



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Laboratorio di Big Data - Introduzione

**Gustavo Marfia**

Dipartimento delle Arti

# Di cosa parliamo oggi

- Una breve presentazione dei docenti e delle loro attività
- Una breve presentazione del corso
- Le origini dei Big Data
- Definizione e nascita dei Big Data
- Internet, un po' di storia
- Web 2.0 e oltre
- Evoluzione
- Rischi
- Discussione





ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Una breve presentazione

# Chi sono

- Laurea in Ingegneria delle Telecomunicazioni a Pisa nel 2003, Dottorato da Universita' della California, Los Angeles, nel 2009
- Lavoro nel settore dell'informatica dal 2002. Negli anni ho lavorato e/o speso periodi di ricerca presso Siemens Mobile Communications, Istituto Superiore Mario Boella di Torino, Microsoft Research Cambridge, Los Alamos National Laboratories
- Dal 2012 sono a Rimini, prima come Ricercatore e poi come Professore Associato
- Nel 2019 fono il laboratorio di ricerca VARLAB: Virtual and Augmented Reality, con due sedi, una a Rimini e una a Bologna
- Dal 2019 dirigo il Master in Innovation in eXtended Reality (MIXR) dell'Universita' di Bologna



# VARLAB: VIRTUAL AND AUGMENTED REALITY LAB

The VARLAB laboratory is the result of an ongoing collaboration between the Department for Life Quality Studies and the Department of Computer Science and Engineering aiming at designing and building distributed and immersive environments to experiment with virtual reality and augmented reality applications.

The establishment of the laboratory has been proposed by Gustavo Marfia (Principal Investigator) and endorsed by the following faculty members, Alessandro Amoroso, Ozalp Babaoglu, Daniela Baroncini, Paolo Bellavista, Alan Albert Bertossi, Daniela Calanca, Antonio Corradi, Lorenzo Donatiello, Fabriano Fabbri, Luca Foschini, Maurizio Gabbielli, Mariella Lorusso, Alberto Malfitano, Simone Martini, Antonella Mascio, Roy Menarini, Edoardo Mollona, Rebecca Montanari, Federica Muzzarelli, Aldopaoolo Palareti, Fabio Panzieri, Marco Patella, Giampaolo Proni, Simona Maria Segre Reinach, Ines Tolic.

The lab has been funded by the University of Bologna, under the AlmaAttrezature 2017 initiative with ca 120K euro and co-financed by the Departments of Computer Science (15K) and for Life Quality Studies (5K).

# VARLAB: VIRTUAL AND AUGMENTED REALITY LAB

## Director of the research group / Direttore del gruppo di ricerca

- [Gustavo Marfia](#)

## Faculty members / Docenti partecipanti

- [Daniela Calanca](#)
- [Lorenzo Donatiello](#)
- [Elena Piccolomini](#)

# People

## Researchers / Assegnisti di ricerca

- [Elena Morotti](#)
- [Shirin Hajahmadi](#)

## Ph.D. Students / Studenti di dottorato

- [Alessia Angeli](#)
- [Valentina Pellicioni](#)
- [Lorenzo Stacchio](#)
- [Maria Elena Tassinari](#)

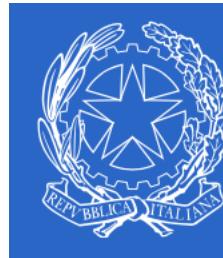


ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# VARLAB: VIRTUAL AND AUGMENTED REALITY LAB



FONDAZIONE  
CASSA DI RISPARMIO  
IN BOLOGNA



Ministero dello  
sviluppo economico



Programma operativo regionale  
Fondo europeo di sviluppo regionale

AEFFE



FONDAZIONE  
GOLINELLI  
l'intelligenza  
di esserci

IMAGE LINE<sup>®</sup>



YOOX

VARLAB: VIRTUAL AND AUGMENTED REALITY LAB



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO DI  
**SCIENZE PER LA QUALITÀ DELLA VITA**



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO  
**DELLE ARTI - DAR**

## Ambiti di ricerca

All'interno del Dipartimento di Scienze per la Qualità della Vita prima e del Dipartimento delle Arti dopo, l'attività di ricerca nell'ambito della realtà virtuale, aumentata e mista (VR/AR/MR), si concentra primariamente sulla progettazione, analisi e implementazione di applicazioni nelle aree del benessere, della moda, degli stili di vita e dell'arte, che integrino metodologie e tecniche di VR/AR/MR con tecnologie di raccolta e analisi dei dati (Internet delle cose, reti di sensori, reti sociali, intelligenza artificiale).



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO DI  
**SCIENZE PER LA QUALITÀ DELLA VITA**



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO DI  
**INFORMATICA - SCIENZA E INGEGNERIA**



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO  
**DELLE ARTI - DAR**



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

CIRI  
**ICT - TECNOLOGIE DELL'INFORMAZIONE  
E DELLA COMUNICAZIONE**



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

**ALMA MATER RESEARCH INSTITUTE FOR HUMAN-CENTERED  
ARTIFICIAL INTELLIGENCE**



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO DI  
**SCIENZE PER LA QUALITÀ DELLA VITA**



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO DI  
**INFORMATICA - SCIENZA E INGEGNERIA**



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO  
**DELLE ARTI - DAR**



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

DIPARTIMENTO DI  
**ARCHITETTURA - DA**



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

MASTER IN INNOVATION IN EXTENDED REALITY



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI



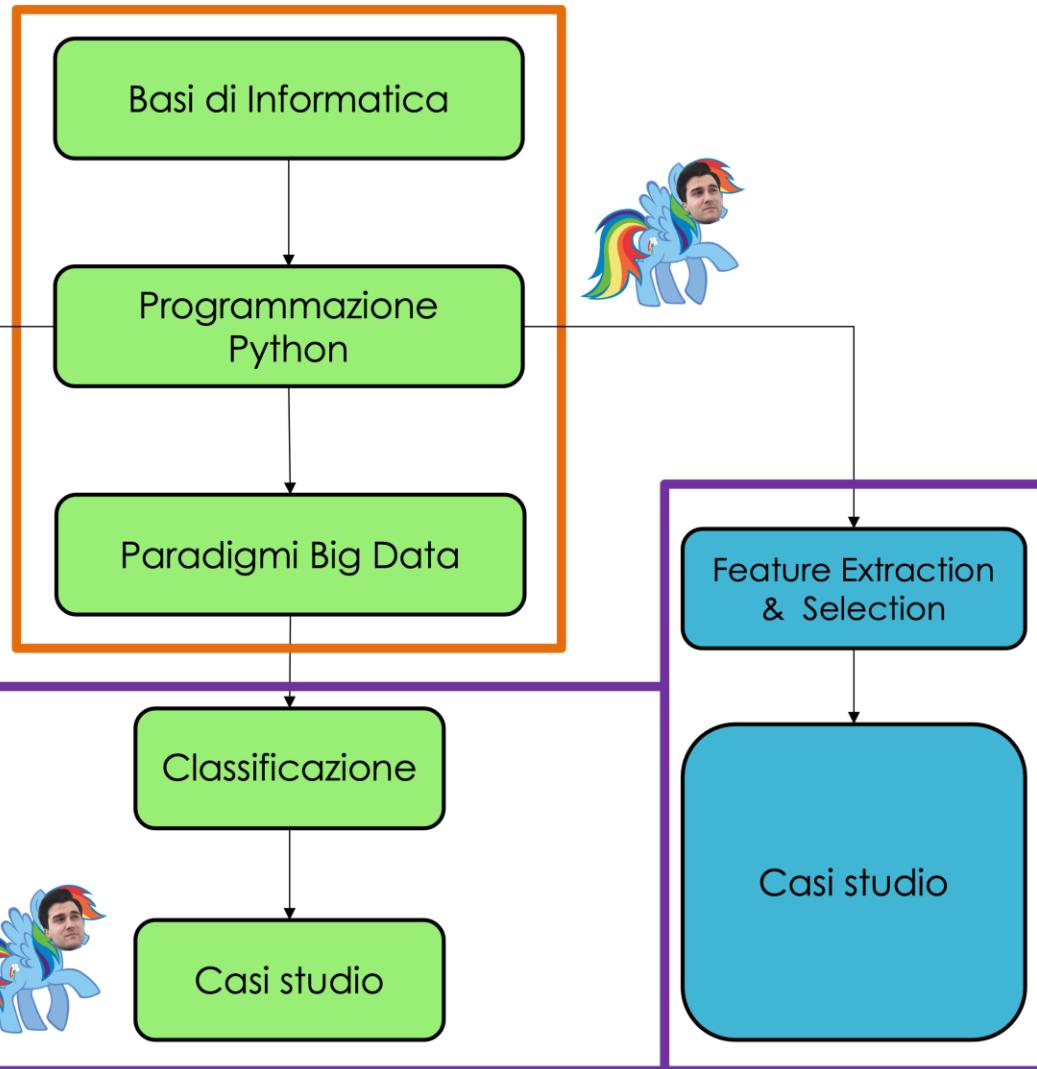
ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Il corso

# Programma

1

**BIG DATA**



**Data  
Science**



# Lorenzo Stacchio



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

VARLAB: VIRTUAL AND AUGMENTED REALITY LAB



## Random stuff lover



## CV breve

- **Studente di dottorato in CS;**
- **Laurea Mag. in CS – Tecniche del software;**
- **Laurea Triennale in CS;**

## Aree di ricerca

- **Machine and Deep Learning;**
- **Computer Vision;**
- **Augmented Reality;**

## Contatti



# Alessia Angeli – alessia.angeli2@unibo.it



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

VARLAB: VIRTUAL AND AUGMENTED REALITY LAB



## CV (in breve)

- Studente di dottorato in Data Science and Computation
- Laurea Magistrale in Matematica, curriculum generale/applicativo
- Laurea Triennale in Matematica

## Contatti



## Principali aree di ricerca

- Machine e Deep Learning
- Data Visualization e Augmented Reality



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

**Valentina Pellicioni - valentina.pellicion2@unibo.it**



**ALMA MATER STUDIORUM**  
UNIVERSITÀ DI BOLOGNA

VARLAB: VIRTUAL AND AUGMENTED REALITY LAB



## Contatti



## CV (in breve)

- Studentessa di dottorato in AI applicata al drug discovery
- Laurea magistrale a ciclo unico in Farmacia

## Principali aree di ricerca

- Computer Aided Drug Discovery
- Machine Learning e Deep Learning



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI



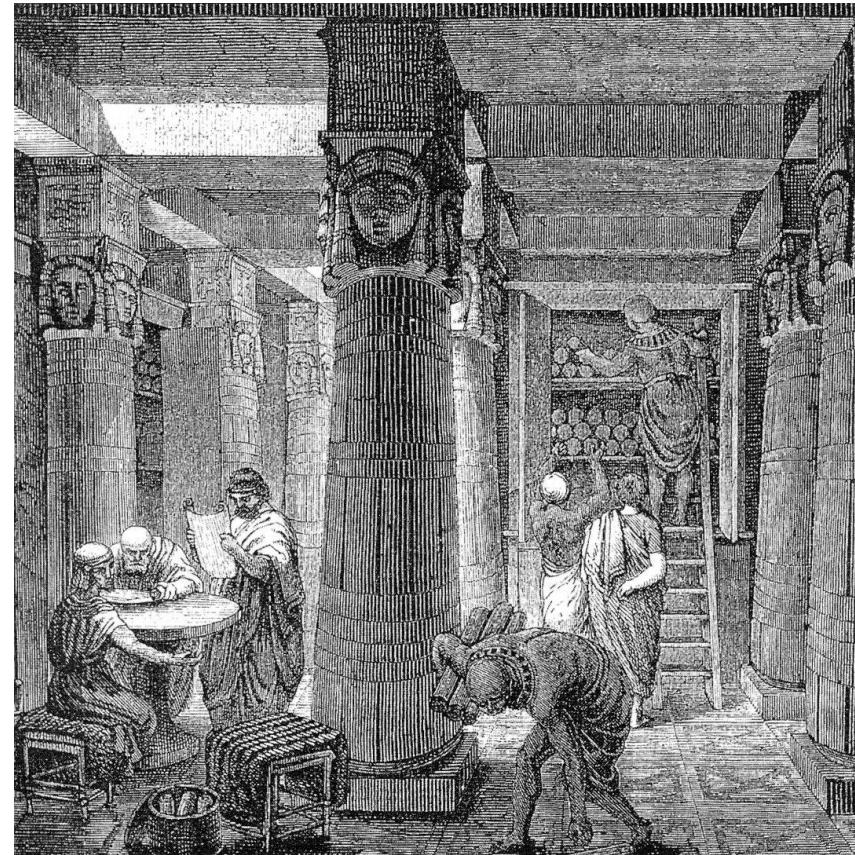
ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Le origini

# Storia dei Big Data: raccolta di volumi

La Bibliotecareale di Alessandria fu la più grande e ricca biblioteca del mondo antico ed uno dei principali poli culturali ellenistici. Andò distrutta nell'antichità, probabilmente più volte tra l'anno 48 a.C. e il 642 d.C.

Oggi giorno tutti i testi potrebbero facilmente essere salvati in una memoria USB da 500 GB (costo poche decine di euro).



# Storia dei Big Data: nascita della demografia

Il 27 febbraio 1662, John Graunt presentò alla Royal Society la prima tavola della vita, che produsse dai dati raccolti a Londra. Questa data è considerata la data di nascita della demografia.

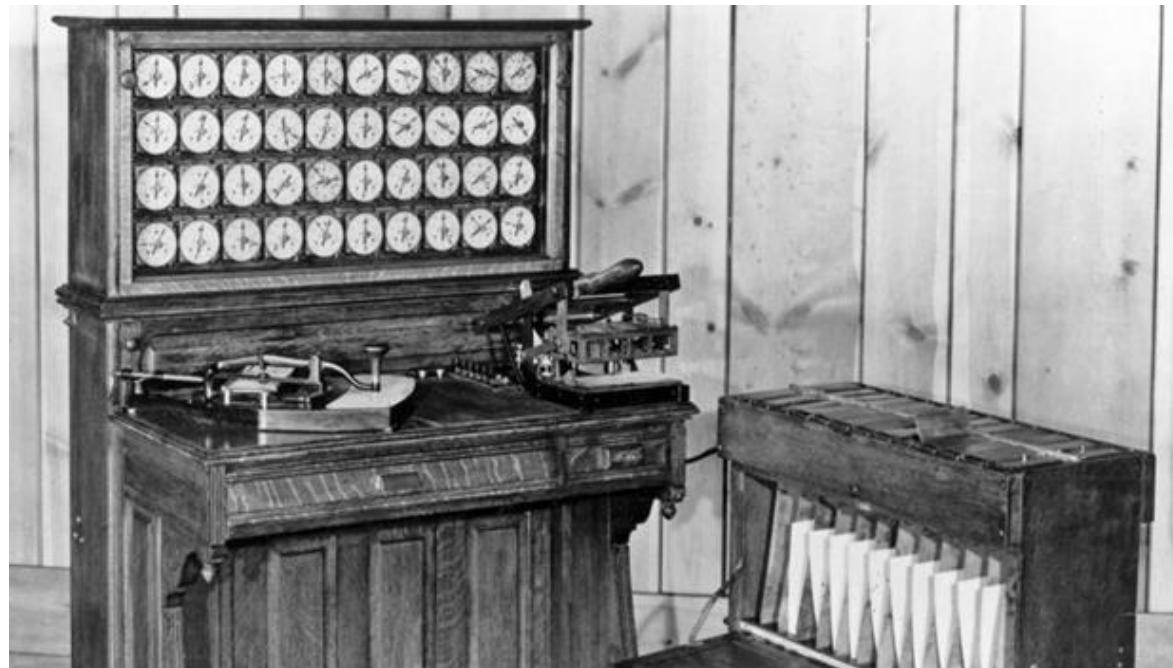
Graunt, insieme a William Petty, sviluppò i primi metodi statistici e di censimento umani che in seguito fornirono un quadro per la demografia moderna. È accreditato per aver prodotto la prima tavola della vita nel 1662, dando probabilità di sopravvivenza a ciascuna età. Graunt è anche considerato uno dei primi esperti in epidemiologia, poiché il suo famoso libro si occupava principalmente di statistiche sulla salute pubblica.

Il suo libro "Natural and Political Observations Made upon the Bills of Mortality" (1662) utilizzava l'analisi dei registri di mortalità nella Londra moderna quando Carlo II e altri funzionari tentarono di creare un sistema per avvertire dell'insorgenza e della diffusione della peste bubbonica nel città. Sebbene il sistema non sia mai stato veramente creato, il lavoro di Graunt nello studio dei registri ha portato alla prima stima statistica della popolazione di Londra.



# Storia dei Big Data: nascita dell'elaborazione dei dati

Le prime tabulatori di Herman Hollerith hanno aperto gli occhi al mondo sull'idea stessa di elaborazione dei dati. Queste macchine hanno anche posto le basi per la nascita di IBM. Il design di base di Hollerith - contatori elettromeccanici che leggono schede perforate di cartone - è stata la forma dominante di elaborazione dei dati dal 1890 fino all'arrivo dei computer elettronici commerciali negli anni '50.



# Storia dei Big Data: primo data center

Times Daily - Mar 31, 1971

Il governo degli Stati Uniti progetta il primo data center al mondo per memorizzare 742 milioni di dichiarazioni dei redditi e 175 milioni di set di impronte digitali su nastro magnetico.

MARTINSBURG, W. Va. (AP) —Armed with a complex of computers and a system of cross-checking that explores angles by the thousands, Uncle Sam keeps close watch here over his 95 million taxpayers.

Four modern computers spend every minute of the year continually pouring over the returns filed by America's bread-winners.

And with tax time at hand, and about half of the country yet to file, the machines at the National Computer Center here—the pride and chief detective of the International Revenue Service—are whirring away to pay refunds, process returns and catch citizens who make mistakes, either honestly or otherwise.

The IRS has centered its income tax memory bank in a one-story, plain-looking brick building outside this eastern West Virginia town since 1962.

# Storia dei Big Data: nascita del termine

Nel 1989, Erik Larson, in seguito autore di bestseller tra cui "The Devil in the White City" e "In The Garden of Beasts", ha scritto un pezzo per Harper's Magazine, che è stato ristampato sul Washington Post.

L'articolo inizia con l'autore che si chiede come tutta quella posta indesiderata arrivi nella sua casella di posta.

L'articolo include queste due frasi: "I detentori dei big data affermano di farlo a vantaggio del consumatore. Ma i dati possono essere utilizzati per scopi diversi da quelli originariamente previsti".



# Storia dei Big Data: crescita

Nel 2010 Eric Schmidt, presidente esecutivo di Google, racconta a una conferenza che ogni due giorni vengono creati tanti dati quanti ne sono stati creati dall'inizio della civiltà umana fino al 2003.



# Ma si parla anche di Small Data

Si sente molto parlare di come i big data, i dispositivi intelligenti e tutte le "omiche" (ad esempio genomica, proteomica, metabolomica e così via) trasformeranno la medicina, e lo faranno.

C'è un'altra forza che cambierà il modo in cui pensiamo e praticchiamo la salute, e sono i nostri piccoli dati (Small Data), piccoli dati derivati dalle nostre tracce digitali individuali.

*Estrin, Deborah. "Small data, where n= me." Communications of the ACM 57, no. 4 (2014): 32-34.*



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Definizione e nascita dei Big Data

# Definizione di Big Data

**big data** loc. s.le m. pl. Ingente insieme di dati digitali che possono essere rapidamente processati da banche dati centralizzate. ♦ L'immagine più suggestiva per comprendere Big Data la offre Dave Menninger di Greenplum: “Il pianeta è diventato un organismo vivente, che comunica continuamente e Internet ne rappresenta il sistema nervoso” [...] Tornando all'immagine del pianeta come organismo vivente, diventa lecito chiedersi quale sia il volto umano dei Big Data. (Luca Indemini, *Stampa.it*, 3 ottobre 2012, Tecnologia) • Pur essendo già piuttosto ambigua, l'espressione Big Data è spesso «chiarita» con aggettivi altrettanto nebulosi. Si prenda ad esempio «Big Data territoriali»: secondo gli scienziati informatici dell'Università del Minnesota, questo nuovo sottoinsieme di Big Data – costituito principalmente di insiemi di dati collegati alla posizione nello spazio, generati da sensori costantemente attivi, ottenuti da telecamere sempre migliori e conservati in banche dati in continua espansione – renderà i nostri spostamenti «sicuri, efficienti ed economici». (Evgeny Morozov, *Corriere della sera*, 21 ottobre 2012, La Lettura, p. 7, traduzione di Maria Sepa).

Espressione ingl. composta dall'agg. *big* ('grande') e dal s. *data* ('dati').



## Altra definizione

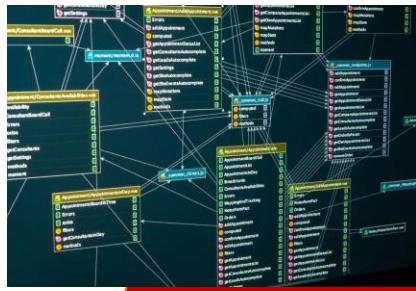
Big data è un termine generico per set di dati così grandi o complessi che le applicazioni di elaborazione dati tradizionali sono inadeguate.

Le sfide includono analisi, acquisizione, cura, ricerca, condivisione, archiviazione, trasferimento, visualizzazione e riservatezza delle informazioni.

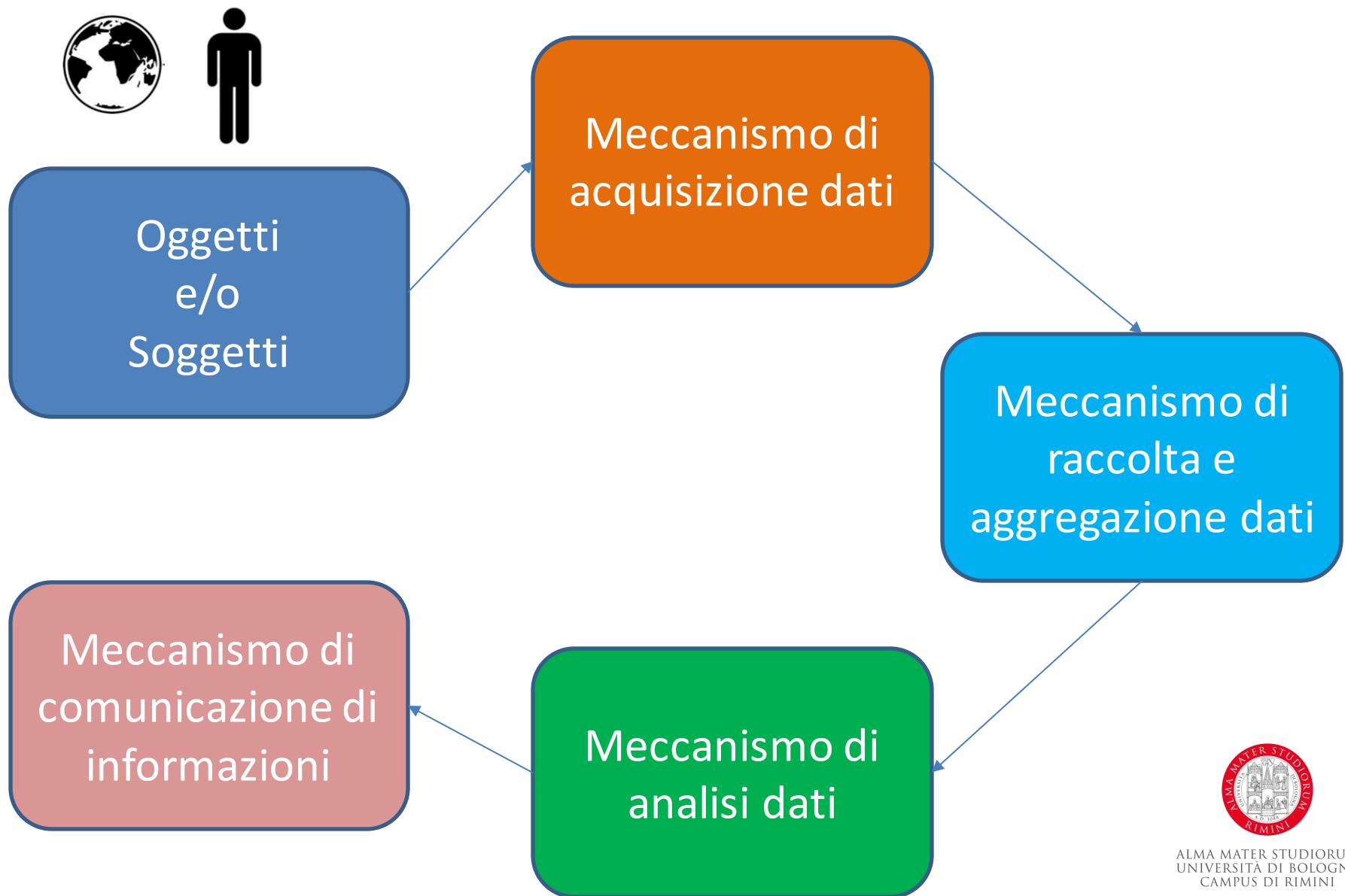
Il termine si riferisce spesso semplicemente all'uso di analisi predittive o altri determinati metodi avanzati per estrarre valore dai dati e raramente a una particolare dimensione del set di dati. – Wikipedia



# Ecosistema tecnologico dei Big Data



# Un semplice modello di acquisizione e analisi di dati



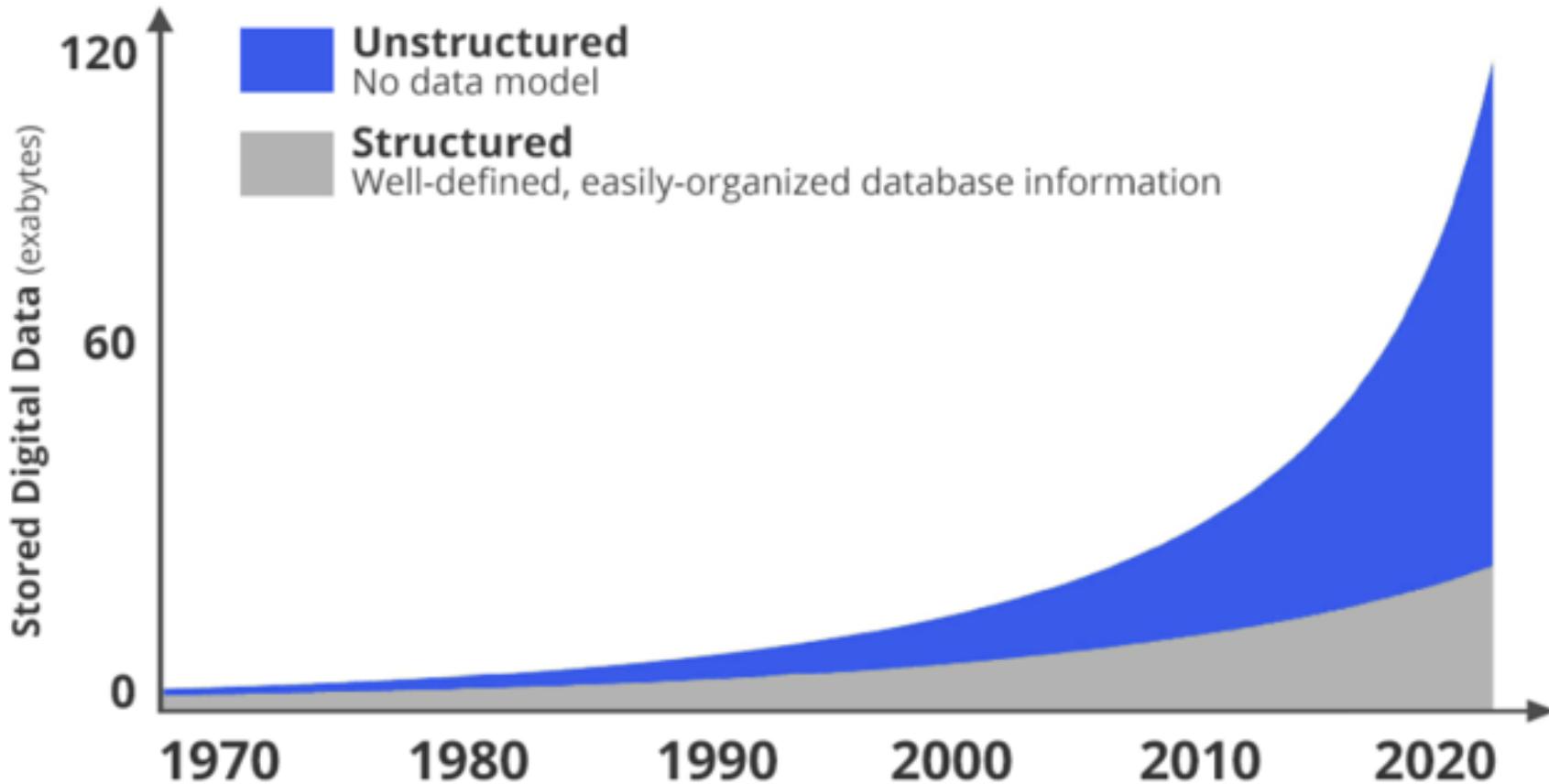
# E' una esempio di cio' che oggi accade



<https://components.omron.com/solutions/mems-sensors/environment-sensor>



# Crescita dei dati: strutturati e non strutturati



<https://seekingalpha.com/article/4317825-elastic-big-data-needs-effective-search-to-drive-value>





ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Internet, un po' di storia

## Le origini

Le origini di Internet sono radicate negli Stati Uniti degli anni '50. La Guerra Fredda era al culmine ed esistevano enormi tensioni tra il Nord America e l'Unione Sovietica. Entrambe le super-potenze erano in possesso di armi nucleari mortali e le persone vivevano nella paura dei lungo raggio attacchi a sorpresa.



Gli Stati Uniti si rendono conto che avevano bisogno di un sistema di comunicazione resiliente agli attacchi nucleari.

A quel tempo, i computer erano macchine grandi e costose utilizzate esclusivamente dai militariscienziati e personale universitario.



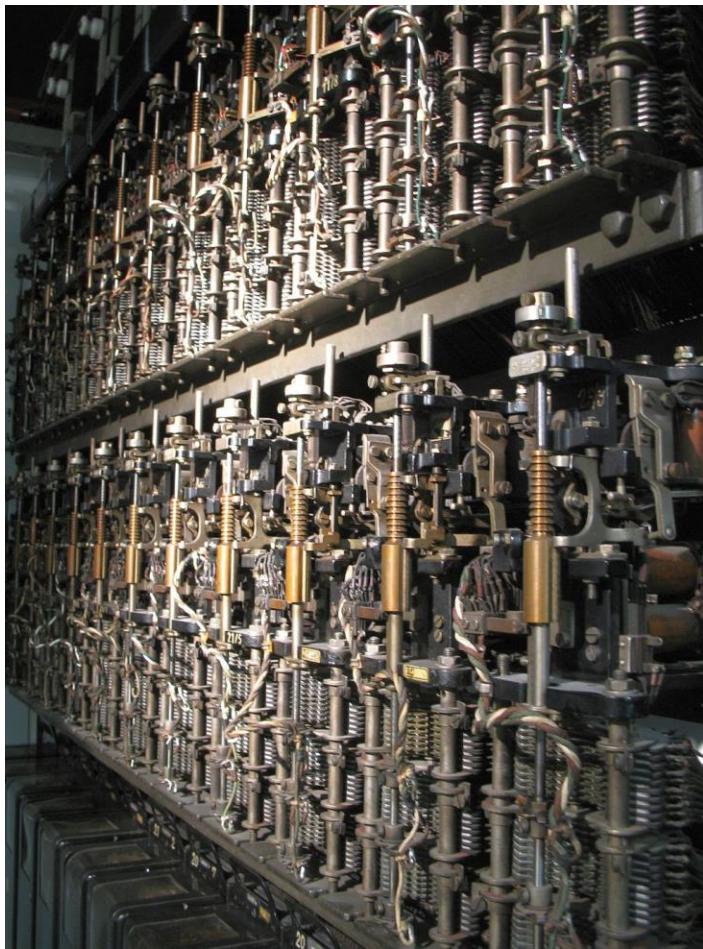
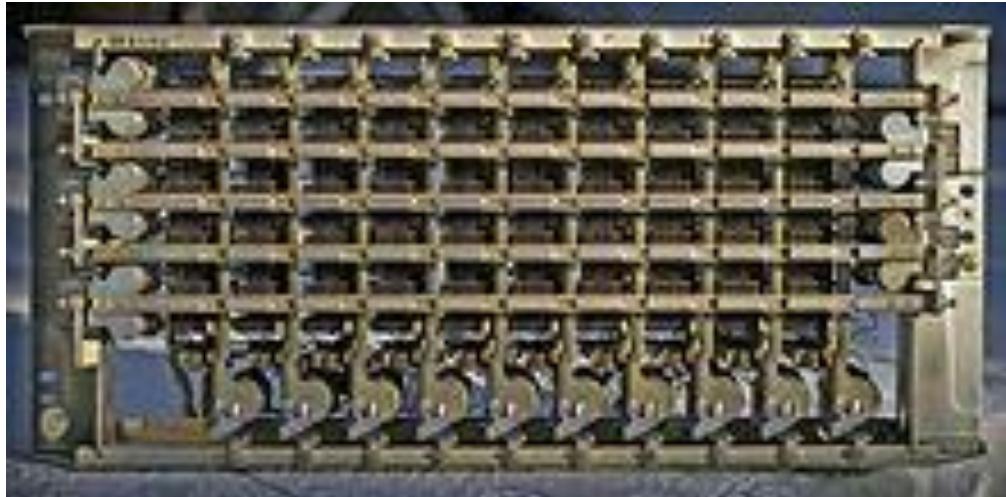
# Perche' questa necessita'?

Ricordiamoci come funzionava una rete di telecomunicazioni (telefono) alla fine degli anni '40, inizio anni '50.

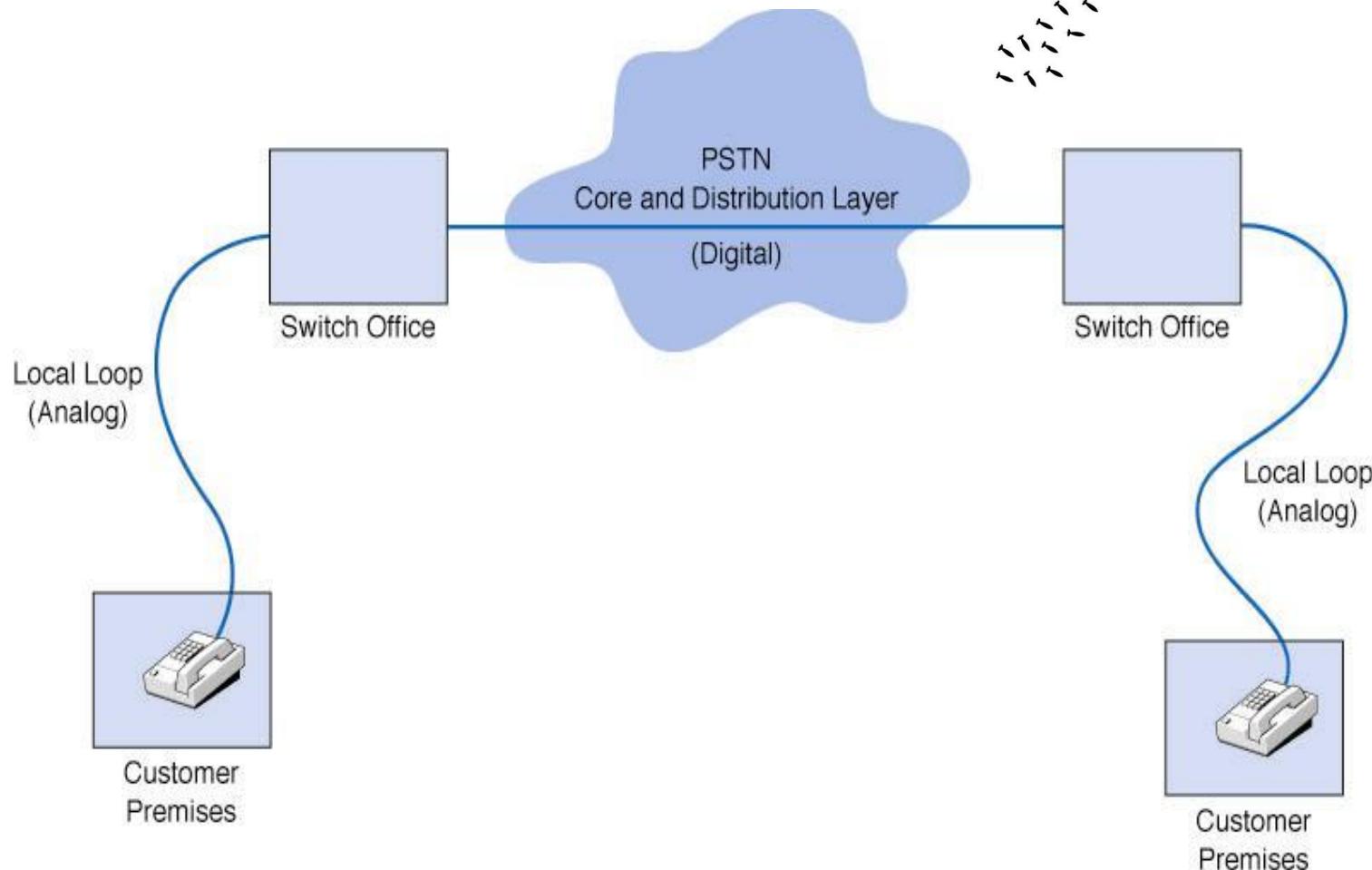


# Perche' questa necessita'?

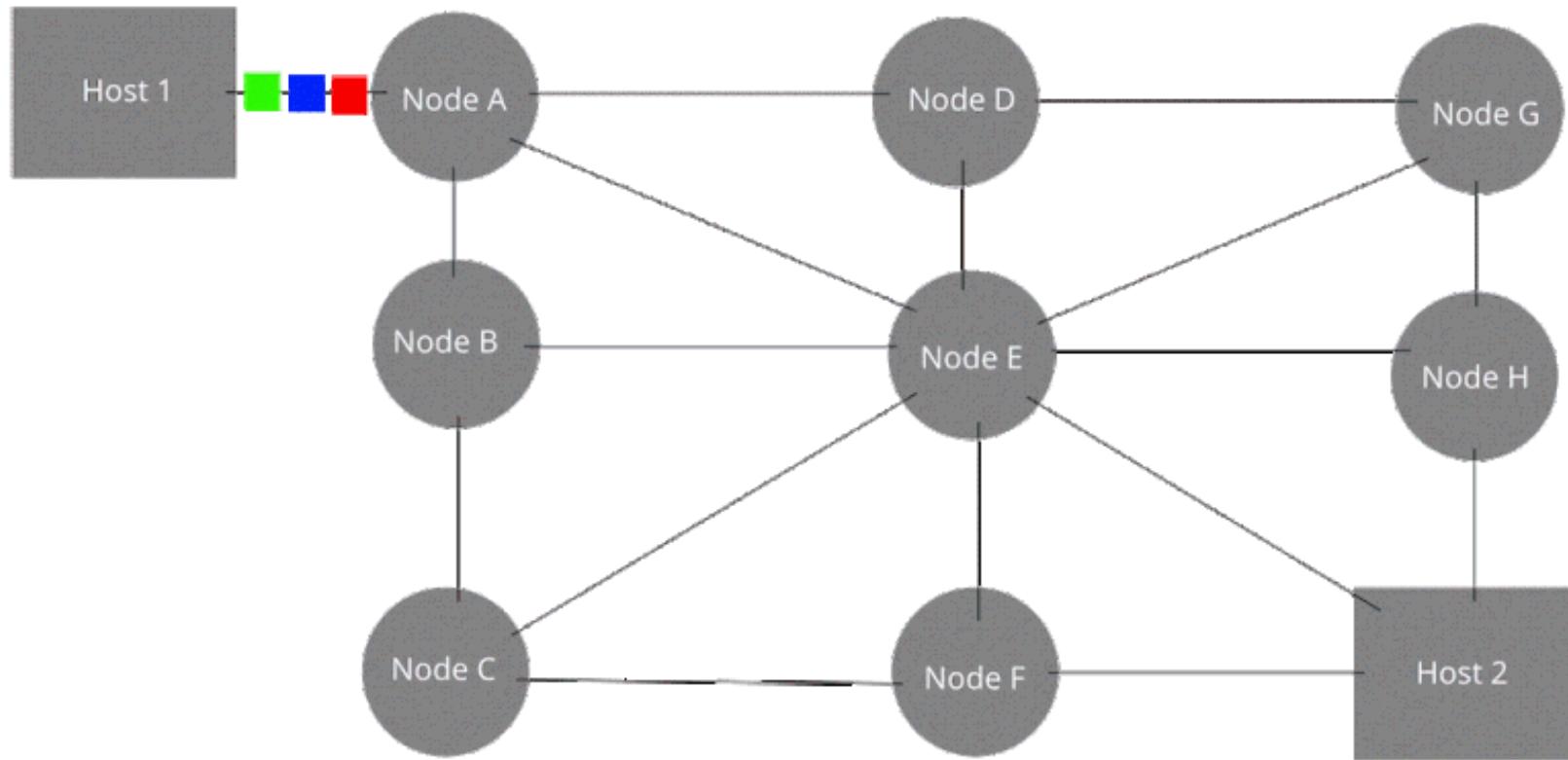
Fino agli anni '70, uso di apparecchi elettromeccanici.



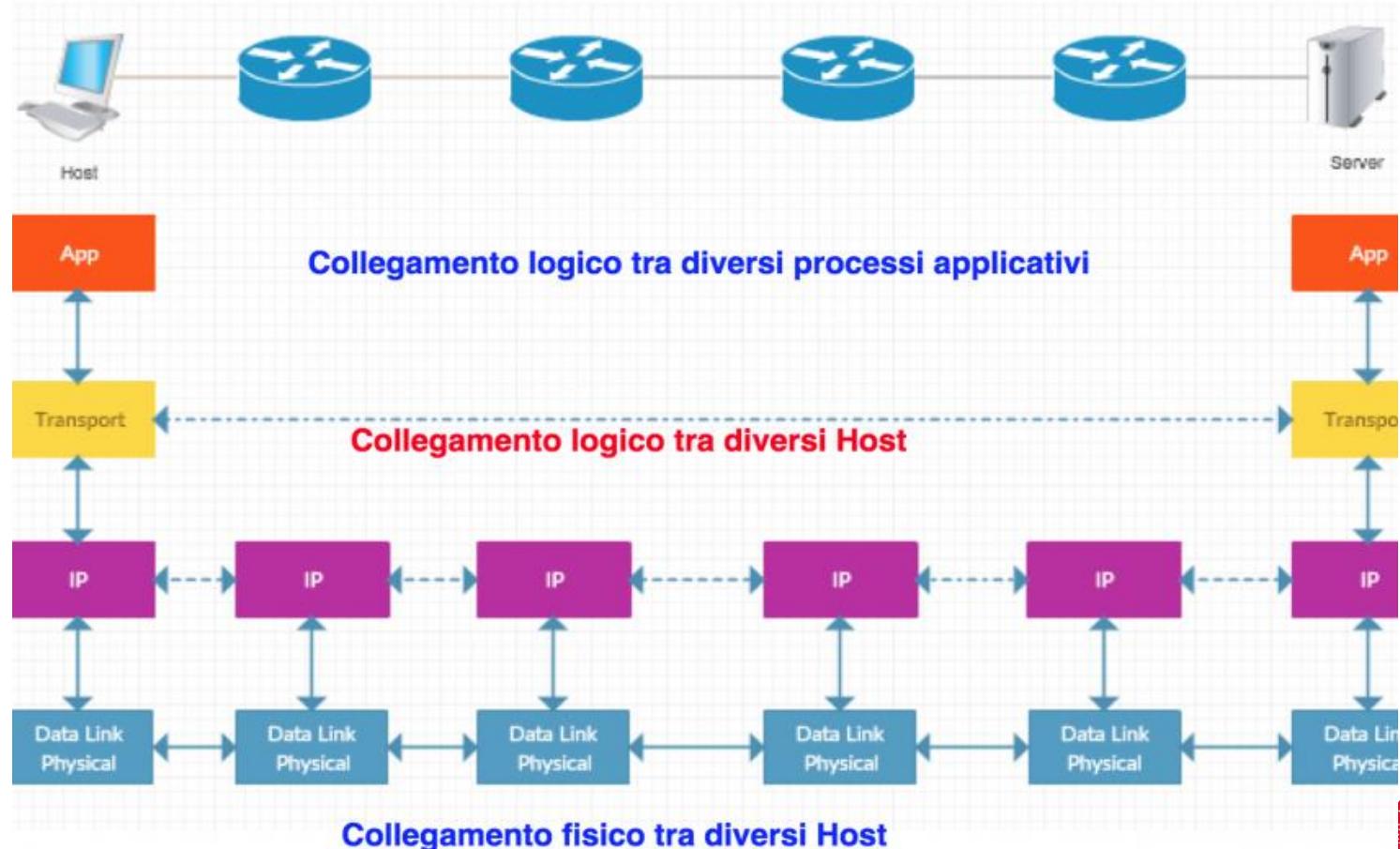
# Perche' questa necessita'?



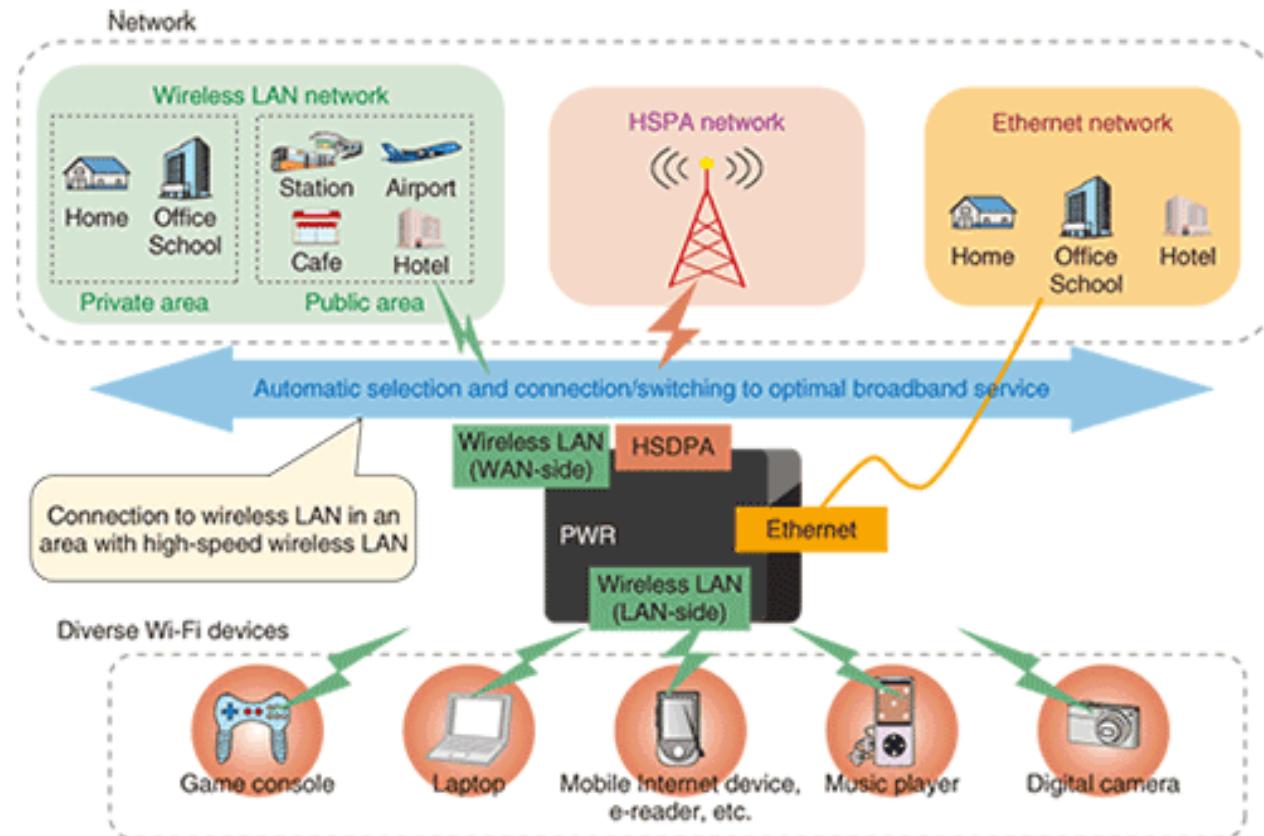
# L'intuizione alla base della resilienza: reti a commutazione di pacchetto



# L'intuizione alla base della flessibilità: stack protocollare TCP/IP



# TCP/IP: Una approccio per diversi tipi di tecnologie di accesso



# C'e' un Padre o Madre di Internet?

- Non uno, ma molti!
- In molti, in compenso, hanno lavorato all'idea a seguito dello stimolo ricevuto dal presidente Eisenhower negli anni '50



# 2012 INTERNET HALL OF FAME INDUCTEES

PIONEERS	GLOBAL CONNECTORS	INNOVATORS
Paul Baran	Randy Bush	Mitchell Baker
Vint Cerf	Kilnam Chon	Tim Berners-Lee
Danny Cohen	Al Gore	Robert Cailliau
Steve Crocker	Nancy Hafkin	Van Jacobson
Donald Davies	Geoff Huston	Larry Landweber
Elizabeth Feinler	Brewster Kahle	Paul Mockapetris
Charles Herzfeld	Daniel Karrenberg	Craig Newmark
Robert Kahn	Toru Takahashi	Raymond Tomlinson
Peter Kirstein	Tan Tin Wee	Linus Torvalds
Leonard Kleinrock		Philip Zimmermann
John Klensin		
Jon Postel		
Louis Pouzin		
Lawrence Roberts		

# 2013 INTERNET HALL OF FAME INDUCTEES

PIONEERS	GLOBAL CONNECTORS	INNOVATORS
David Clark	Karen Banks	Marc Andreessen
Dave Farber	Gihan Dias	John Perry Barlow
Howard Frank	Anriette Esterhuysen	François Flückiger
Kanchana Kanchanasut	Steve Goldstein	Stephen Kent
J.C.R. Licklider	Teus Hagen	Anne-Marie Eklund Löwinder
Bob Metcalfe	Ida Holz	Henning Schulzrinne
Jun Murai	Qiheng Hu	Richard Stallman
Kees Neggers	Haruhisa Ishida	Aaron Swartz
Nii Quaynor	Barry Leiner	Jimmy Wales
Glenn Ricart	George Sadowsky	
Robert Taylor		
Stephen Wolff		
Werner Zorn		

# 2014 INTERNET HALL OF FAME INDUCTEES

## PIONEERS

Douglas Engelbart

## GLOBAL CONNECTORS

Dai Davies

Susan Estrada

Demi Getschko

Frank Heart

Masaki Hirabaru

Dennis Jennings

Erik Huizer

Rolf Nordhagen

Steve Huter

Radia Perlman

Abhaya Induruwa

Dorcus Muthoni

Mahabir Pun

Srinivasan Ramani

Michael Roberts

Ben Segal

Douglas Van Houweling

## INNOVATORS

Eric Allman

Eric Bina

Karlheinz Brandenburg

John Cioffi

Hualin Qian

Paul Vixie

# 2017 INTERNET HALL OF FAME INDUCTEES

## GLOBAL CONNECTORS

Nabil Bukhalid

Ira Fuchs

Shigeki Goto

Mike Jensen

Ermanno Pietrosemoli

Tadao Takahashi

Florencio Utreras

Jianping Wu

## INNOVATORS

Jaap Akkerhuis

Yvonne Marie Andrés

Alan Emtage

Ed Krol

Tracy LaQuey Parker

Craig Partridge



## Prossimo futuro

- Piu' banda
- Minor latenza
- Sempre piu' host grazie a IPv6





ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Web 2.0 e oltre

# Ma la storia continua...

## Web 3.0: From Data**base** to Data**bank**



### Global Library

- Existing information was gathered into a single database
- Brick and mortar libraries were disrupted

### Global Publishing

- Individuals gained the ability to create information in a global database
- Companies like Facebook and Google monetized data

### Web 3.0

### Global Databank

- Individuals have the potential to monetize their own data
- Cryptocurrencies and blockchains are the foundation of the global data**bank**

FOR DISCUSSION ONLY NOT FOR DISTRIBUTION CONFIDENTIAL/SUBJECT TO CHANGE

Prosumer



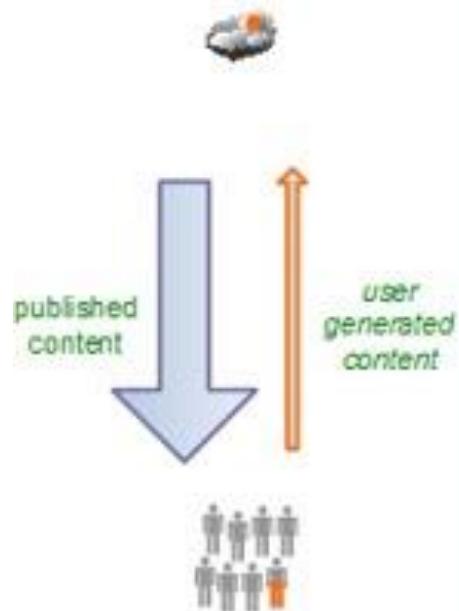
ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Prosumer: Producer + consumer

## Web 1.0

"the mostly read-only Web"

250,000 sites



45 million global users

1996

## Web 2.0

"the wildly read-write Web"

80,000,000 sites



1 billion+ global users

2006

## Web 3.0

"the wildly write-read Web"

800,000,000 sites



8 billion+ global users

2016



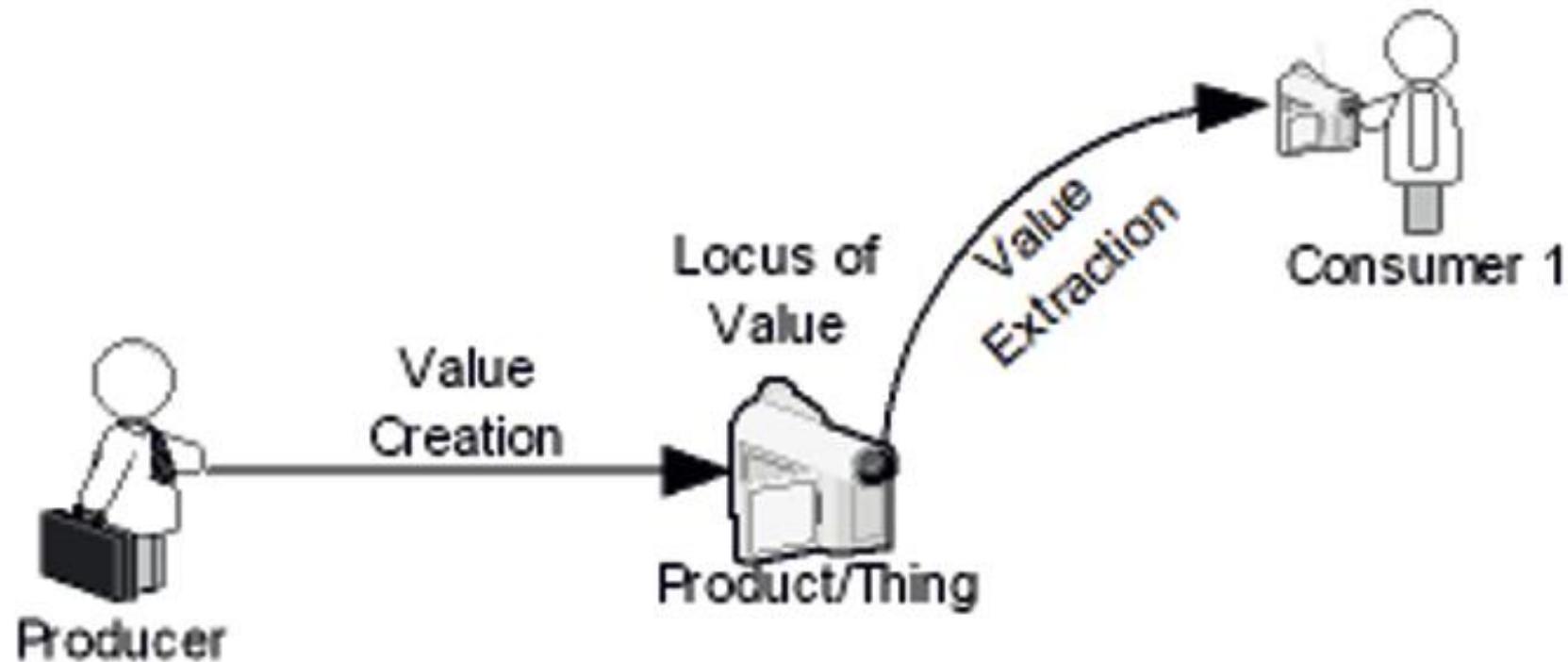


Fig 4. Phases of Value Chain.



# Creazione del valore post Web 2.0

Nell'era del Web 2.0, la creazione e la ricerca del  
valore ha un senso molto più ampio  
odels



# Web 2.0, alcuni esempi

Web 1.0	Web 2.0
Ofoto	Flickr
Akamai	BitTorrent
Mp3.com	Napster, iTunes
Britannica Online	Wikipedia
Personal Websites	Blogging
Publishing	Participation
Content Management Systems	Wikis
Directories (taxonomy)	Tagging (folksonomies)

T. O'Reilly, online: <http://oreilly.com/web2/archive/what-is-web-20.html>





ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Web 2.0 e processi decisionali

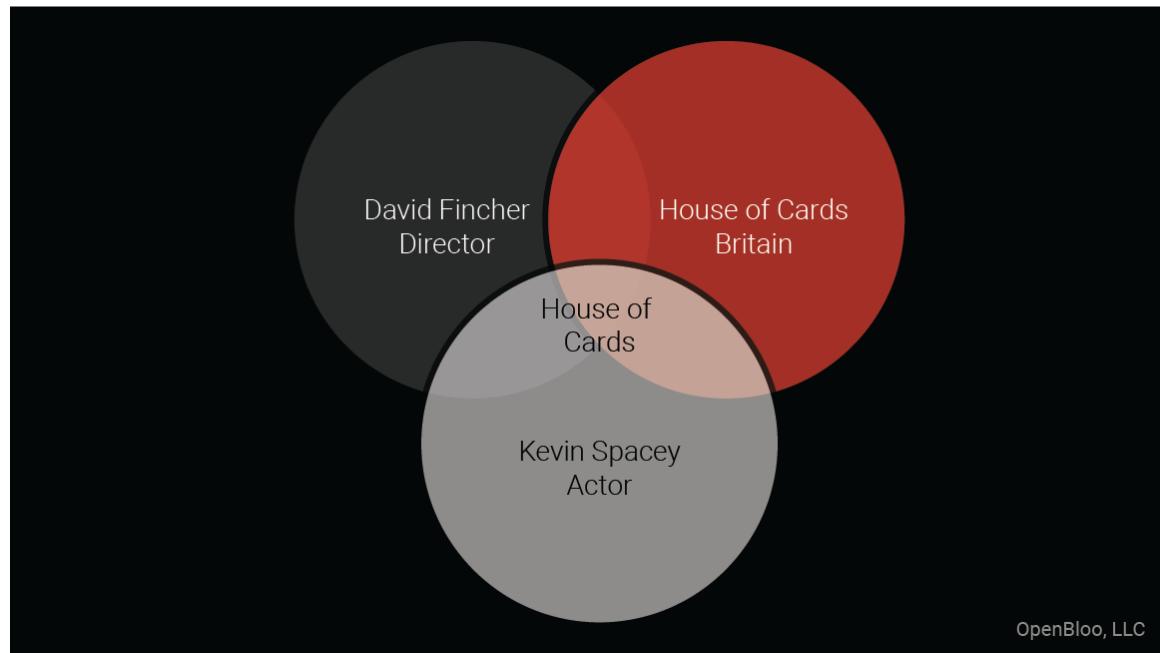
# Netflix prize

Al tempo, Netflix pubblica oltre 100 milioni di valutazioni, di 480 mila utenti per circa 18 mila film al fine di predire rating futuri



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Netflix: House of Cards



Produzione di House of Cards individuando regista, il produttore e gli attori



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Goldcorp challenge

Goldcorp rilascia i rilievi geologici in suo possesso, promettendo \$575K a chi potesse aiutarla nella ricerca nella sua carta geologica.

- Partecipare al challenge
- I partecipanti inviano dati per la ricerca
- A seguito della campagna di crowd sourcing sono stati rinvenuti 8 milioni di nuovi siti (precedentemente inaccessibili)

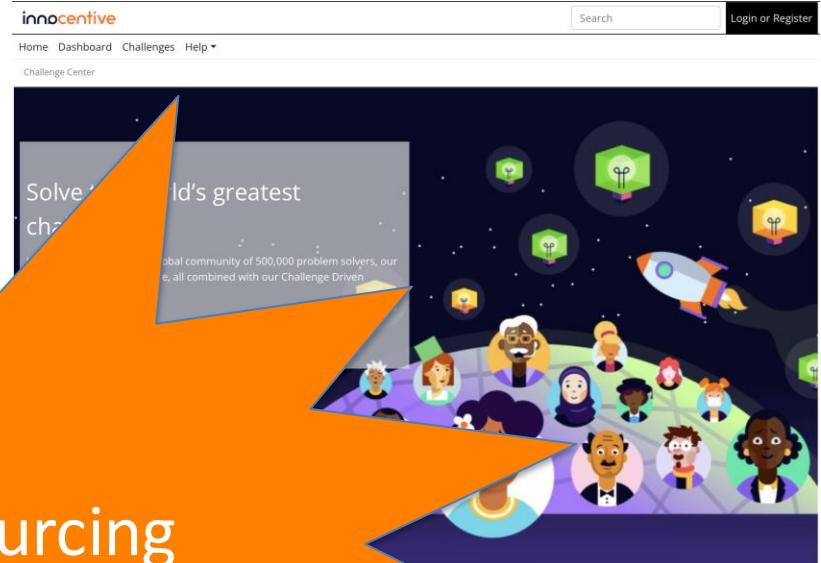
# DARPA Challenge

## Experimental Crowd-derived Combat-support Vehicle Design Challenge



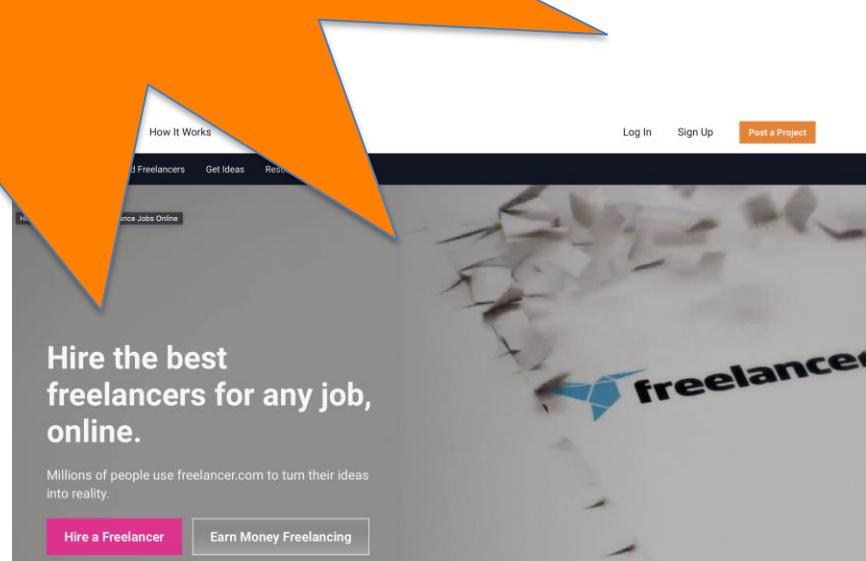
# Nascono nuove piattaforme collaborative

- <http://www.innocentive.com>



Crowd sourcing

- <http://www.vwork.it>



# Manifattura e logistica

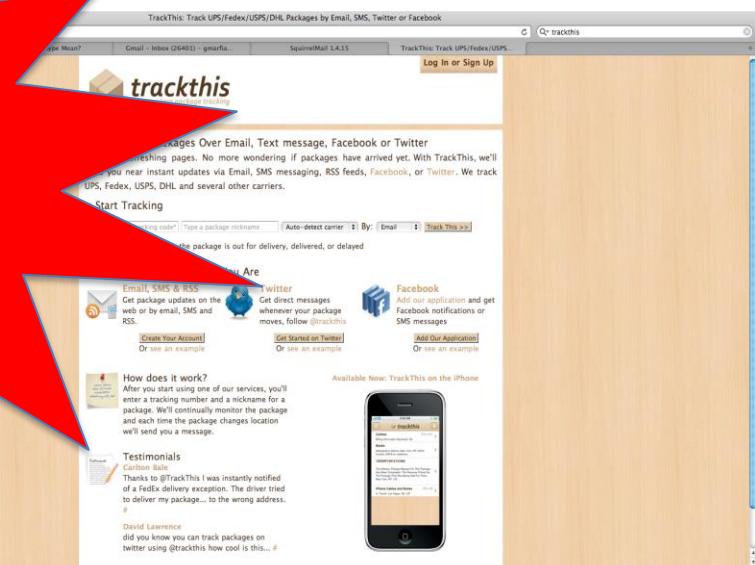
- <http://www.mfg.com>



- <http://www.trackthis.com>

– Inoltra i dati di spedizione su  
realtà aumentata, Facebook

Mash up, ma il  
servizio e'  
dismesso ora



# Marketing e vendite

Nuovo modo per creare contenuti per siti aziendali



# Marketing e vendite

- Facebook, twitter  
– Contenuto generato dai utenti  
– Fidelizzazione
- Nuove piattaforme (ad esempio HubSpot)  
– Ottimizzano il marketing online  
– Semplificano l'utilizzo di strumenti web 2.0

Social networking,  
virtual community

Search engine  
optimization, blogs



# Tesi

*La ricerca stessa, intesa come utilizzo di strumenti online che diano risposte a problemi reali, è un valore*

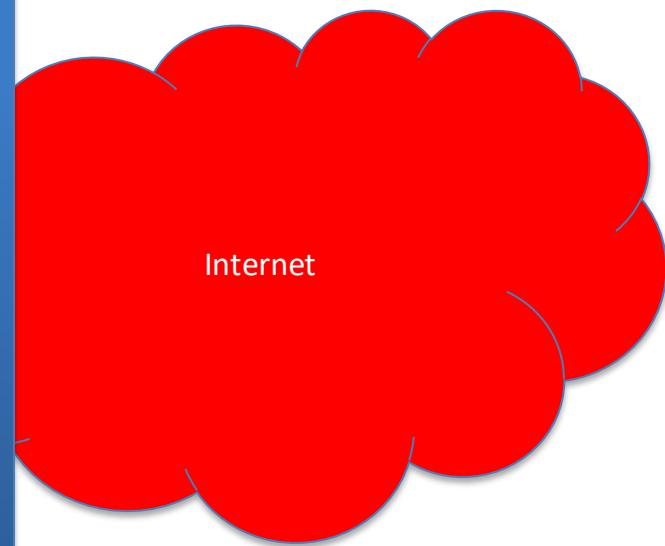


# Nuovo ecosistema

Decisore

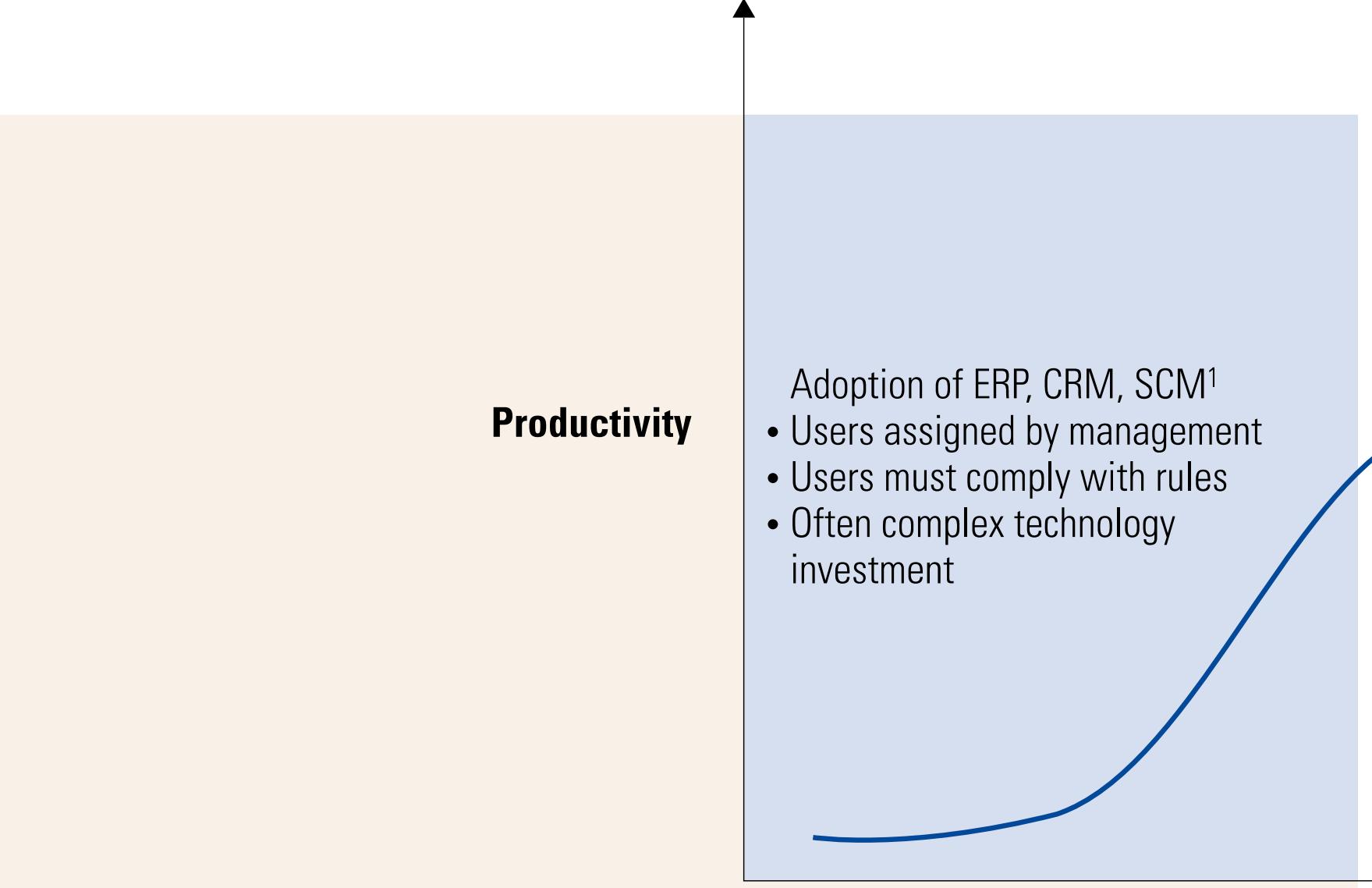


Ricerca di  
soluzioni di  
alta qualità



# Concludendo...

## Productivity

- 
- Adoption of ERP, CRM, SCM<sup>1</sup>
  - Users assigned by management
  - Users must comply with rules
  - Often complex technology investment

# Concludendo...

## Participations

For what purpose?

Content generation

the organization to collaborate on codifying/ managing knowledge, sharing best practices, communicating, and coordinating activities.

- **Broad collaboration/ communication, metadata**

Community building

**Large-scale community building**  
Enable creation of large-scale communities where experts a

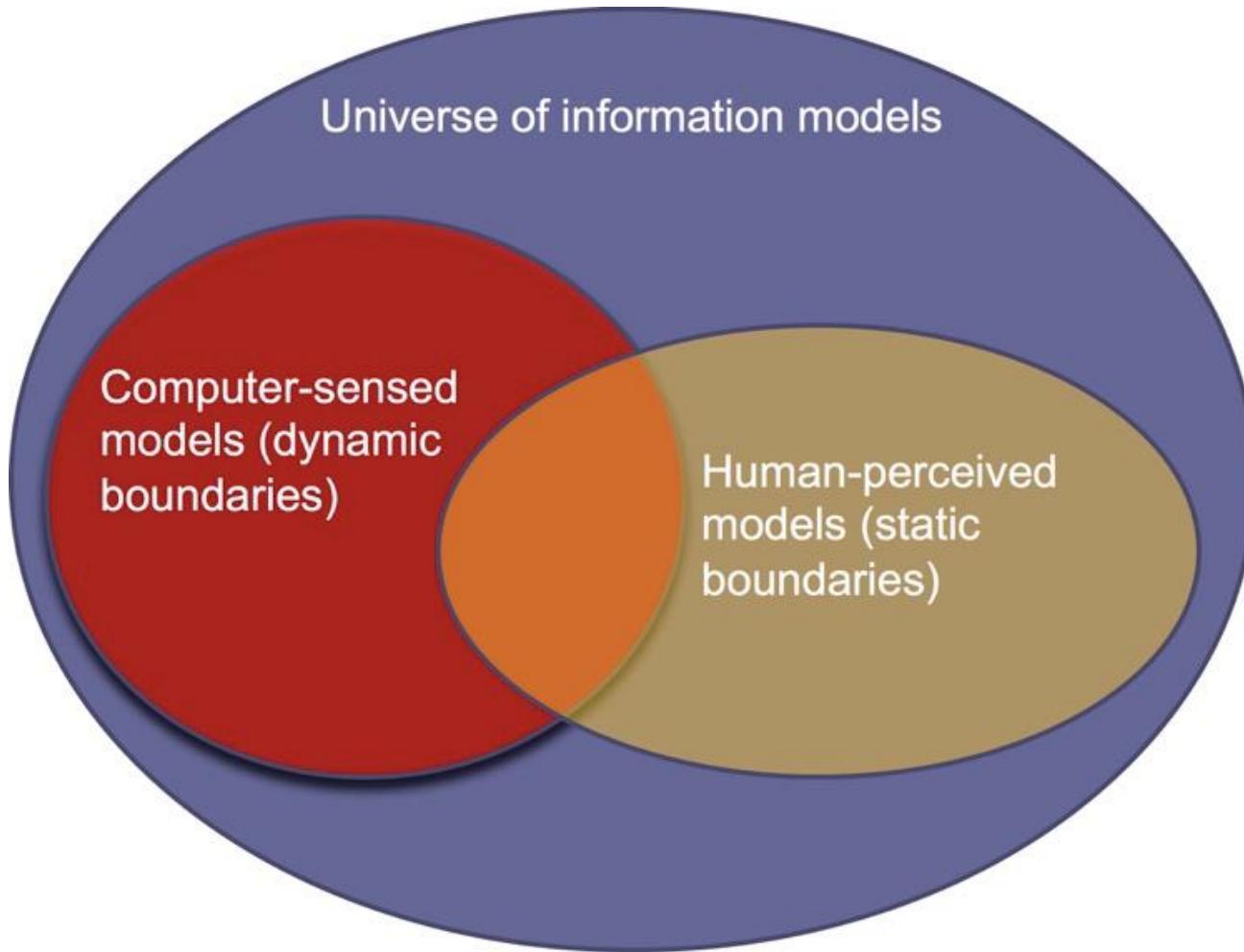
- **Broad collaboration/communication, metadata, social graphing**



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Evoluzione

# Modelli informativi



Marfia, Gustavo, and Giovanni Matteucci. "Some remarks on aesthetics and computer science." *Studi di estetica* 12 (2018).



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Personal computer

## Primi anni '80

- Dimensioni importanti
- Utilizzato per lo più da esperti

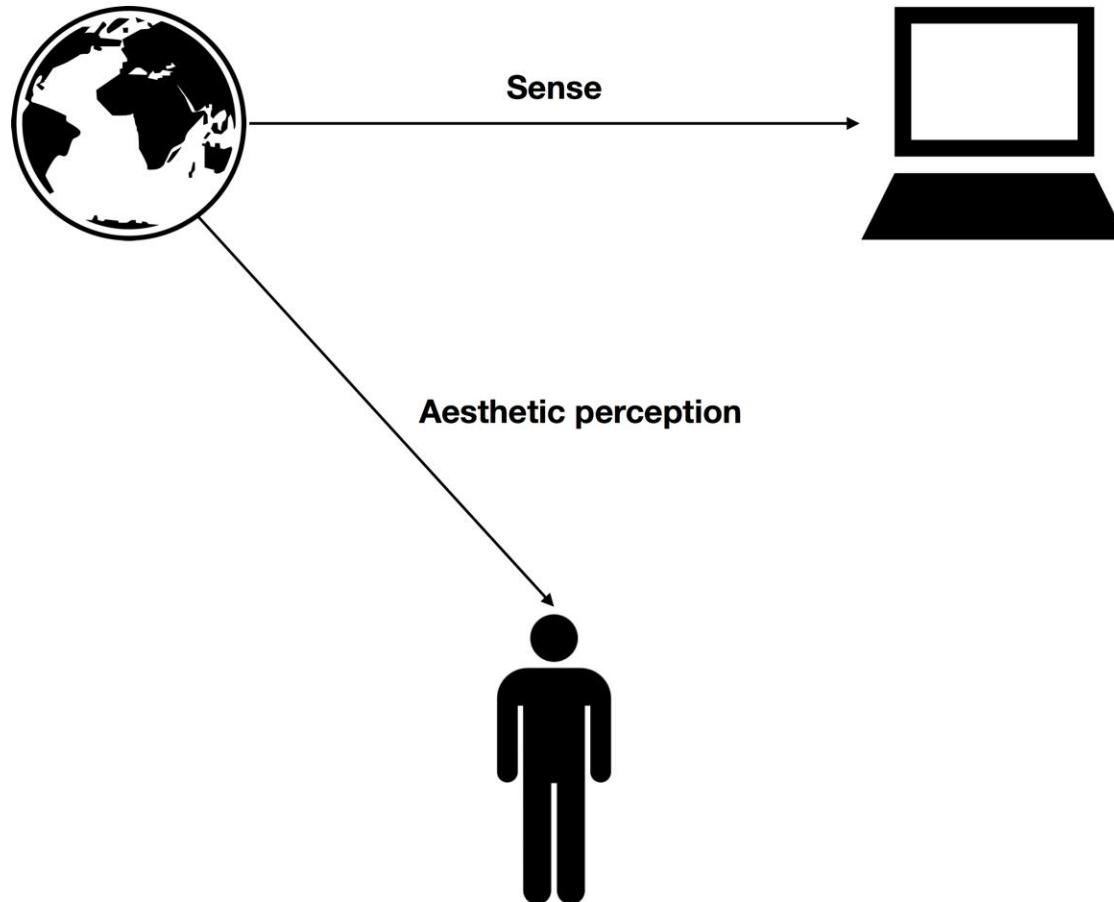


## Anni 2000

- Leggero
- Utilizzato da chiunque



# Dimensione cognitiva e percettiva

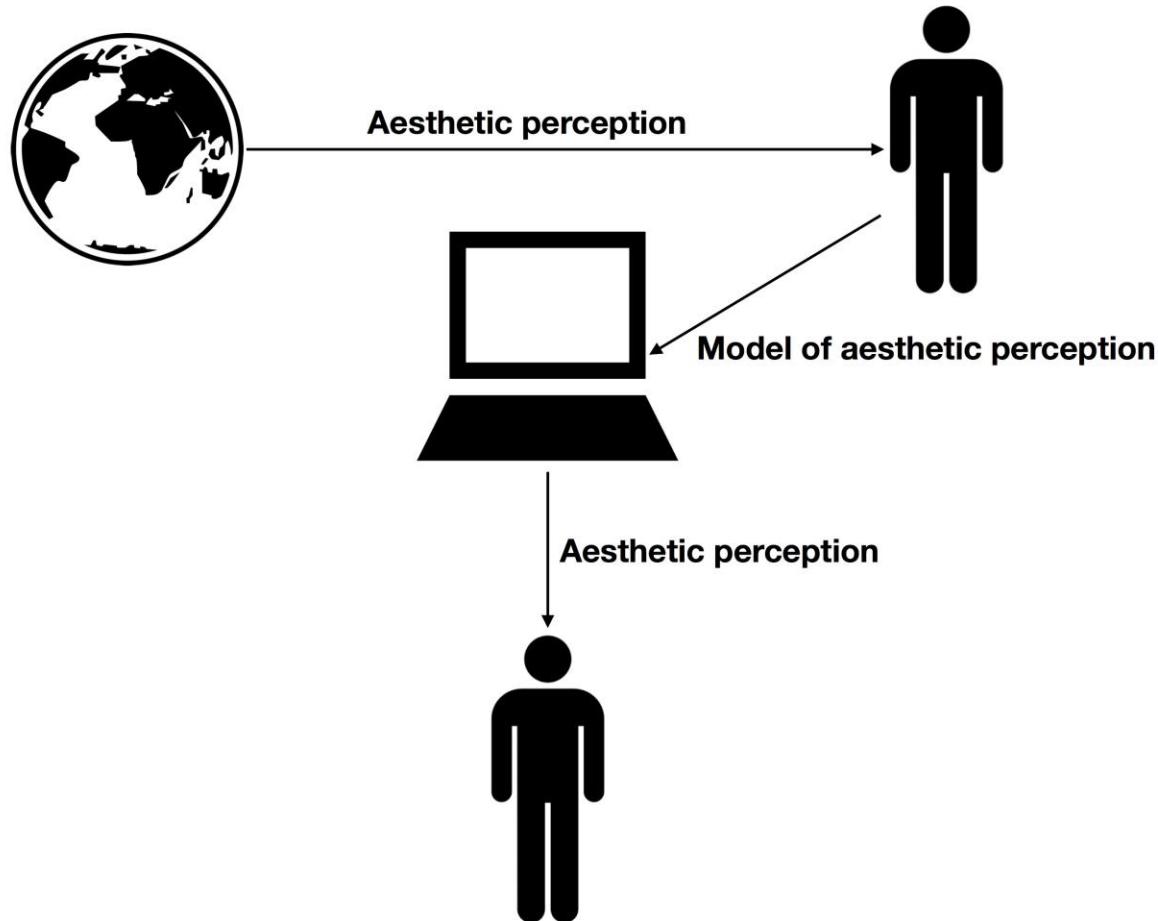


Marfia, Gustavo, and Giovanni Matteucci. "Some remarks on aesthetics and computer science." *Studi di estetica* 12 (2018).



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Verso una dimensione percettiva

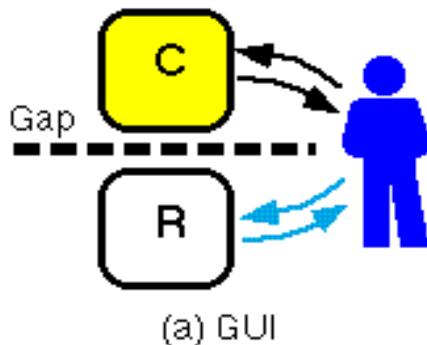


Marfia, Gustavo, and Giovanni Matteucci. "Some remarks on aesthetics and computer science." *Studi di estetica* 12 (2018).

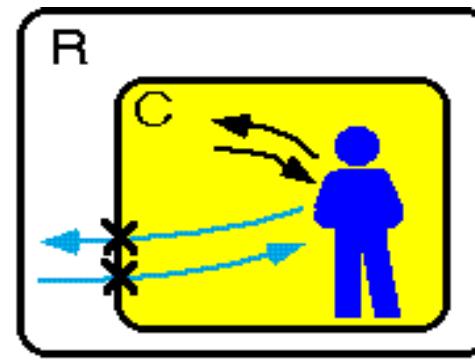


ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

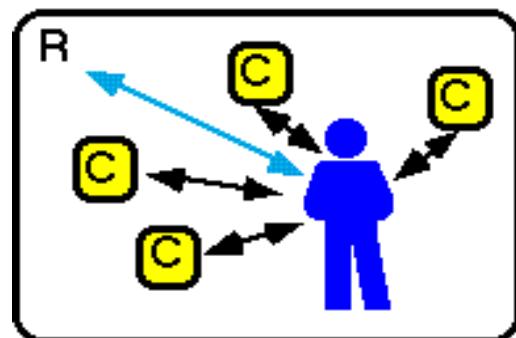
# Cambiano le modalita' di interazione



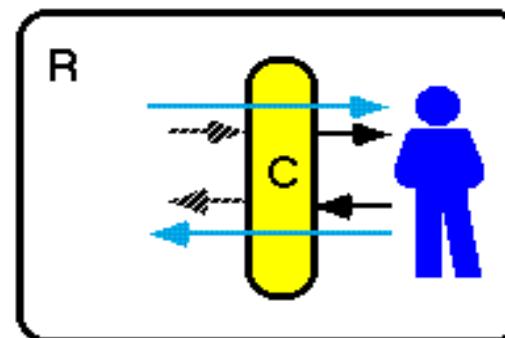
(a) GUI



(b) Virtual Reality



(c) Ubiquitous Computers



(d) Augmented Interaction

Rekimoto, J. and Nagao, K. 1995. *The world through the computer: computer augmented interaction with real world environments*. In *Proceedings of the 8th Annual ACM Symposium on User interface and Software Technology. UIST '95*. ACM, New York, NY, 29-36.



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# WebXR



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Rischi

# Big data: solo dati o vera conoscenza?

La prima è che le tecniche computazionali e il valanga di numeri diventati fini a se stessi, scollegato da ciò che è importante.

Cioè, le tecniche e i numeri diventano feticizzati, indossano un piedistallo, apprezzati per quello che sono piuttosto che per cosa fanno. Questo va all'importante distinzione tra dati e conoscenze. Raccogliamo dati per amore dei dati? Perché sono lì. Perchè possiamo.

Nate Silver, lo statistico dei big data che con successo ha previsto le elezioni americane del novembre 2012 risultati in ciascuno dei 50 stati ha detto di recente, "la maggior parte" dei dati è solo rumore, come lo è la maggior parte dell'universo riempito di spazio vuoto.

I big data produrranno sempre più rumore? Poiche' descrivibili matematicamente, e poiché questo è il segno distintivo della scienza ("la matematica è il linguaggio della natura" diceva Galileo), i big data verranno propagandati come conoscenza.

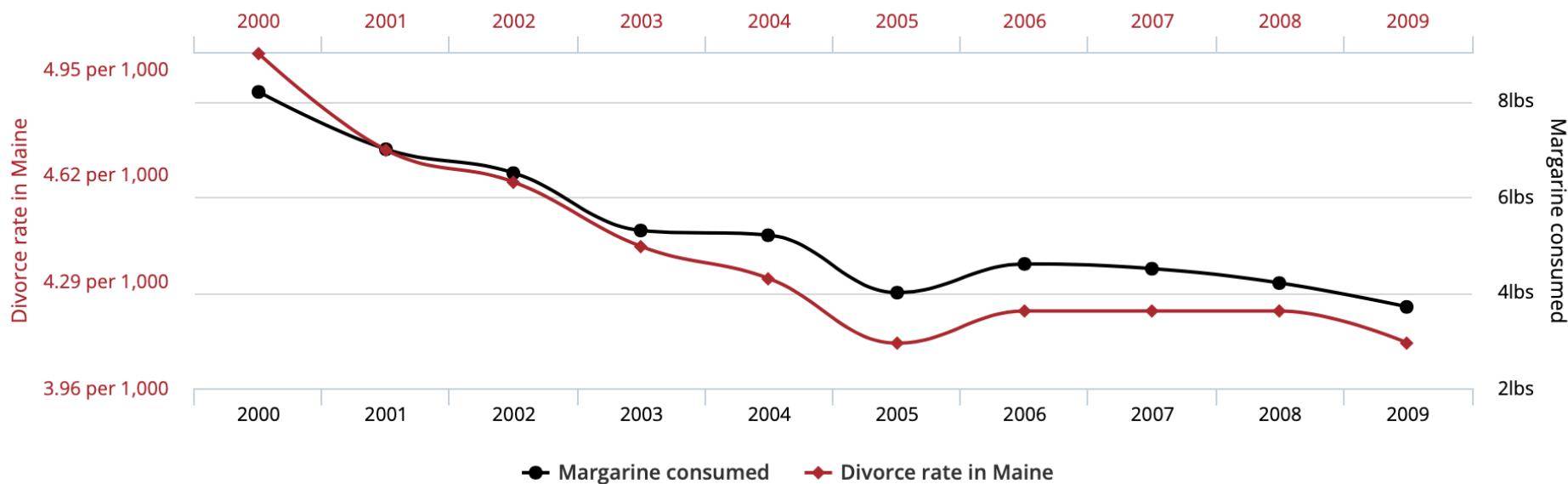


# Rischi: solo dati o vera conoscenza?



## Divorce rate in Maine correlates with Per capita consumption of margarine

Correlation: 99.26% ( $r=0.992558$ )



tylervigen.com

Data sources: National Vital Statistics Reports and U.S. Department of Agriculture



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

<http://tylervigen.com/>

## Rischi: solo quantitativo

Affinché possano essere inclusi all'interno di big data, le informazioni deve essere convertito in numeri, in dati.

Francis Galton, l'inventore vittoriano della regressione, disse: "Quando puoi, conta".

Tale affermazione è particolarmente importante nell'attuale regime dei big data. Se qualcosa non può essere contato, quantificato, non ha senso includerlo. Nello specifico, ciò che spesso si perde è il contesto, che non può essere messo in un'equazione.



# Rischi: i dati possono sostituire il metodo scientifico?

Terzo, nella visualizzazione dei big data, i numeri sono i storia, spogliata del bisogno di qualsiasi interpretazione, spogliata della necessità di attribuire anche il nesso di causalità.

Chris Anderson, nel suo editoriale del 2008, dice, "con abbastanza" dati, i numeri parlano da soli". Cioè, i dati da soli determinano la conoscenza senza interpretazione o causalità.

Anderson scrive, "Correlazione» sostituisce la causalità, e la scienza può avanzare anche senza modelli coerenti, teorie unificate, o davvero alcuna spiegazione meccanicistica» (2008).

La correlazione non risponde alla grande domanda che le scienze sociali più vuole rispondere: perché succedono le cose? Supporre che coefficienti di correlazione siano sostitutivi della causalità è illudere noi stessi e gli altri.

Barnes, Trevor J. "Big data, little history." *Dialogues in Human Geography* 3, no. 3 (2013): 297-302.

Anderson C (2008) *The end of theory: the data deluge makes the scientific method obsolete*. *Wired Magazine* 23 June. Available at: [http://www.wired.com/science/discoveries/magazine/16-07/pb\\_theory/](http://www.wired.com/science/discoveries/magazine/16-07/pb_theory/)



## Rischi: mancanza di originalità

L'ultimo punto è che i big data come progetto sono intrinsecamente conservatori.

Utilizzando i numeri man mano che vengono forniti, i big data sono bloccati con ciò che è, piuttosto che con ciò che dovrebbe essere.

Ecco perché Netflix ha ripreso House of Cards. Non sapremo mai se qualcosa di completamente nuovo sarebbe potuto essere meglio, non solo in termini qualitativi, ma anche per il fatturato dell'azienda.

Barnes, Trevor J. "Big data, little history." *Dialogues in Human Geography* 3, no. 3 (2013): 297-302.



# Rischi: cosa perdiamo nei luoghi virtuali?

Un database non è un luogo e gli esseri umani sono creature per le quali l'apprendimento, inteso come un mix di procedure di assemblaggio, lettura, analisi e condivisione di informazioni, è un atto complesso e incarnato, un'esperienza multistrato e situata, che non è riducibile al fatto che i dati siano semplicemente disponibili in tasca o nella libreria.

Le biblioteche sono sempre stati spazi sociosensoriali in cui si esercita l'intera gamma delle capacità cognitive dell'umanità, sia quelle che si occupano di ciò che è facilmente osservabile e quantificabile, sia quelle che invece sono alle prese con percezioni sfuggenti, ambigue, effimere.

Schnapp, Jeffrey. "The Permanent Library of the Now." *KNOW: A Journal on the Formation of Knowledge* 2, no. 2 (2018): 303-320.



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

# Discussione

# Discussione, con alcune domande

- Quali sono le tecnologie e i metodi alla base di Big Data?
- La correlazione implica la causalità?
- Quali sono gli elementi che hanno determinato il successo di Internet?
- Quali aspetti, secondo voi, non abbiamo toccato nel seminario di oggi?
- Quali campi del sapere saranno più influenzati dall'uso di tali tecnologie?





ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA  
CAMPUS DI RIMINI

## Gustavo Marfia

Dipartimento di Scienze per la Qualità della Vita

[gustavo.marfia@unibo.it](mailto:gustavo.marfia@unibo.it)

[www.unibo.it](http://www.unibo.it)