Università degli Studi di Milano
Jean Monnet Centre of Excellence
"The impact of European Union Research and Innovation Policy upon Services of General Interest"
With the support of the Erasmus+ Programme of the European Union

Massimo Florio

BIG SCIENCE, BIG DATA, TECH GIANTS

SEEDS from CEEDS

Centre of Excellence in Economics and Data Science

4 Maggio 2022

OUTLINE

1. Dalla Big Science ai Big Data
2. Dai Big Data alle Tech Giants
3. Opportunità di benessere sociale sprecate
4. L'Unione Europea
5. La proposta di un'infrastruttura pubblica europea

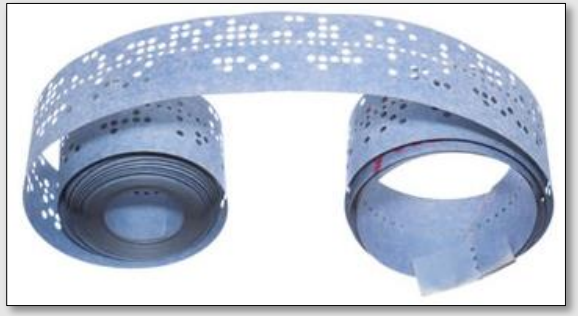
1. BIT AND BYTES: UN PO' DI STORIA

A	·–	J	·–––	S	···	2	··–––
B	–···	K	–·–	T	–	3	···–––
C	–·–·	L	·–··	U	··–	4	····–
D	–··	M	––	V	···–	5	·····
E	·	N	–·	W	·––	6	–····
F	··–·	O	–––	X	–··–	7	––···
G	––·	P	·––·	Y	–·––	8	–––··
H	····	Q	––·–	Z	––··	9	––––·
I	··	R	·–·	1	·–––––	0	–––––

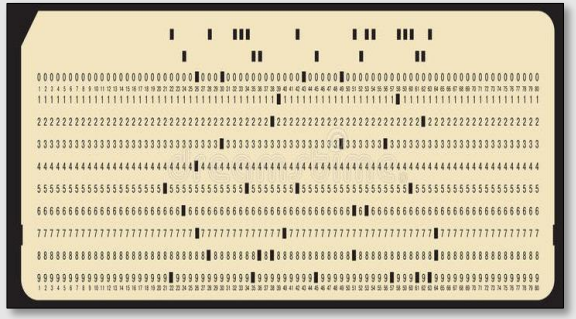
Codice Morse per telegrafo

Il concetto di 'bit' compare per la prima volta in un celebre articolo di Shannon (1949), "A Mathematical Theory of Communication".

- Dimensione definita in termini di bytes: una sequenza tipicamente di otto bit, cioè 2^3
- L' uso del byte come misura nasce in IBM nel 1956
- Gigabyte: 10^9 bytes
- Petabytes: 10^{15} bytes
- Exabytes: 10^{18} bytes
- Zettabytes: 10^{21} bytes
- Yottabytes: 10^{24} bytes



Nastri perforati per telescriventi



Schede perforate per i telai Jacquard e per le macchine calcolatrici

Una puntata di una serie Netflix ad alta definizione corrisponde a 7 Gigabytes

10 miliardi di foto su Facebook corrispondono a 1,5 Petabyte

1. SCIENZA E BIG DATA: TRE ESEMPI

- **Broad Institute** (Harvard and MIT)

  70 Petabytes genomic data

- **CERN**

  1 Petabyte per day of data

- **Square Kilometre Array SKA**

  4.9 Zettabytes per year in Australia

- **YouTube**

  24 Terabytes

- **Facebook (10 billion photos)**

  1,5 Petabyte

- **Google**

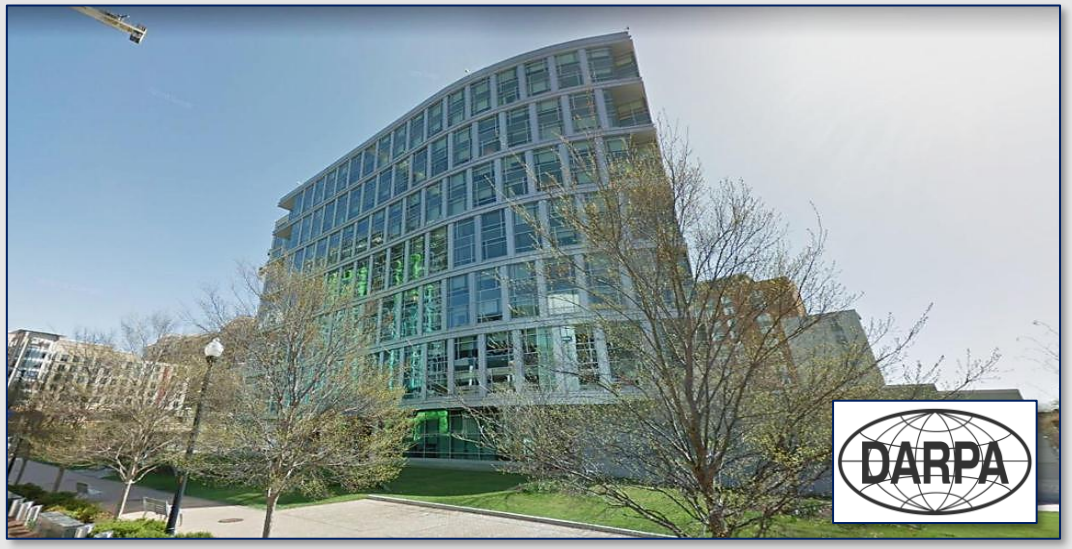
  15 Exabytes

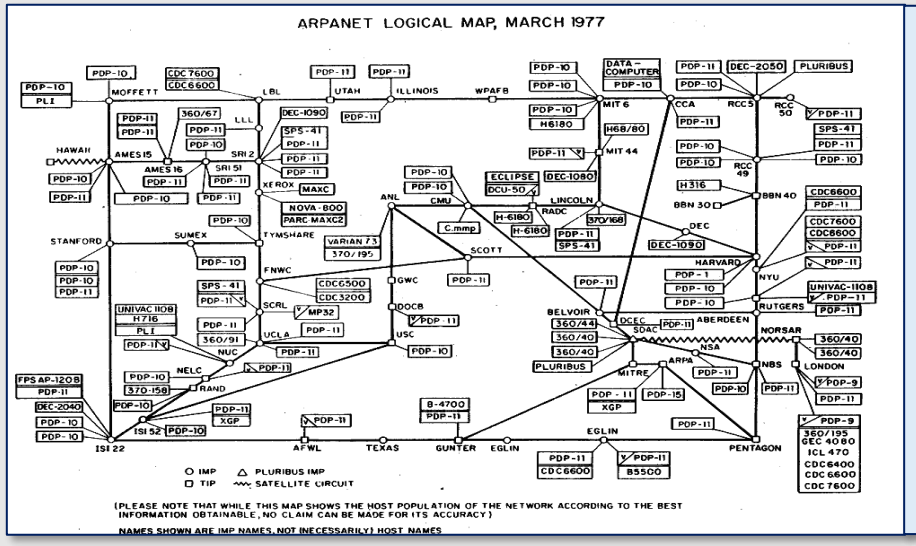
- **23andMe**

  10 million genome data

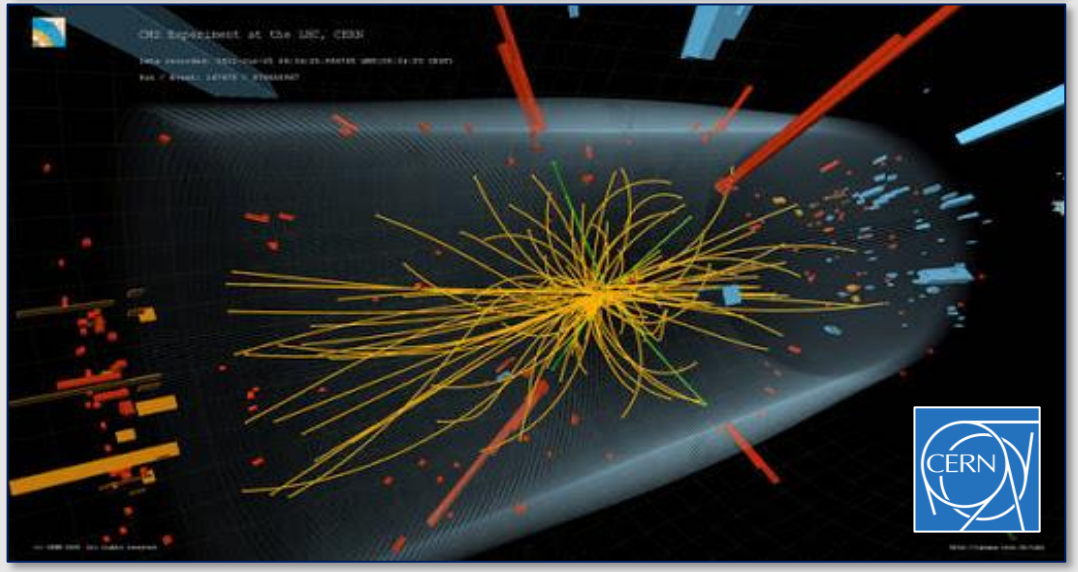
1. DALLA SCIENZA AI PROFITTI

I profitti delle Tech Giant si basano spesso su innovazioni create e donate dagli scienziati e da organizzazioni pubbliche o no-profit





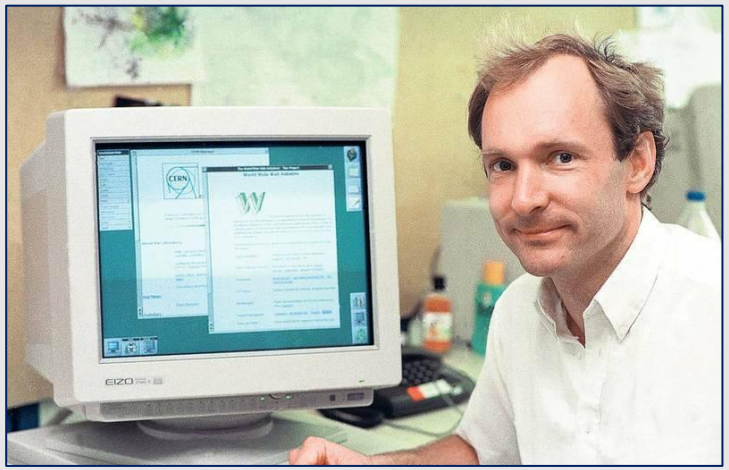
The Arpanet Logical Map in March 1977, the prototype of internet



The computer used by Tim Berners-Lee to build the first web server, 1980

2. IL NUOVO OLIGOPOLIO

- Come è successo che CERN e Tim Berners-Lee abbiano co-creato la ricchezza di Amazon e Jeff Bezos?



Ranking	Market value*
2. Microsoft	1,966.6
3. Apple	2,252.3
4. Amazon	1,711.8
5. Alphabet	1,538.9
6. Facebook	870.5
7. Alibaba Group	657.5
8. Tencent Holdings	773.8

https://www.statista.com/statistics/263264/top-companies-in-the-world-by-market-capitalization/

*billion $, year 2021

2. IL NUOVO OLIGOPOLIO E LA DISUGUAGLIANZA SOCIALE

1° Jeff Bezos
Amazon

2° Bill Gates
Microsoft

7° Mark Zuckerberg
Facebook

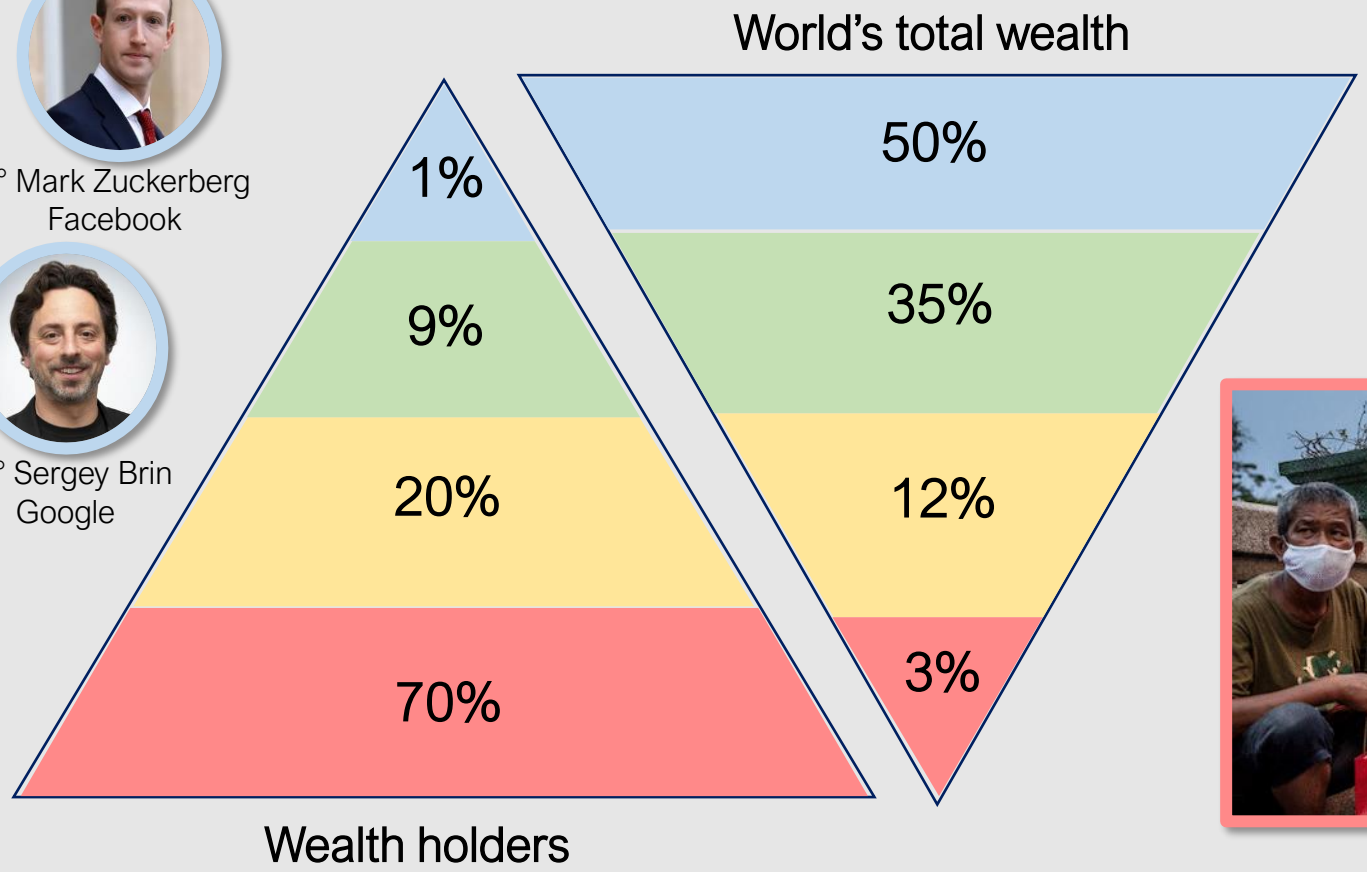
11° Steve Ballmer
Microsoft

13° Larry Page
Google

14° Sergey Brin
Google

17° Jack Ma
Alibaba

20° Ma Huateng
Tencent Holdings



"Nine of the 20 wealthiest people in the world are *tech tycoons* - one more than on the 2019 list. […] The 20 richest people in technology are worth a combined $740 billion [...]. In all, there are 241 technology billionaires on Forbes' 2020 billionaires list, up from 215 last year. Combined, they are worth a whopping $1.4 trillion - 16% of the total net worth for all 2,095 billionaires on the list this year." https://www.forbes.com/billionaires/

2. CINQUE TECH GIANTS USA: VALORE DI MERCATO

Name	Sales	Profit	Assets	Market Value
Apple	294	63.9	354.1	2,252.3
Amazon	386.1	21.3	321.2	1,711.8
Alphabet	182.4	40.3	319.6	1,538.9
Microsoft	153.3	51.3	304.1	1,966.6
Facebook	86	29.1	159.3	870.5
Tencent	70	23.3	203.9	773.8
Alibaba	93.8	23.3	250.1	657.5

https://www.forbes.com/lists/global2000/#28fd45b15ac0

Billion $

Altogether, the [177] technology companies on *Forbes*' Global 2000 … represent a staggering $17.9 trillion in market value—surging 73% year over year to nearly a third of the value of the entire U.S. stock market. The firms posted nearly $3.3 trillion in combined annual sales, up from about $3 trillion last year, when 161 tech companies made the list. Assets and profits also swelled, climbing 19% and 11%, respectively, to $5.2 trillion and $434.7 billion

https://www.forbes.com/sites/jonathanponciano/2021/05/13/worlds-largest-tech-companies-2021/?sh=2f12a3ef69bc

2. GLI INVESTIMENTI DELLE TECH GIANTS USA

- Tech Giants Five: market value triplicato tra il 2015 e il 2020
- Spesa pubblica e privata R&S USA 2020: $713 miliardi
- Tech Giants Five: $149 miliardi (25%del totale)
- Investimenti fissi: $131 miliardi
- Priorità R&S: metaverso, veicoli autonomi, salute, spazio, robotica, fin tech, cripto, quantum computing
- Negli ultimi tre anni acquisite circa 110 società
- Esempio, Microsoft Aprile 2020 acquista Nuance Communications (cloud di dati sanitari e software): 19.7 miliardi $. Google acquista Fitbit 2.1 miliardi $

Fonte: The Economist, 22 Gennaio 2022

2020

- In risposta alla pandemia: 170mila dipendenti e 82 aerei in più
- USA, il 6% di tutte le vendite al dettaglio
- Assets tangibili per 104 miliardi
- 11mila dollari al secondo di transazioni, 3,5 miliardi di pacchi (2019)
- Price/earning ratio è pari 118, tre-quattro volte Microsoft ed Apple
- Dinamica: il valore di mercato è quadruplicato circa dal 2016
- Ricavo pubblicitario: 11 miliardi di revenues
- 7% del mercato online in rapida salita
- I servizi di cloud (AWS) vendono 35 miliardi di dollari nell'ultimo anno ma competizione crescente

3. OPPORTUNITÀ SPRECATE

- I dati non sono di per sé un bene pubblico: **escludibili**
- **Consumo non rivale**: il loro uso da parte di un soggetto non toglie nulla alla possibilità di un altro soggetto di usarli a sua volta
- Come potranno essere definiti i diritti di proprietà e di accesso ai dati?
- Il movimento a favore degli '**open data**' si contrappone a quello che propone di creare delle istituzioni di mercato, che necessariamente implica diritti di proprietà

3. OPPORTUNITÀ SPRECATE

- Fenomeno '**data gravity**': l'attrazione che i dati esercitano su altri dati per le ovvie economie di scala, di densità, e di varietà che derivano dal fatto che la base informativa cui attingere sia la più ampia e diversificata possibile
- Opportunità sprecate dallo scollamento fra interessi privati e pubblici
- Tre sfide:

1. Nuove visioni della salute umana
2. Nuove visioni della salute del pianeta
3. Accesso universale all'informazione e alla conoscenza, e l'interconnessione fra persone

4. L'UNIONE EUROPEA

"Per quanto riguarda i dati personali – business to consumer- l'Europa è stata lenta e ora dipende da altri. Questo non può accadere con in dati industriali... L' ammontare di dati industriali nel mondo quadruplicherà nei prossimi cinque anni – e così pure le opportunità che ne derivano. Dobbiamo dare alle nostre imprese, PMI, start-up e ai ricercatori l'opportunità di estrarre tutto il loro potenziale… Ma la realtà è che l'80% dei dati industriali è raccolta e mai usata... Abbiamo bisogno di spazi comuni di dati, ad esempio nel settore dell'energia e della sanità. Questo supporterà ecosistemi innovativi in cui università, imprese e ricercatori potranno accedere ai dati e collaborare. Ed è per questo che costruiremo un cloud Europeo come parte di Next generation EU – basato su GAIA X"

Ursula von der Leyen

4. L'UNIONE EUROPEA

- Nel discorso sullo Stato dell'Unione 2020, von der Leyen ha annunciato che il 20% dei fondi di Next Generation EU (150 miliardi di Euro) andranno a progetti per una Europa digitale che comprenderebbe un "cloud" europeo per i dati industriali
- 8 miliardi di Euro:
  1. supercomputer del futuro
  2. investimenti nelle tecnologie 5G e 6G
  3. infrastruttura di banda larga
  4. e-identity europea
  5. investimenti nella intelligenza artificiale
  6. spazi comuni digitali

4. GAIA X

GAIA-X

- Federazione di clouds: adozione di certe regole e standards
- Modello “cloud dei cloud”: standard unico per la condivisione di dati, con una forma flessibile
- Ecosistema user-friendly
- 40 ‘use case’ in settori come Industria 4.0, Smart Living, Finanza, salute, settore pubblico, mobilità agricoltura
- Gaia X ed EOSC progetti in evoluzione

- Vantaggi Gaia-X:

1. soluzioni affidabili per l’uso di dati sanitari, nel rispetto delle normative, della protezione dei dati, della cybersecurity
2. capacità di archiviazione sicura dei dati, risorse computazionali a disposizione di ospedali e piccole-medie imprese
3. servizi di anonimizzazione o di pseudonimizzazione standardizzati

5. LA PROPOSTA

- **Soggetto sovranazionale europeo** che non abbia solo funzioni di coordinamento, ma abbia a tutti gli effetti l'autonomia manageriale, di bilancio, di capitale tangibile e intangibile, di personale dedicato
- Combinazione di **infrastruttura di ricerca e di impresa pubblica**
- L'agenzia sovranazionale che gestisse questa infrastruttura dovrebbe certamente ricorrere a **collaborazioni con terze parti** ma in una chiara distinzione di ruoli

5. LA PROPOSTA

- Piattaforma alternativa che dia le garanzie che le tech giants non danno
- Studi di fattibilità e dei modelli ‘proof of concept’
- Autonomo cloud europeo gestito direttamente dall’agenzia digital europe
- Sviluppo settore ICT europeo competitivo: Digital Europe deve investire in proprio
- Adeguamenti tecnologici delle reti di trasmissione dei dati
- Soluzioni tecnologiche innovative per portare internet a banda larga ovunque via satellite
- Progettare macchine o sistemi in grado di oltrepassare la soglia di 10^{18} FLOPS (floating point operations per second): medicina personalizzata, modelli metereologici, simulazioni del funzionamento del cervello umano ecc.

PER SAPERNE DI PIÙ

- https://www.data-infrastructure.eu/GAIAX/Navigation/EN/Home/home.html (video)
- The Economist, Artificial intelligence and its limits, Technology Quarterly, 13 Giugno 2020
- The Economist, The Economist, Special report: the data economy, Feb 22 2020
- The Economist 18 luglio Chipmaking. The incredbly shrinking machine
- Ruggiero 2019 https://www.agendadigitale.eu/infrastrutture/gaia-x-perche-e-importante-il-primo-cloud-su-scala-europea-litalia-ne-fara-parte/
- https://data-infrastructure.eu
- https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_20_1655
- https://www.agendadigitale.eu/infrastrutture/gaia-x-perche-e-importante-il-primo-cloud-su-scala-europea-litalia-ne-fara-parte/
- https://www.euractiv.com/section/digital/opinion/gaia-x-a-trojan-horse-for-big-tech-in-europe/
- https://ec.europa.eu/digital-single-market/en/europe-investing-digital-digital-europe-programme
- https://www.vox.com/2018/4/9/17204004/amazon-research-development-rd
- La privatizzazione della conoscenza (2021), M. Florio: https://www.laterza.it/scheda-libro/?isbn=9788858145371