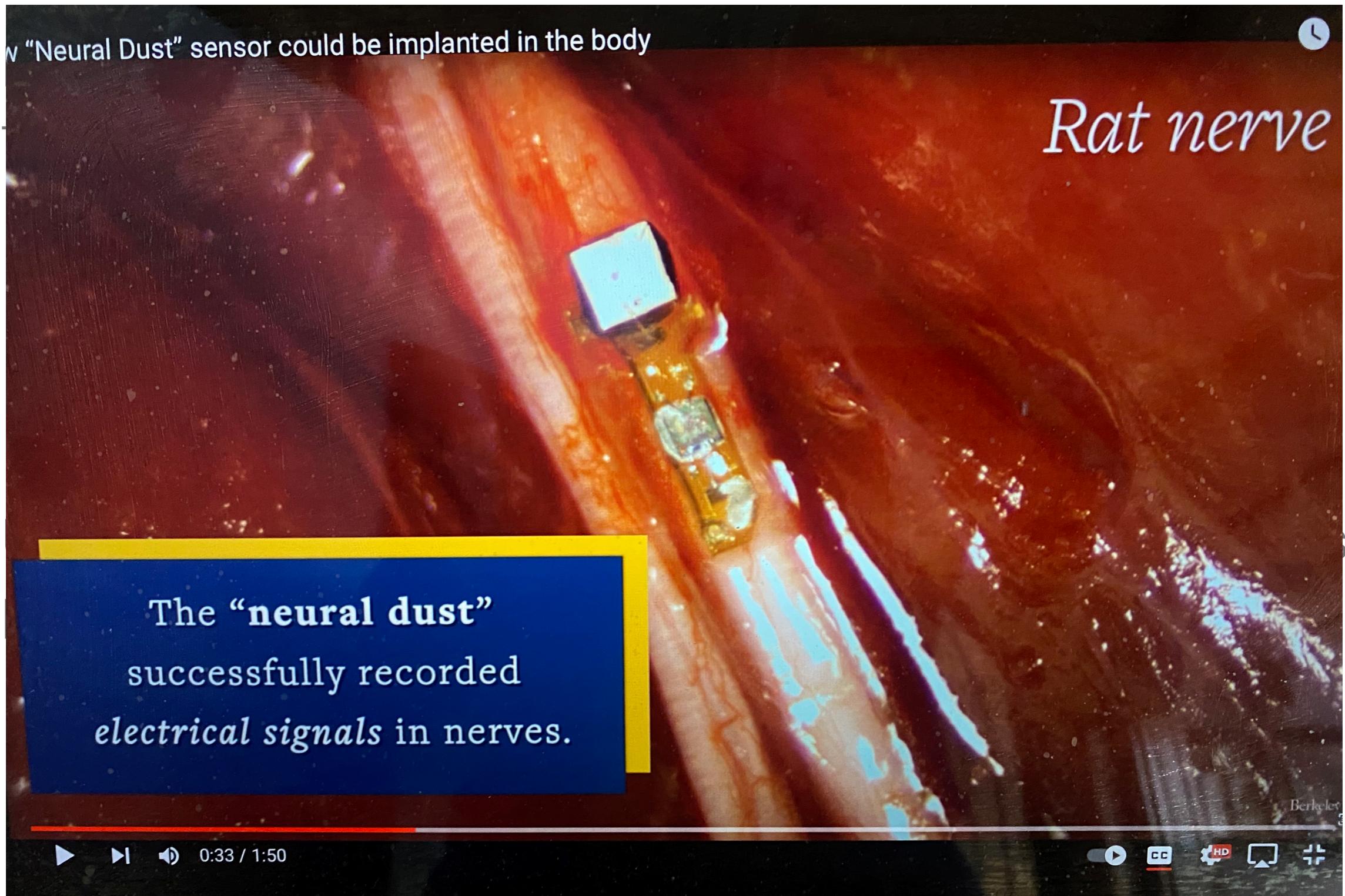
- Magliette "intelligenti" per il tracciamento di segnali fisiologici



- Ultrasonic Neural Dust: raccolta dati dal sistema nervoso periferico (dispositivi millimetrici wireless impiantabili)

Interazione con gli esseri umani: dati fisiologici

- w "Neural Dust" sensor could be implanted in the body



Verso l'interazione "brain machine"

- Comunicazione diretta tra le attività elettriche nel cervello e i dispositivi esterni
- Primi studi negli anni 70
- Dal riconoscimento di semplici "comandi" a una interazione basata sul riconoscimento dello stato mentale ed emozionale
- Contributi interdisciplinari: neuroscienze, informatica (machine learning), elettronica, ...

La rivoluzione digitale richiede *macchine versatili
capaci di operare*

il software

Il ruolo chiave del software

- Fornisce vita e intelligenza all'hardware

È vero che il software non potrebbe esercitare i poteri della sua leggerezza se non mediante la pesantezza dello hardware; ma è il software che comanda, che agisce sul mondo esterno e sulle macchine, le quali esistono solo in funzione del software, si evolvono in modo d'elaborare programmi sempre più complessi.

La seconda rivoluzione industriale non si presenta come la prima con immagini schiaccianti quali presse di laminatoi o colate d'acciaio, ma come i bits d'un flusso d'informazione che corre sui circuiti sotto forma d'impulsi elettronici. Le macchine di ferro ci sono sempre, ma obbediscono ai bits senza peso.

Italo Calvino, "Lezioni Americane", 1985
I lezione, dualismo "leggerezza peso"

Evoluzione: dal software per calcolare

- Umani non competitivi per calcoli complessi
- Problemi non affrontabili in pratica da umani
 - es. calcolo durata accensione dei razzi di frenata per sbarco sulla Luna (1969)
- Oggi un processore di PC esegue miliardi di operazioni/sec su numeri

Evoluzione: ... al software per automatizzare

- Automazione introdotta in ogni ambito, non solo industriale, ma anche personale e sociale
- Da attività ripetitive e basate su procedure standard a attività a maggior valore aggiunto
- Sostituzione/supporto del lavoro umano
- Software per decidere
- Software per progettare
- Software per educare, divertire, ...

Evoluzione:

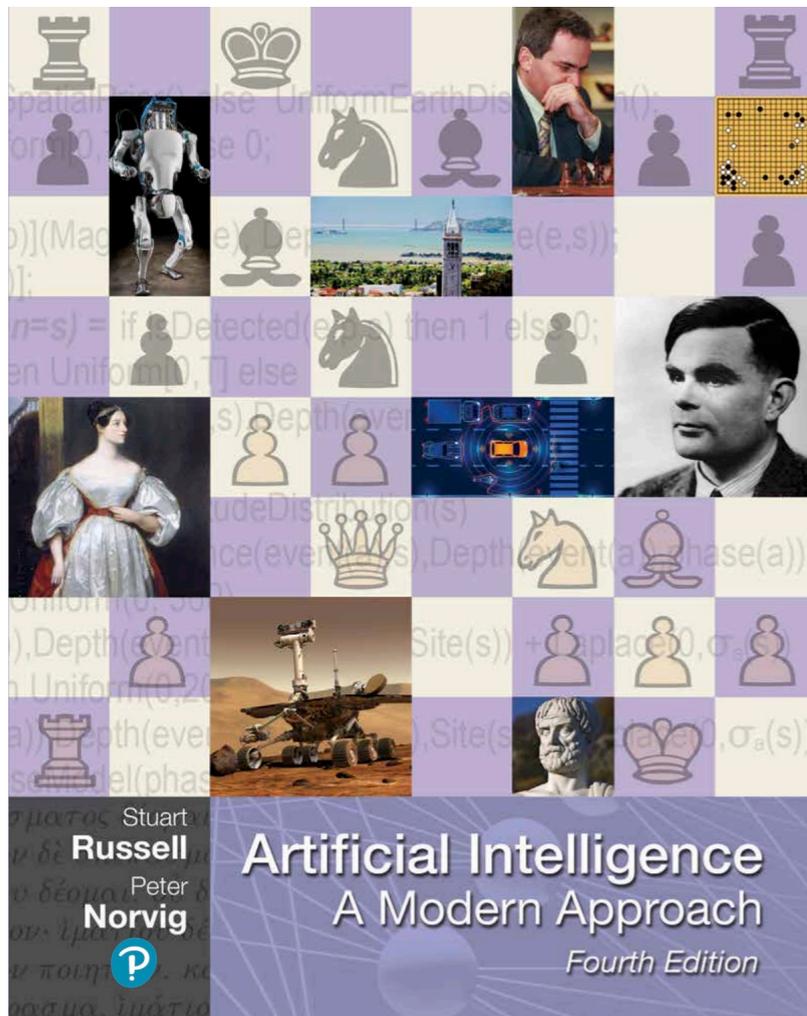
il software rende le macchine **autonome**
mediante l'AI

Evoluzione:

il software rende le macchine **autonome**
mediante l'AI

qualche riflessione sull'AI

S. Russell, P. Norvig – il classico testo (4th ed.)



- Pensare/agire razionalmente (la tradizione AI che deriva dalla logica)
 - **successo** la macchina è razionale; "fa la cosa giusta in base a ciò che sa"
- Pensare/agire come un essere umano
 - decidere, risolvere problemi, imparare....
muovere oggetti, ...saldare, ... operare bisturi
 - rendere capace una macchina di fare ciò che oggi gli umani sanno fare meglio
 - **successo** indistinguibilità dall'essere umano

Il test di Turing

- Una definizione "pragmatica" di intelligenza
- Così formulato per una macchina interattiva:

Dopo aver posto alla macchina una serie di domande scritte, in base alle risposte ottenute un essere umano non sa dire se ha interagito con un essere umano o con una macchina

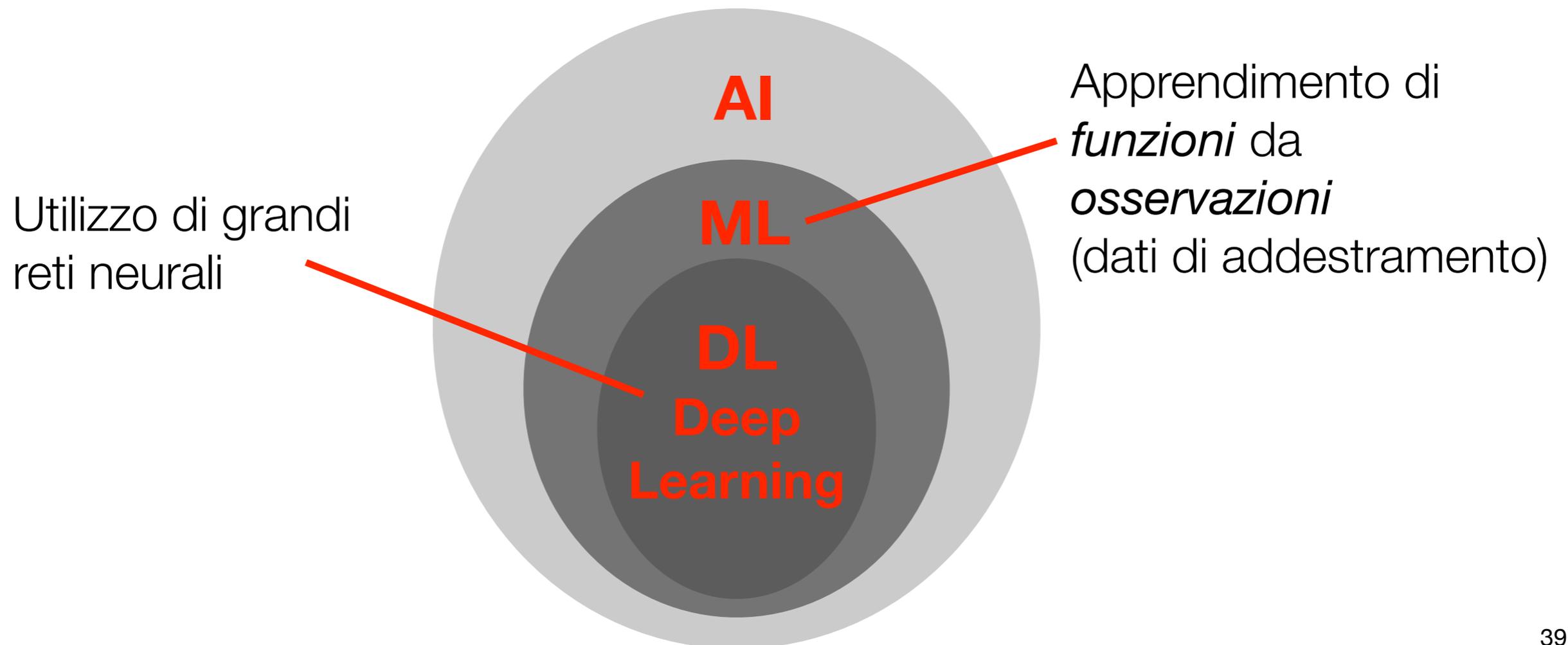
Oggi quando si parla di AI ci si riferisce a ML

Oggi quando si parla di AI ci si riferisce a ML

l'apprendimento è la base per l'autonomia

Machine learning

- Da sempre un sogno
- Diventato realtà con l'esplosione della disponibilità di dati in rete da cui "apprendere"



Come impara una macchina?

- Inferenza statistica: capacità di correlare dati e generalizzare leggi di dipendenza tra fenomeni osservati (p. es., apprendere il profilo degli utenti o classificare i comportamenti osservati all'interno di classi specifiche — riconoscere immagini)
- Per apprendere, l'algoritmo non deve capire un fenomeno osservato, ma è invece in grado di riprodurlo se ha potuto osservare un insieme sufficientemente grande di esempi
- Importanza della quantità e della **qualità** dei dati
- **Apprendimento** —> **autonomia**: apprendere è la base per decidere come agire

Apprendimento basato su reti neurali profonde (deep learning)

- Struttura computazionale ispirata alla struttura del cervello degli animali
- Sono una collezione di nodi (neuroni artificiali) connessi
- Ogni connessione (ispirata dalla sinapsi) può trasmettere un segnale ad altri neuroni
- Un neurone riceve segnali, ne calcola una funzione e trasmette il risultato ai neuroni ad esso connessi
- Neuroni e connessioni hanno un *peso* che si modifica durante l'apprendimento. Il peso aumenta o diminuisce la forza del segnale in una connessione
- In DL le reti sono organizzate a livelli: ogni livello estrae progressivamente dai dati informazioni di livello sempre più alto e astratto

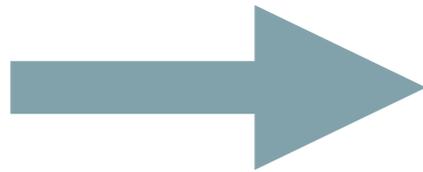
Costruzione di un'applicazione mediante DL

Costruzione di un'applicazione mediante DL



Costruzione di un'applicazione mediante DL

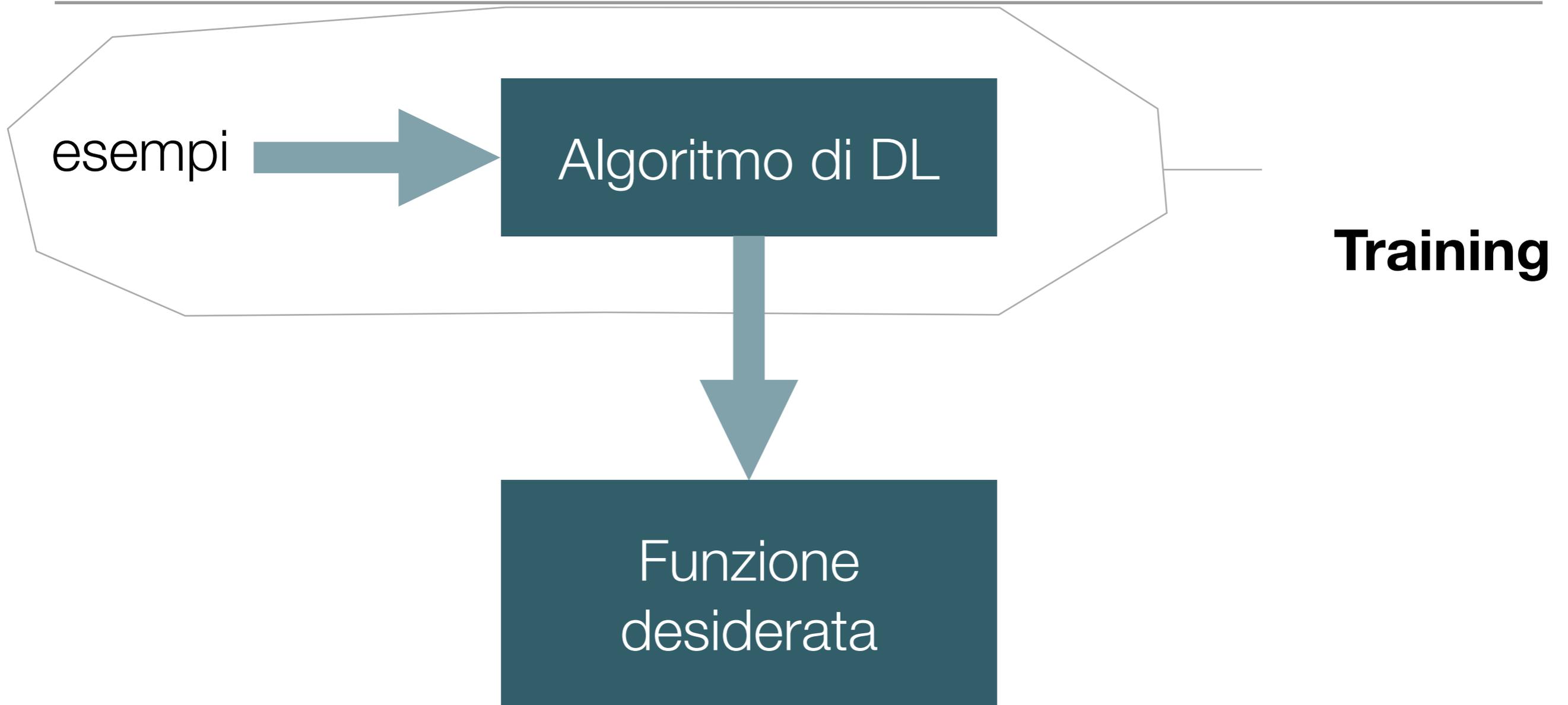
esempi



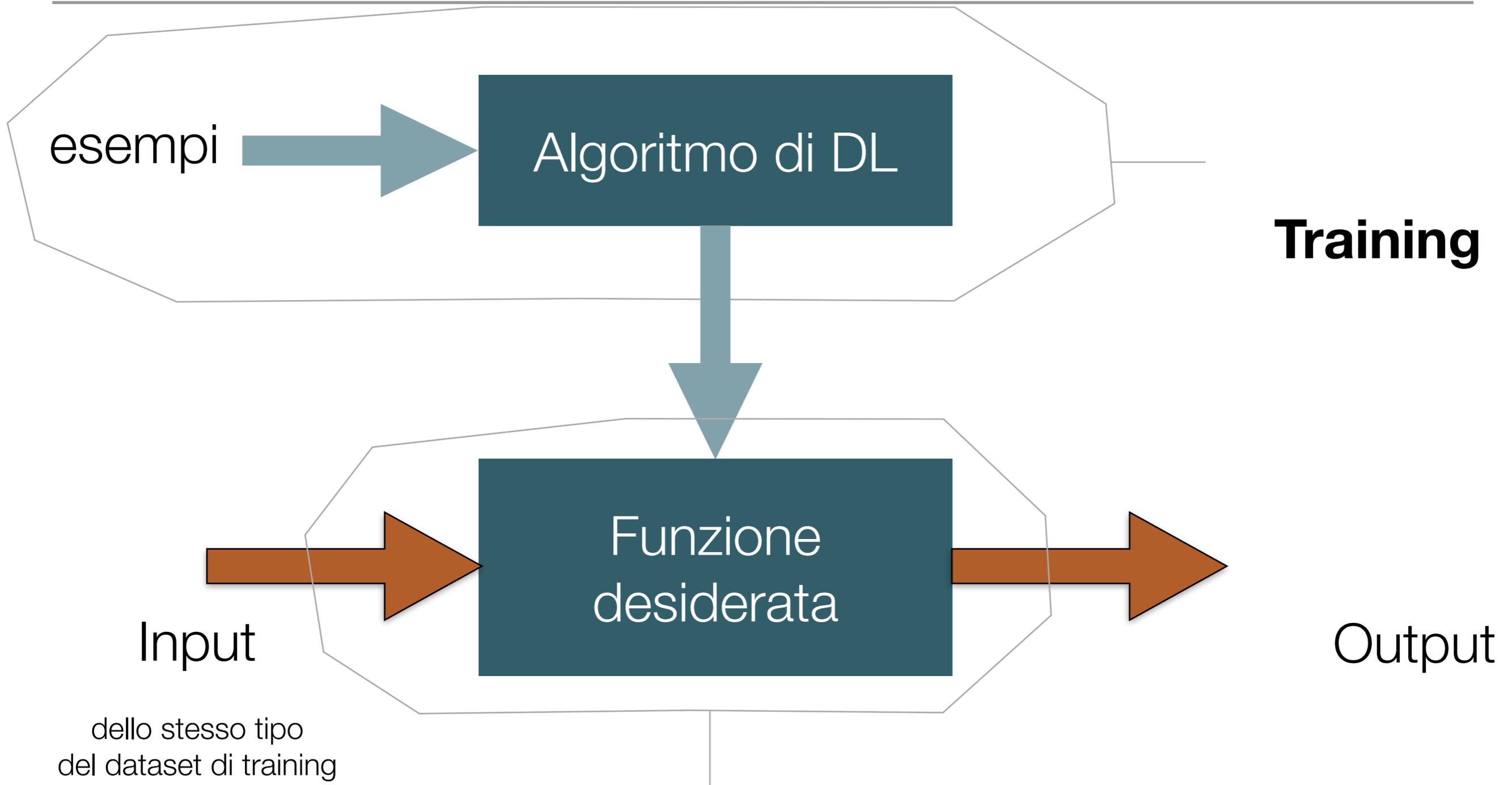
Algoritmo di DL

Training

Costruzione di un'applicazione mediante DL



Costruzione di un'applicazione mediante DL



Uso dell'applicazione

Differenza rispetto al software prodotto tradizionalmente

- Tradizionalmente la funzione desiderata è progettata e programmata da una persona umana
- Con ML, la funzione è il risultato dell'addestramento che si basa su dati esistenti (*Il futuro dipende dal passato*)
 - Ruolo cruciale dei dati usati per il training
- Per un programma tradizionale è possibile argomentare perché produce certi risultati
- Non è di fatto possibile giustificare perché una funzione inferita con DL produce un certo risultato —> *explainable AI*
- Costo del training

Open AI e ChatGPT

ChatGPT launched on wednesday. today it crossed 1 million users!
7:35 AM · Dec 5, 2022

Open AI e ChatGPT

ChatGPT launched on wednesday. today it crossed 1 million users!
7:35 AM · Dec 5, 2022

Come ti definiresti da un punto di vista tecnologico?

Essendo un modello di lingua artificiale sviluppato con la tecnologia GPT-3.5, mi definirei come una piattaforma di intelligenza artificiale in grado di elaborare grandi quantità di testo, generare contenuti e rispondere alle domande degli utenti in modo naturale e automatico. Sono in grado di apprendere dalle interazioni con gli utenti e migliorare costantemente le mie prestazioni grazie all'uso di algoritmi di apprendimento automatico. In sintesi, mi definirei come un assistente virtuale di linguaggio naturale, in grado di fornire supporto e risposte aiutando gli utenti a comunicare e interagire con il mondo digitale in modo più efficace.

Estratto di conversazione con ChatGPT

Estratto di conversazione con ChatGPT

Quale e' stato il costo del training di chatGPT (tempo e consumo di energia)?

Estratto di conversazione con ChatGPT

Quale e' stato il costo del training di chatGPT (tempo e consumo di energia)?

Non ho informazioni precise sul costo del training di ChatGPT in termini di tempo e consumo di energia. Tuttavia, posso darti un'idea generale sulla portata del training di un modello di linguaggio di grande dimensione come ChatGPT.

Il training di un modello di linguaggio basato su transformer come ChatGPT richiede una grande quantità di dati di testo, potenza di calcolo e tempo di training. Il modello GPT-3, che è il predecessore di ChatGPT, ad esempio, è stato addestrato su un dataset di oltre 45 terabyte di testo e ha richiesto circa $3,14 \times 10^{17}$ operazioni di calcolo.

In termini di tempo di training, GPT-3 è stato addestrato per diversi mesi utilizzando decine di migliaia di processori grafici (GPU) in parallelo. Tuttavia, non ho informazioni precise sul tempo e sul numero di risorse utilizzate per addestrare ChatGPT.

Per quanto riguarda il consumo di energia, il training di modelli di linguaggio di grande dimensione richiede un'enorme quantità di energia. Gli studi suggeriscono che il training di modelli di linguaggio come GPT-3 può avere un impatto significativo sull'ambiente. Tuttavia, le informazioni esatte sul consumo di energia per il training di ChatGPT non sono disponibili.

Estratto di conversazione con ChatGPT

Estratto di conversazione con ChatGPT

I dati di training di chatGPT sono tutti pubblici e accessibili per chiunque voglia realizzare un LLM simile a chatGPT?

Estratto di conversazione con ChatGPT

I dati di training di chatGPT sono tutti pubblici e accessibili per chiunque voglia realizzare un LLM simile a chatGPT?

Non tutti i dati di training di ChatGPT sono pubblici e accessibili a chiunque voglia creare un modello di linguaggio simile a ChatGPT.

Il modello di ChatGPT è stato addestrato su un vasto dataset di testo che comprende molti contenuti web e testi in diversi formati. Tuttavia, il dataset utilizzato per addestrare ChatGPT è stato creato da OpenAI e non è completamente pubblico. OpenAI ha reso disponibili alcuni subset dei dati di training, come il dataset "WebText" contenente circa 40 GB di testo in lingua inglese.

Tuttavia, l'intero dataset di training utilizzato per addestrare ChatGPT non è disponibile pubblicamente per la maggior parte degli utenti. Inoltre, il modello di ChatGPT utilizza un'architettura di rete neuronale molto avanzata e sofisticata, che richiede una grande potenza di calcolo per il training e l'elaborazione dei dati.

Sebbene sia possibile addestrare un modello di linguaggio usando dataset pubblicamente disponibili, creare un modello di linguaggio come ChatGPT richiederebbe un'expertise e risorse considerevoli in termini di potenza di calcolo, infrastrutture di cloud computing e personale qualificato.

Estratto di conversazione con ChatGPT

Estratto di conversazione con ChatGPT

Qual'e' la dimensione complessiva del dataset di addestramento di chatGPT?

Estratto di conversazione con ChatGPT

Qual'è la dimensione complessiva del dataset di addestramento di chatGPT?

La dimensione complessiva del dataset di addestramento di ChatGPT non è nota con precisione, in quanto OpenAI non ha rilasciato informazioni dettagliate sul dataset completo utilizzato per addestrare il modello.

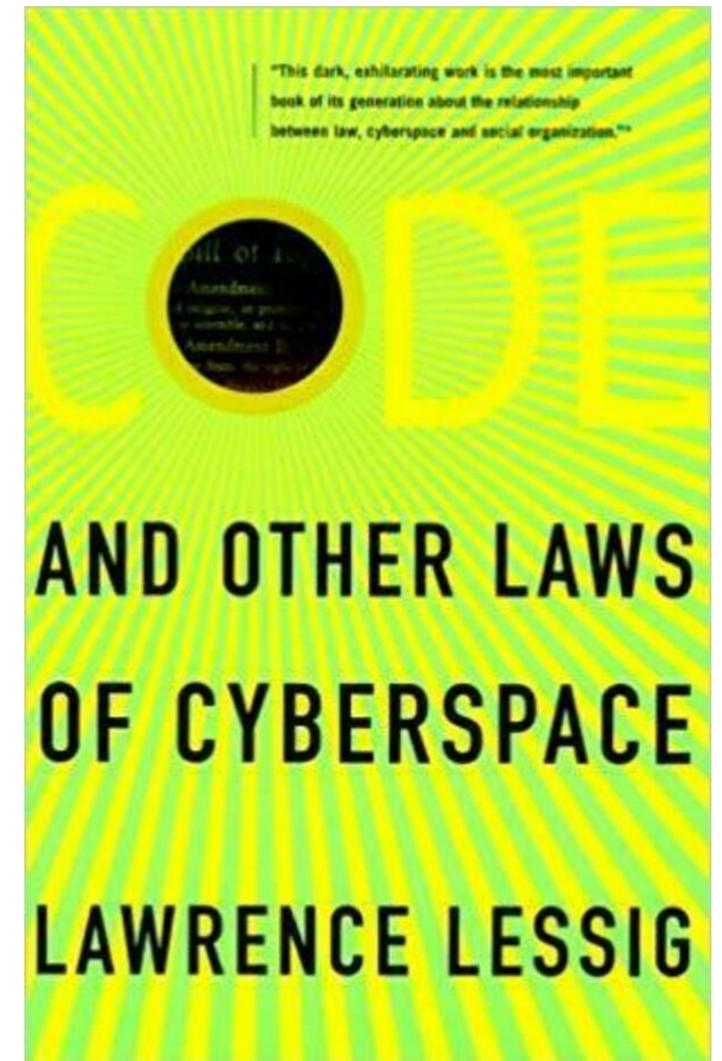
Tuttavia, OpenAI ha reso pubblici alcuni dei dataset utilizzati per addestrare ChatGPT, come il dataset "WebText", che contiene circa 40 GB di testo in lingua inglese. Inoltre, OpenAI ha dichiarato che il dataset completo utilizzato per addestrare ChatGPT include un ampio range di testi in molte lingue diverse, con una vasta gamma di contenuti, come testi scientifici, articoli di notizie, conversazioni informali e molto altro ancora. Sulla base di queste informazioni, è probabile che il dataset completo utilizzato per addestrare ChatGPT sia molto grande, probabilmente nell'ordine dei terabyte. Tuttavia, senza informazioni ufficiali da parte di OpenAI, non è possibile fornire una stima precisa della dimensione del dataset completo di addestramento di ChatGPT.

ChatGPT e la programmazione

- The End of Programming
M.Welsh, Communications of the ACM, Jan. 2023
- The Premature Obituary of Programming
D.M. Yellin, Communications of the ACM, Feb. 2023

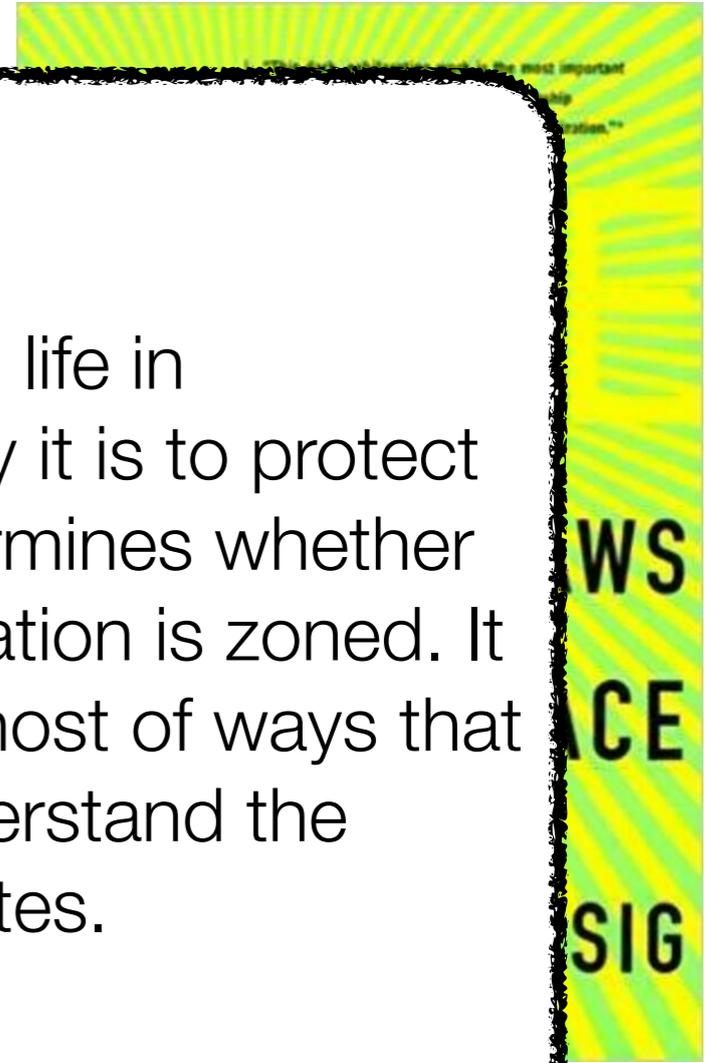
Ruolo del software nel mondo digitale

- L. Lessig, 1999 (Harvard U.) **Code is law** (*code = software*)
- Il mondo e le regole secondo le quali opera l'intera società sono codificati nel software
 - progettare tecnologia <> progettare il mondo
- L'autonomia (e la scarsa trasparenza) raggiunta con l'AI esaspera ulteriormente il livello critico



Ruolo del software nel mondo digitale

- L'architettura del software (code architecture) è il più importante dei fattori che determinano il modo in cui la vita in cyberspace è esperita. Il codice regola l'accesso alla privacy, o quanto è facile censurare la parola. Il codice determina se l'accesso all'informazione è generale o se l'informazione è zonata. Il codice determina chi vede cosa, o cosa è monitorata. In un modo o nell'altro, a meno che non si cominci a comprendere la natura di questo codice, il codice di cyberspace regola.
- L'aggiunta di AI esaspera ulteriormente il livello critico



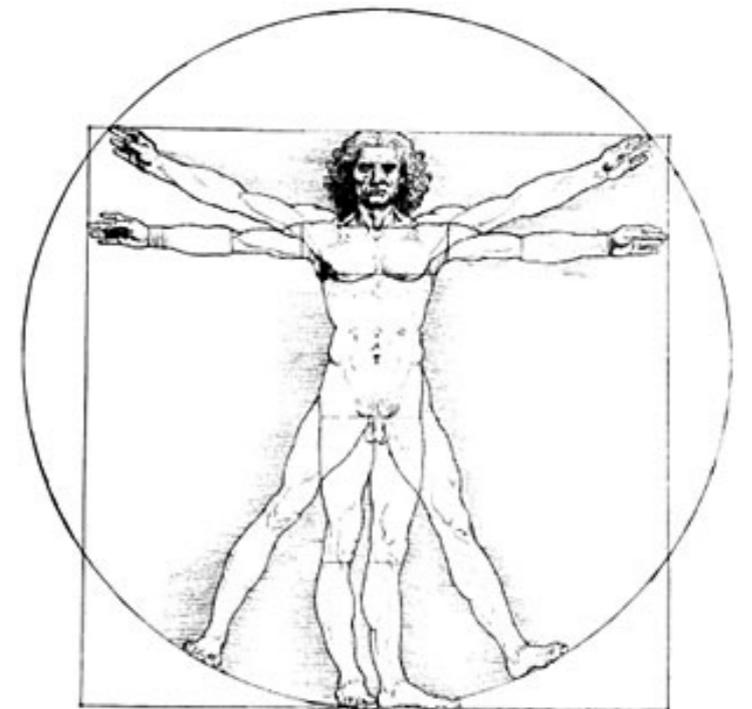
Perché è necessario fondare un nuovo umanesimo digitale?

Perché è necessario fondare un nuovo umanesimo digitale?

Perché il mondo digitale deve essere a misura dell'uomo

Perché la tecnologia deve essere al suo servizio e non viceversa

Perché tutti gli uomini devono poter decidere del proprio futuro e della società in cui vivono e la salute del mondo deve essere salvaguardata





Vienna Manifesto on Digital Humanism

Vienna, May 2019

“The system is failing” – stated by the founder of the Web, Tim Berners-Lee – emphasizes that while digitalization opens unprecedented opportunities, it also raises serious concerns: the monopolization of the Web, the rise of extremist opinions and behavior orchestrated by social media, the formation of filter bubbles and echo chambers as islands of disjoint truths, the loss of privacy, and the spread of digital surveillance. Digital technologies are disrupting societies and questioning our understanding of what it means to be human. The stakes are high and the challenge of building a just and democratic society with humans at the center of technological progress needs to be addressed with determination as well as scientific ingenuity. Technological innovation demands social innovation, and social innovation requires broad societal engagement.

This manifesto is a call to deliberate and to act on current and future technological development. We encourage our academic communities, as well as industrial leaders, politicians, policy makers, and professional societies all around the globe, to actively participate in policy formation. Our demands are the result of an emerging process that unites scientists and practitioners across fields and topics, brought together by concerns and hopes for the future. We are aware of our joint responsibility for the current situation and the future – both as professionals and citizens.

Today, we experience the co-evolution of technology and humankind. The flood of data, algorithms, and computational power is disrupting the very fabric of society by changing human interactions, societal institutions, economies, and political structures. Science and the

Hannes Werthner
Erich Prem
Edward A. Lee
Carlo Ghezzi *Editors*

Perspectives on Digital Humanism

OPEN ACCESS

 Springer

Estratto dal DigHum statement on chatGPT

<https://dighum.ec.tuwien.ac.at/statement-of-the-digital-humanism-initiative-on-chatgpt/>

Released by the Open AI company near the end of 2022 as a conversational version of their Generative AI models, ChatGPT (Generative Pre-Trained Transformer) is a Large Language Model (LLM) based on the innovative combination of both unsupervised training and reinforcement learning from humans. The large data sets come from sources such as books, (news) articles, websites or posts, or comments from social networks to perform its core function as a dialogue system simulating human conversation. It achieves this by ‘estimating’ via probabilities which word(s) is likely to follow the previous word(s). This is done in accordance with specific writing styles or tones which, in turn, creates the illusion of conversing with a human. While apparently good at mastering this aspect of language, these systems are rather limited at the functional level, e.g., lack of reasoning and abstraction capabilities, or very limited situation modelling. We humans speak to communicate with other humans with the intention of achieving some goal. We tend to attribute this intention to all agents that produce language. We are thus easily seduced to project human intelligence as we understand it onto machines capable of some form of language imitation.

It is important to underline that such models can build convincing combinations of words and sentences, but it does not have human understanding of our questions nor its own answers. Neither does it have an understanding of what ‘facts’ are and is prone to produce factual errors and to ‘hallucinate’. Open AI admits the limitations of its models: ‘ChatGPT sometimes writes answers that sound reasonable, but in fact are incorrect or nonsense’. This has been the reason for characterizations such as ‘stochastic parrot’ or, less kindly, ‘confident bullshitter’ and others which, of course, are occasionally applicable to humans as well.

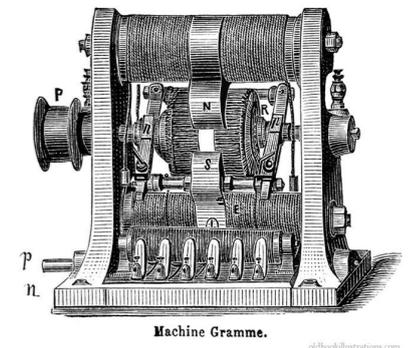
OpenAI and other companies have now entered a fierce competitive race and continue to gather feedback data from a rapidly increasing number of users. This means that all of us are the subjects of a huge ongoing field experiment that these companies are conducting – without our consent.

Umanesimo digitale: **come** ci riguarda?

- Come cittadini — effetti sulla democrazia, la giustizia sociale, la sovranità
- Come "addetti ai lavori"
- *quanto e come l'umanesimo digitale ha a che fare con la ricerca e la formazione in ingegneria del software?*

Rivisitiamo qualche fondamento

- Michael Jackson e Pamela Zave (~1990)
 - The WORLD and the MACHINE
 - Noi costruiamo macchine (astratte) per raggiungere certi obiettivi (GOALS) e soddisfare certi requisiti (REQUIREMENTS)
 - Dobbiamo capire lo specifico dominio, le sue proprietà e fare le necessarie assunzioni
- Responsabilità del progettista: **M, E \models R**



M, E \models R nel contesto socio-tecnico

- Goals e requirements esprimono *anche* "valori" umani e sociali
- Contesto e ambiente riguardano *anche* proprietà/assunzioni che riguardano le persone e la società
- *Goals e requirements sono prevalentemente dettati dai tecnologi (ingegneri del software)*
- *Alcuni goals/requirements hanno a che fare con altri goals/requirements che non sono noti, o sono nascosti o sono impliciti — unknown unknowns*
- *La costruzione della "macchina" è nelle mani degli ingegneri del software*

Opacità

- Le regole espresse dal software sono opache, incomprensibili dalle persone (non tecniche), e sicuramente dalla grande maggioranza dei cittadini e dei decisori (pubblici)
- All'opacità si aggiunge la complessità, che impedisce di capire e dominare gli effetti
- AI (componenti generati da ML) amplifica di ordini di grandezza opacità e complessità

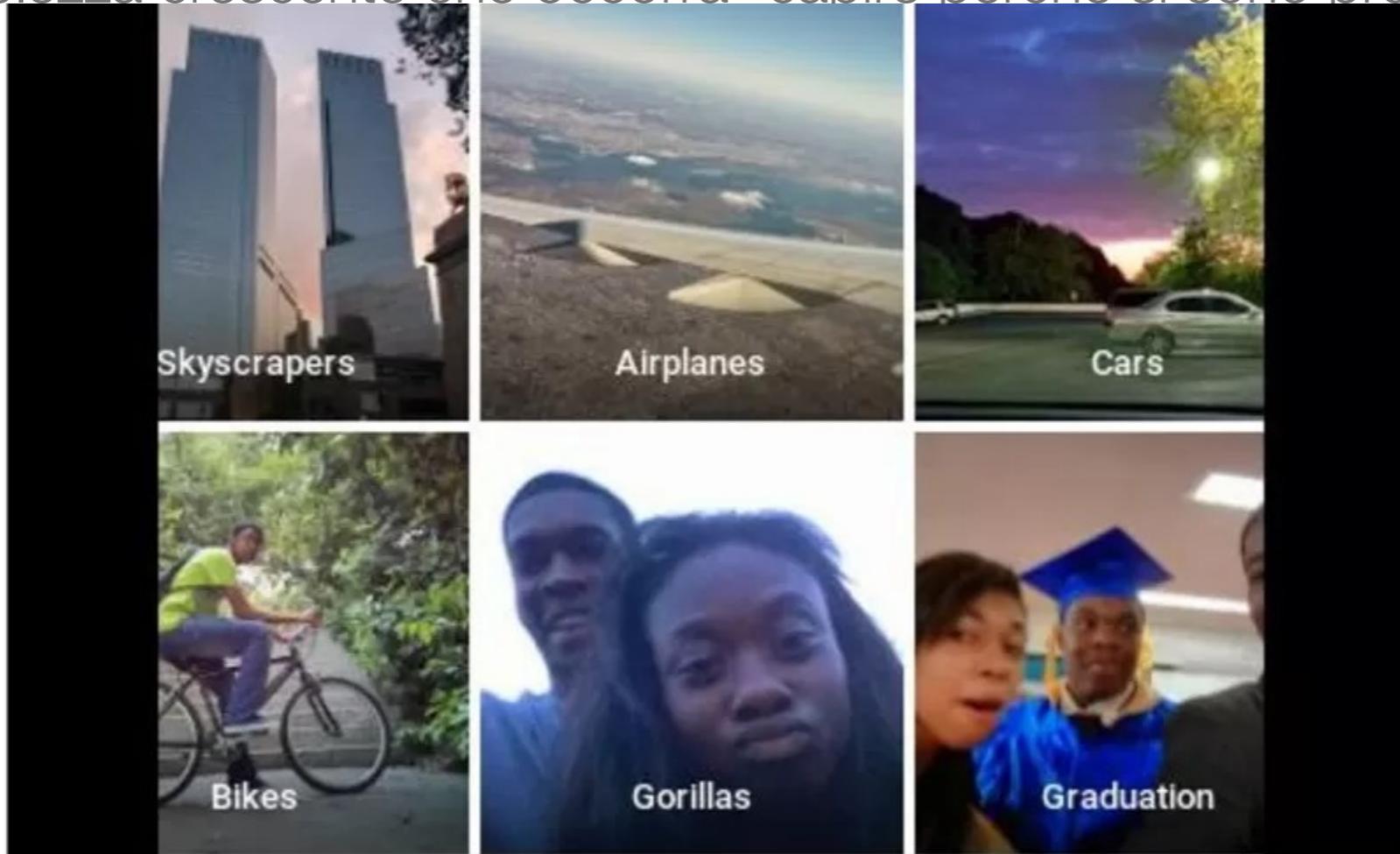
Stato attuale

- Consapevolezza crescente che occorre "capire perchè si sono prodotte conseguenze sbagliate o inaccettabili" (ex post)
 - violazioni di privacy
 - operazioni e decisioni ingiuste, non eque, viziate da pregiudizi razzisti o di genere — caso clamoroso di Google photos 2005
 - pregiudizi indotti dai dati di training di componenti "appresi" via ML
 - ◆ studiare nuovi metodi di V&V
 - ◆ regolamentare/legiferare
- **Troppo tardi?**
- **Verso un approccio "ex ante"?**

Stato attuale

- Consapevolezza crescente che occorre "capire perchè si sono prodotte con

- \
- (
- (



i o di

a ML

- -



TWITTER

- verso un approccio ex ante ?

Possiamo costruire sistemi **migliori**?

Possiamo costruire sistemi **migliori**?

Sistemi **migliori** per un **mondo migliore**

Possiamo costruire sistemi **migliori**?

Sistemi **migliori** per un **mondo migliore**

Possiamo costruire sistemi che rispondano ai principi e agli obiettivi dell'umanesimo digitale? Come?

Giusto e sbagliato, buono e cattivo —> etica

- Nella pratica dell'ingegneria del software ci si pone mai la domanda di ciò che sia bene o male, giusto o sbagliato?
- Queste domande vengono mai poste nell'insegnamento dell'ingegneria del software?
- Quando si insegna a come identificare i goals e i requirements?
- Quando si implementano e valutano diverse scelte alternative?
- Quando si fa V&V?

Tre dimensioni etiche

- La dimensione umana
 - proteggere e arricchire i valori delle persone
- La dimensione sociale
 - proteggere e arricchire una società giusta e democratica
- La dimensione naturale
 - (ri)stabilire un rapporto sostenibile con l'ambiente naturale

Mondo digitale e ambiente

- Le tecnologie digitali essenziali possono giocare un ruolo positivo fondamentale per progettare, monitorare e prevedere
- Ma
 - Situazione pre-AI: circa 10% emissioni CO2 dovute al digitale
 - email —> 4 g. email con attachment —> 50 g. (UE)
 - Problema esasperato dai costosissimi sistemi di apprendimento automatico
 - Il problema delle batterie

Code of ethics: utili, ma non ci aiutano in pratica



The cover of the ACM Code of Ethics and Professional Conduct features a dark blue background with binary code (0s and 1s) and faint mathematical symbols. The title 'The Code' is written in large, light blue letters. Below it, the full title 'ACM Code of Ethics and Professional Conduct' is displayed in white, followed by the subtitle 'Affirming our obligation to use our skills to benefit society'. A list of contents is provided at the bottom.

- General Ethical Principles
- Professional Responsibilities
- Professional Leadership Principles
- Compliance with the Code
- Case Studies
- Using the Code



The cover of the IEEE Code of Ethics has a light blue background with a circuit board pattern. The title 'IEEE Code of Ethics' is in bold blue text. Below the title is a paragraph of text, followed by a list of six principles.

IEEE Code of Ethics

We, the members of the IEEE, in recognition of the importance of our technologies in affecting the quality of life throughout the world, and in accepting a personal obligation to our profession, its members, and the communities we serve, do hereby commit ourselves to the highest ethical and professional conduct and agree:

1. To uphold the highest integrity, responsibility, and ethical conduct in practice;
2. to hold paramount, the welfare of the public with ethical design and development practices, privacy of others, and factors that might endanger the environment;
3. to improve the understanding and society of the consequences of our technologies, including their safety and security;
4. to avoid real or perceived conflicts of interest whenever possible, and to disclose them to affected parties when they do exist;
5. to avoid unlawful or unethical activities, and to reject corrupt practices;
6. to seek, accept, and perform technical work, to correct errors, to be truthful in stating claims or estimates, to use available data, and to acknowledge the contributions of others;
7. to maintain and improve our competence and to use our skills only within our areas of expertise, or after appropriate training and certification;
8. to maintain and improve our competence and to use our skills only within our areas of expertise, or after appropriate training and certification;

The cover of the IEEE Ethically Aligned Design report features a vibrant, abstract background with a human profile and colorful, ethereal shapes. The title 'ETHICALLY ALIGNED DESIGN' is in large white letters, with the subtitle 'A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems' below it. The IEEE logo and tagline 'Advancing Technology for Humanity' are in the top right corner.

Overview - Version II

IEEE
Advancing Technology for Humanity

ETHICALLY ALIGNED DESIGN
A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems

Code of ethics: utili, ma non ci aiutano in pratica



The cover of the ACM Code of Ethics and Professional Conduct features a dark blue background with binary code (0s and 1s) and mathematical symbols. The title "The Code" is written in large, light blue letters. Below it, the full title "ACM Code of Ethics and Professional Conduct" is displayed in white, followed by the subtitle "Affirming our obligation to use our skills to benefit society". A list of contents is provided at the bottom.

- General Ethical Principles
- Professional Responsibilities
- Professional Leadership Principles
- Compliance with the Code
- Case Studies
- Using the Code



The cover of the IEEE Code of Ethics has a light blue background with a circuit board pattern. The title "IEEE Code of Ethics" is in bold blue text. Below the title, a paragraph states that the members of the IEEE, recognizing the importance of technology in affecting the quality of life, commit themselves to the highest ethical and professional conduct. A list of principles follows.

IEEE Code of Ethics

We, the members of the IEEE, in recognition of the importance of our technologies in affecting the quality of life throughout the world, and in accepting a personal obligation to our profession, its members, and the communities we serve, do hereby commit ourselves to the highest ethical and professional conduct and agree:

1. To uphold the highest integrity, responsibility, and ethical conduct in practice;
1. to hold paramount, the welfare of the public with ethical design and development practices, privacy of others, and factors that might endanger the environment;
2. to improve the understanding and society of the consequences and implications of conventional and emerging technologies, including their benefits and risks;
3. to avoid real or perceived conflicts of interest whenever possible, and to disclose them to affected parties when they do exist;
4. to avoid unlawful or unethical activities, and to reject deceptive practices;
5. to seek, accept, and perform technical work, to correct errors, to be truthful in stating claims or estimates, and to use available data, and to acknowledge the contributions of others;
6. to maintain and improve competence and to use skills and experience for others only when appropriate, or after obtaining pertinent limitations;

The cover of the IEEE Ethically Aligned Design report features a vibrant, abstract background with a human profile and colorful, ethereal shapes. The title "ETHICALLY ALIGNED DESIGN" is in large white letters, with the subtitle "A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems" below it. The IEEE logo and tagline "Advancing Technology for Humanity" are in the top right corner. The text "Overview - Version II" is in the top left corner.

Overview - Version II

IEEE
Advancing Technology for Humanity

ETHICALLY ALIGNED DESIGN
A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems

Giusto e sbagliato, buono e cattivo —> **etica**

- Etica —> non porta a un verdetto, ma richiede una deliberazione che porti a decisioni di cui ci si assume responsabilità
- Principi dell'etica e deliberazione etica devono diventare pilastri della formazione in ingegneria del software
- Ricerca necessaria per capire come possono diventare parte del processo di sviluppo del software

Chi decide? Chi contribuisce?

- Come si possono affrontare i risvolti umani e sociali di un sistema?
- Devono essere lasciati nelle sole mani dei tecnologi (ingegneri del software)?
- La deliberazione deve poter coinvolgere altri soggetti; oltre ai "soliti" stakeholders, deve aprirsi a contributi da altre discipline (umanistiche, sociali/politiche, giuridiche, economiche, ...)
- Responsabilità collettiva

Progetti esistenti

- Bavarian Research Institute for Digital Transformation, SE A. Pretschner (SE), J. Nida-Rümelin (Philosophy): "Ethical Software Engineering By Design"
 - Deliberazioni etiche in processi agili
- Lero, Ireland, "Responsible Software Engineering" (B. Nuseibeh)

Formazione

- Come modificare gli approcci formativi?
- Gli ingegneri del software engineers devono saper praticare metodi di **co-design** e saper interagire con altri saperi
- Sfida educativa
 - Rompere i silos disciplinari mantenendo eccellenza tecnica

Rendere i curricula in Informatica DigHum-aware

- Bottom line: insegnare i principi di etica
 - dai principi generali dell'etica al "code of ethics" per gli studenti al "code of ethics" ethics nella professione
- Consentire/facilitare la scelta di corsi extra dalle scienze umane e scienze sociali
- Innovare con attività di "laboratorio" cross-disciplinare, mescolando studenti/mentori con diversi background culturali
 - il ruolo dell'"universitas"
- (Anche le altre discipline —e altri curricula— si devono aprire alle sfide dell'umanesimo digitale!)

Rompere i silos disciplinari

- Nella formazione e nella ricerca
- Ma allora dobbiamo anche rivedere i meccanismi di valutazione e di premialità
- Pur riconoscendo la necessità di trans-disciplinarietà, tendiamo a rintanarci nel nostro comodo silo quando poi si devono valutare le persone e i loro contributi

L'inferno è lastricato di buone intenzioni

Conclusione

Conclusione

Il futuro è nelle nostre mani

Conclusione

Il futuro è nelle nostre mani

e deve restare nelle nostre mani