

Modena – 29 settembre 2018

Datacy: valutare l'impatto
delle tecnologie sulla società

Carlo Batini
Università di Milano Bicocca
batini@disco.unimib.it

Literacy

- *Literacy* is the ability to use written language actively and passively; one *definition of literacy* is the ability to "read, write, spell, listen, and speak". ...
- A basic *literacy* standard in many places is the ability to read the newspaper.



Cosa **e'** la datacy?

Etymology of data from Wiktionary

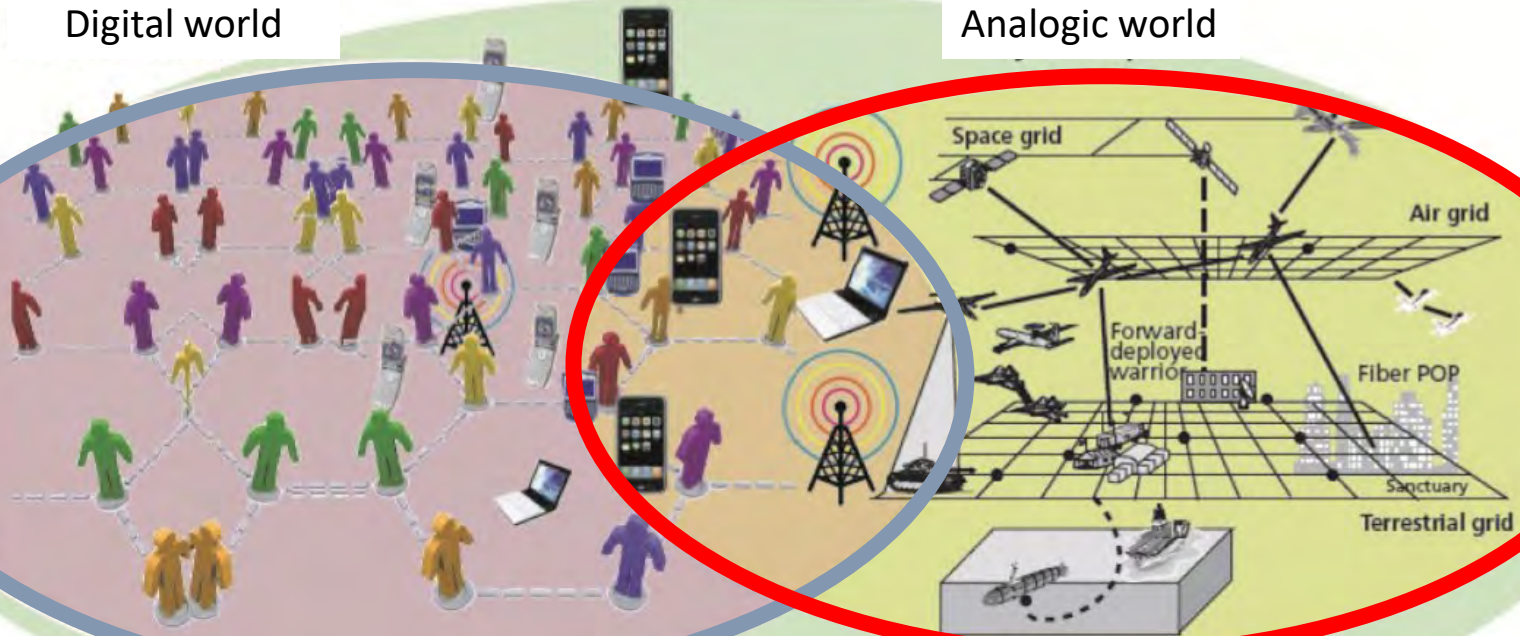
From Latin *data*, nominative plural of *datum* (“that is given”)

Dunque: al passato...

The information ecosystem

Digital world

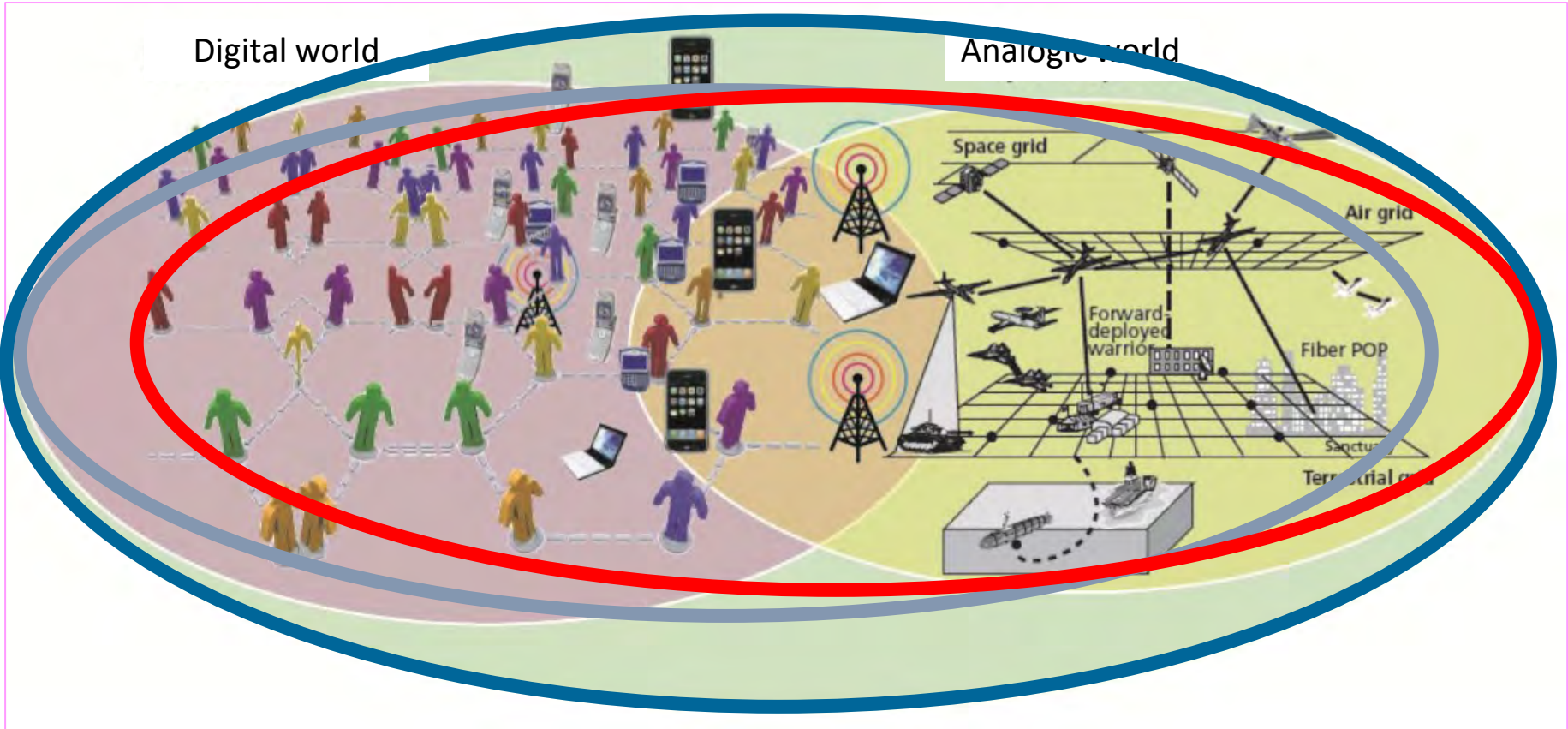
Analogic world



Floridi Infosphere

Digital world

Analogic world



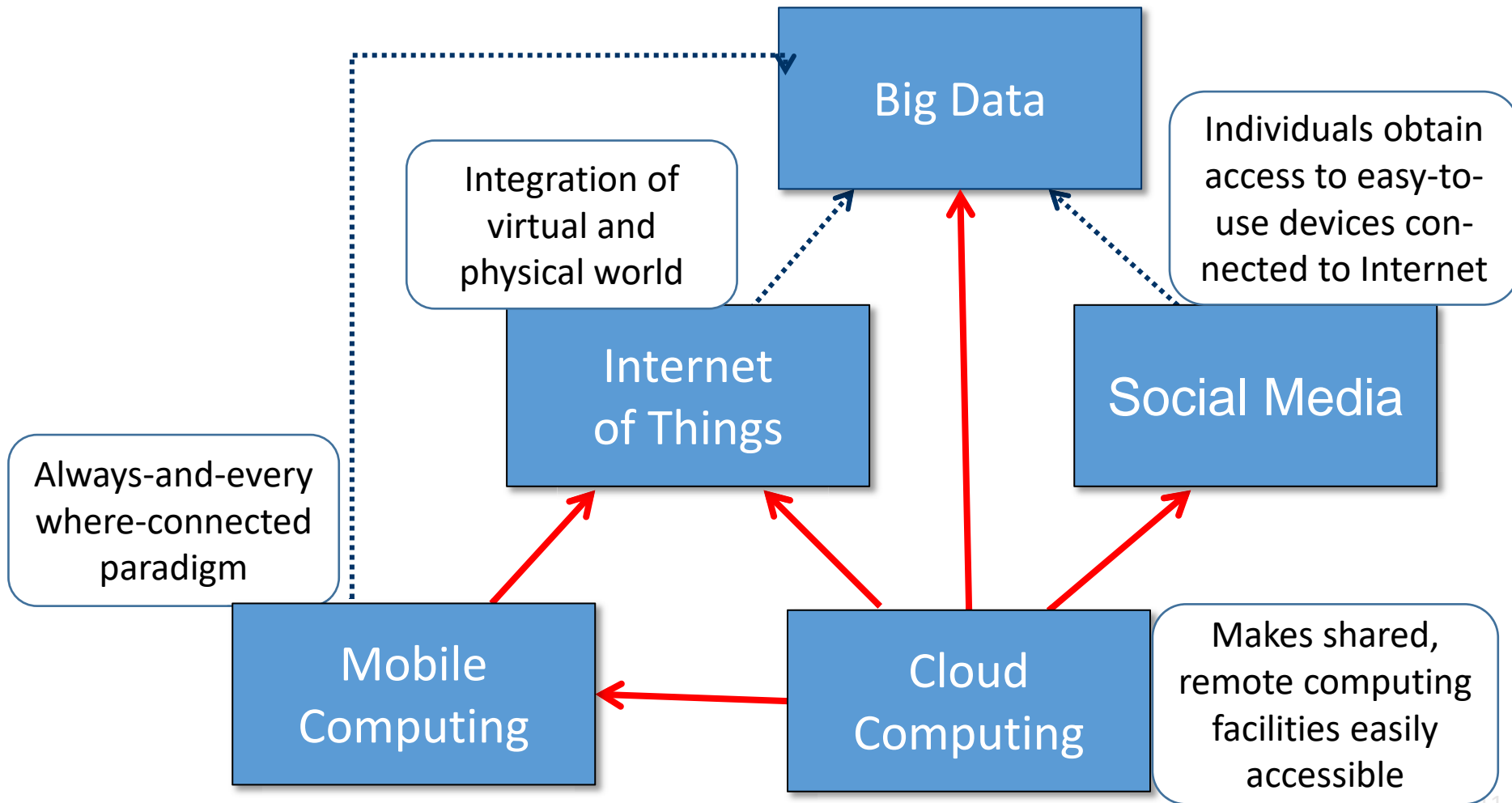
Ready Player One, di Steven Spielberg



Le technologie ICT

The Big Five (from Grefen 2016)

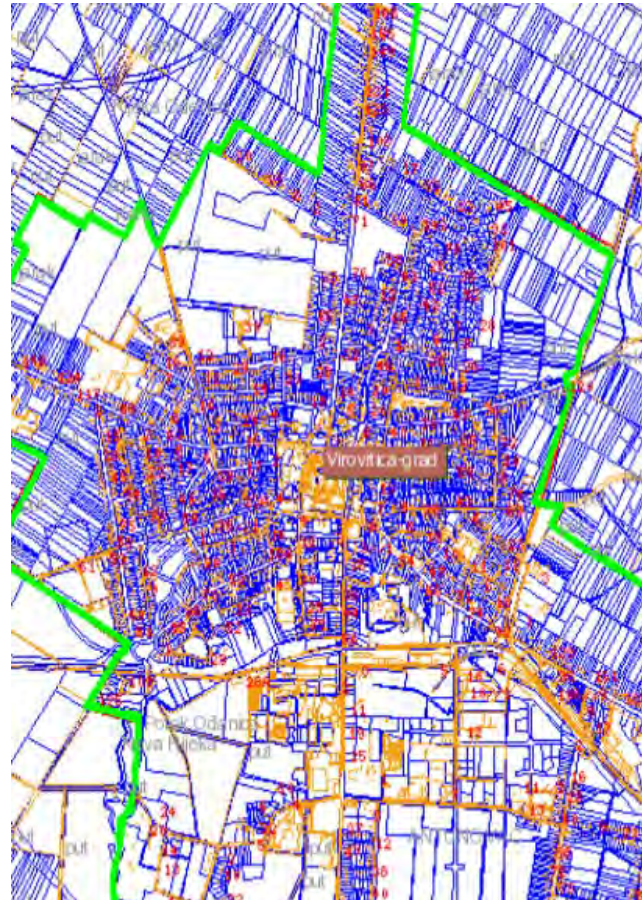
← enables
←..... feeds data



Dati catastali in India



Dati catastali in India



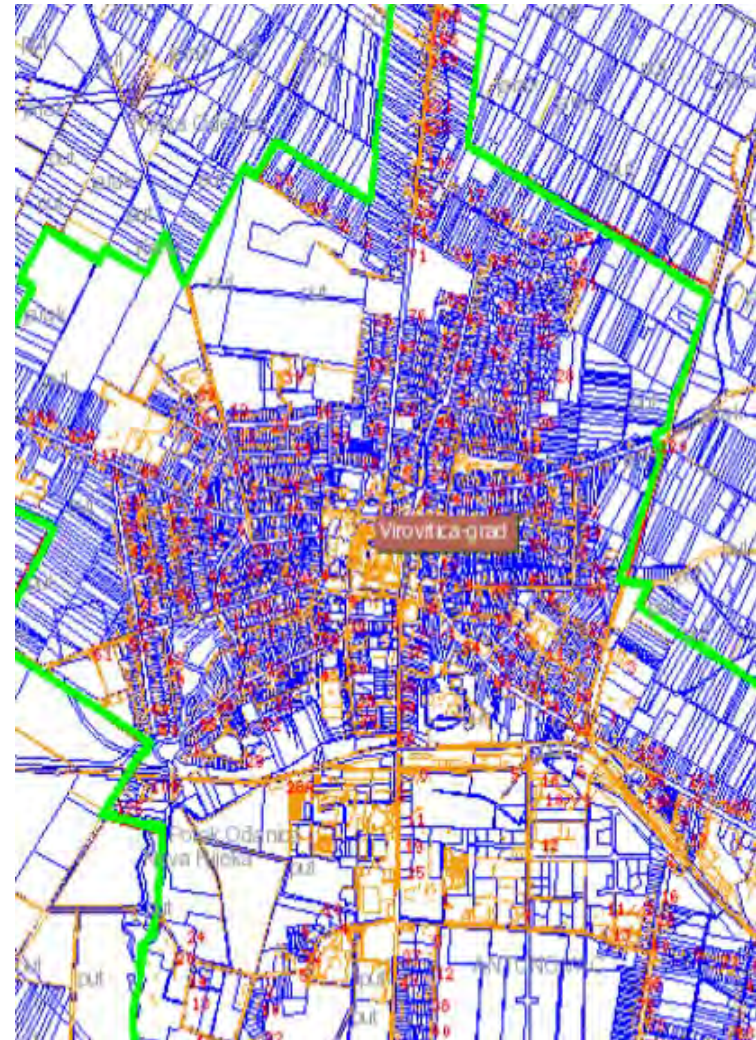
Scopi della digitalizzazione

Dare più garanzie ai cittadini verso:

- Le burocrazie
 - Gli ufficiali corrotti
- attraverso la trasparenza e la tracciabilità

Risultato effettivo:

il contrario di quanto si era sperato



Verba volant, scripta manent



 *Il commento*

Le parole
che restano
ai tempi
di WhatsApp

di **Marco Imarisio**

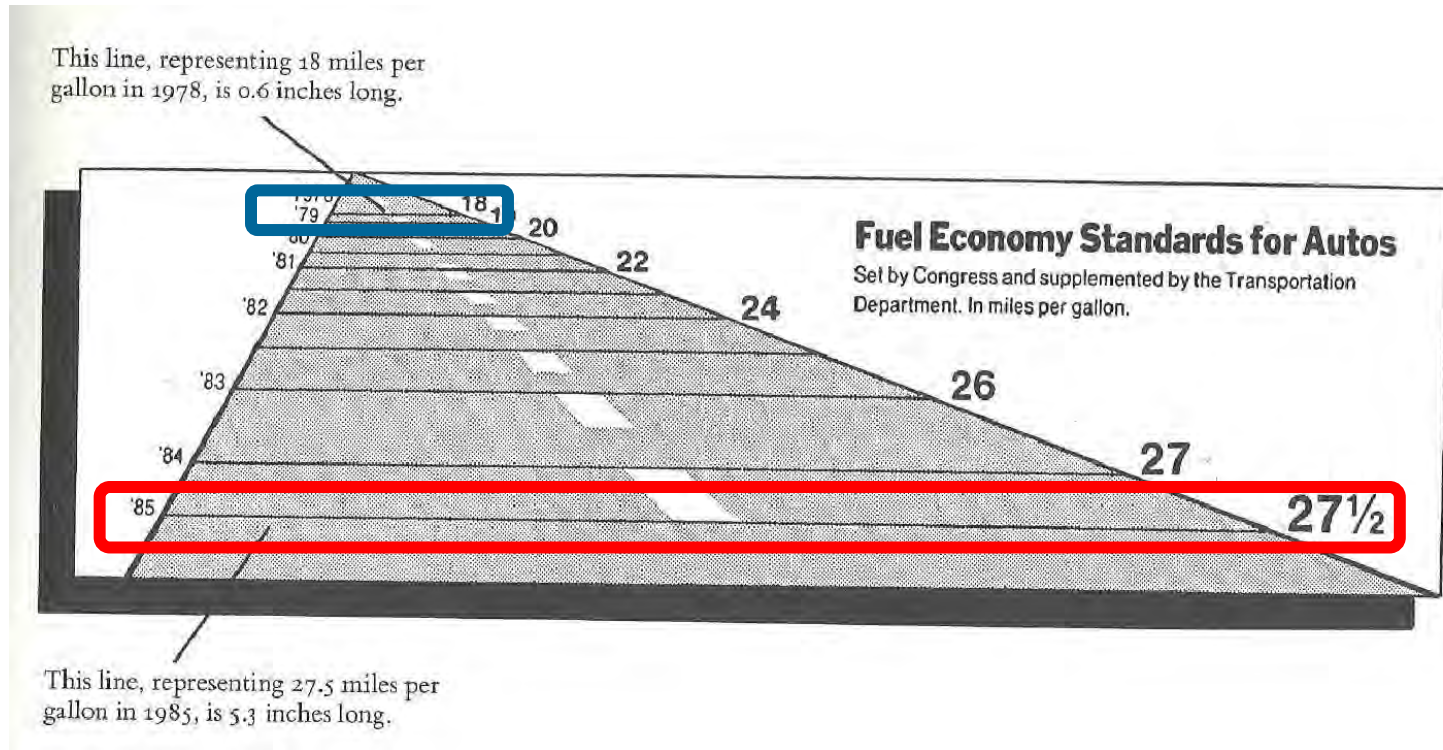
Com'era bello quando le parole si limitavano a volare. Ce lo insegnavano fin da bambini, che verba volant mentre scripta manent. Era una di quelle regole di vita che a nessuno veniva in mente di mettere in dubbio. I tempi sono cambiati, non solo in meglio, non solo in politica. Quel che colpisce dell'affaire Casalino è anche il modo. Non importa se ci troviamo davanti a una

residente del Consiglio Giovanni Conte

Alc...

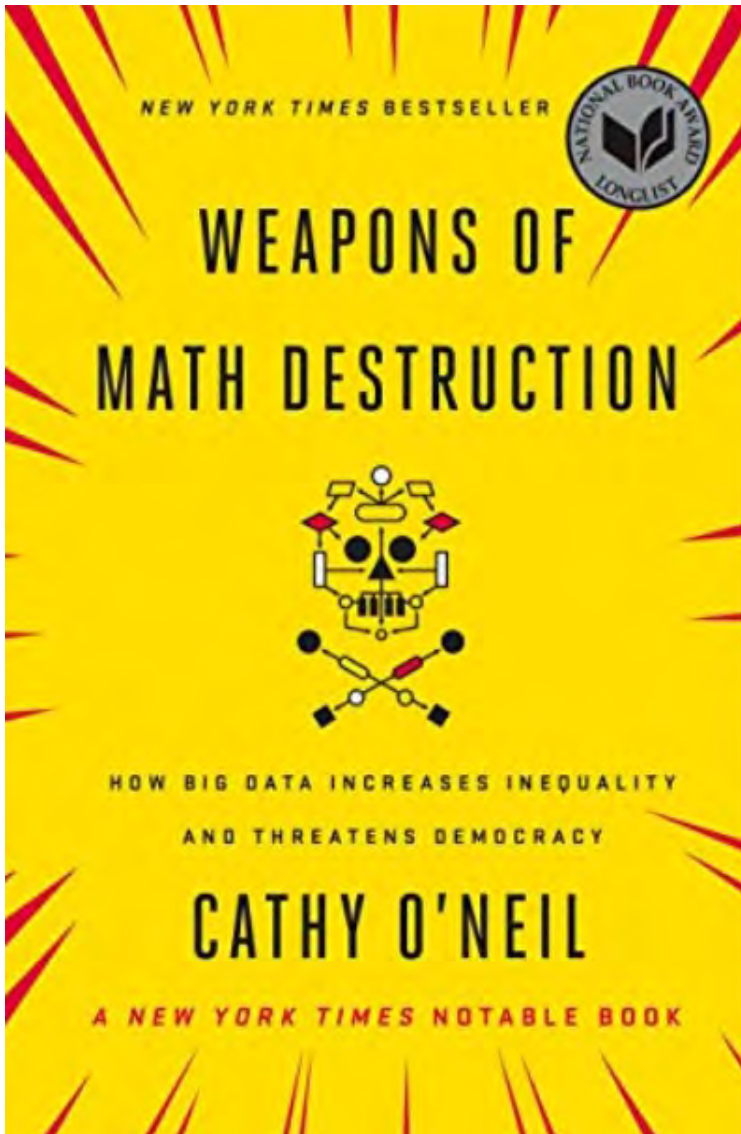
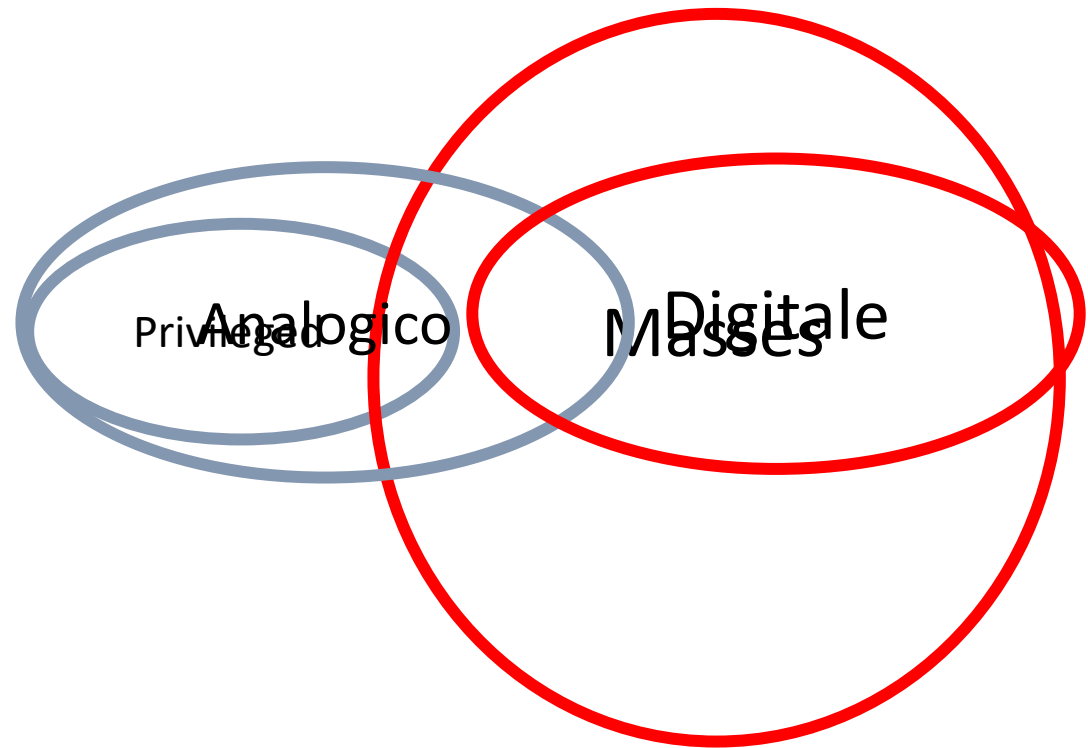
How many lies in visualizations...

Year	Miles per gallon
1978	18
1979	19
1980	20
1981	22
1982	24
1983	26
1984	27
1985	27.5



Lie factor = ratio of numerical values in the real world / ratio in visualization = 15

The privileged are processed more by people, the masses by machines

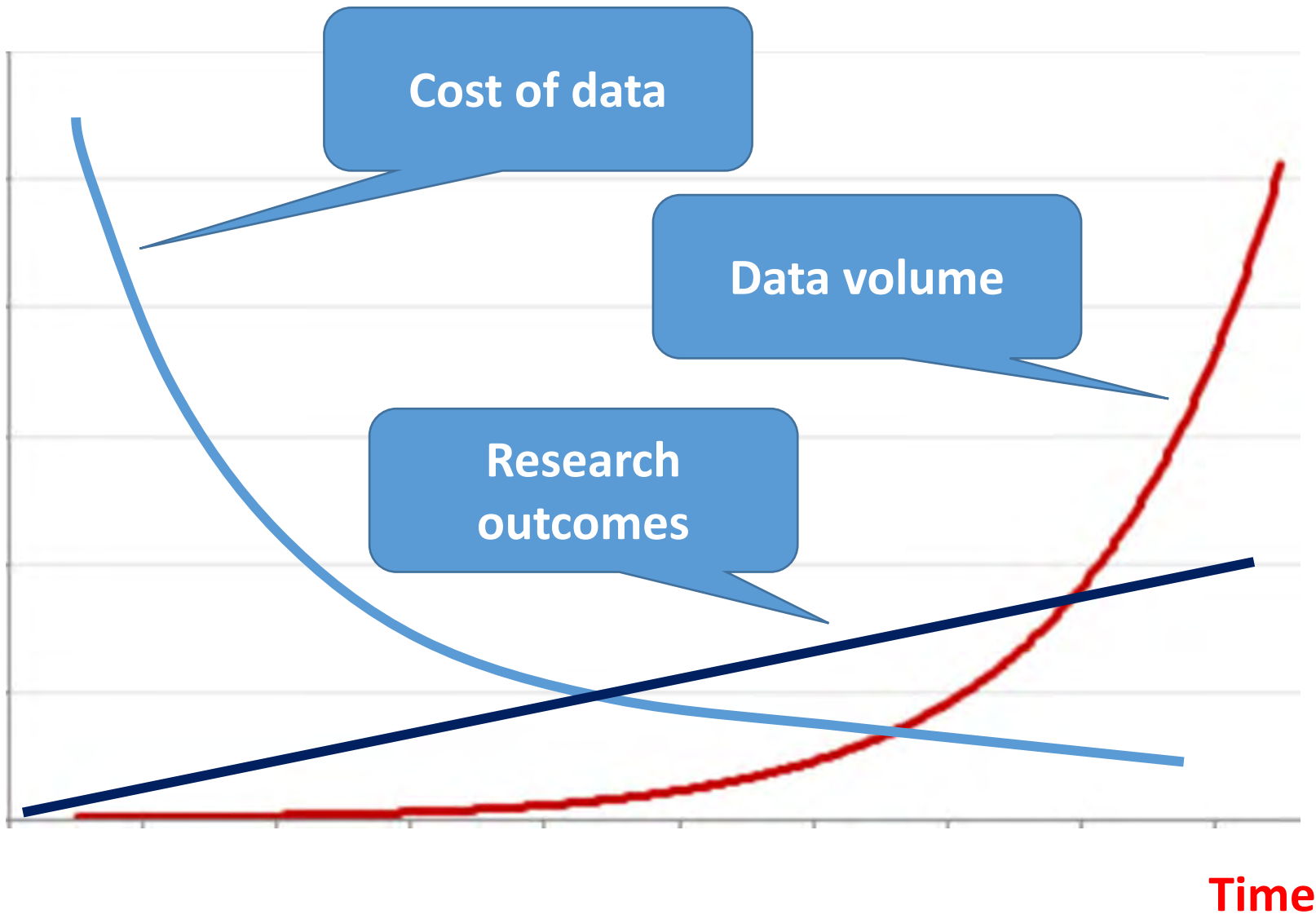


EB = 10^{18} byte Four Domains of Big Data in 2025

Data Phase	Astronomy	Twitter	YouTube	Genomics
Acquisition	25 zetta-bytes/year	0.5–15 billion tweets/year	500–900 million hours/year	1 zetta-bases/year
Storage	1 EB/year	1–17 PB/year	1–2 EB/year	2–40 EB/year
Analysis	In situ data reduction	Topic and sentiment mining	Limited requirements	Heterogeneous data and analysis
	Real-time processing	Metadata analysis		Variant calling, ~2 trillion central processing unit (CPU) hours
	Massive volumes			All-pairs genome alignments, ~10,000 trillion CPU hours
Distribution	Dedicated lines from antennae to server (600 TB/s)	Small units of distribution	Major component of modern user's bandwidth (10 MB/s)	Many small (10 MB/s) and fewer massive (10 TB/s) data movement

Uneven trends of **data growth** and **research & culture on data**

Amount



Beni, dati, servizi e lavori

Goods, services, data



← **Data** ? →



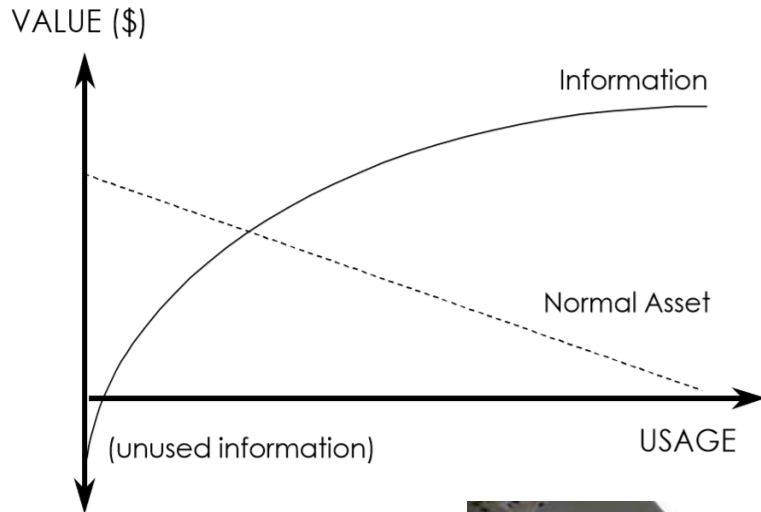
Data are not so visible as goods

```
Source: query [8.074e+07 x 5]
Database: spark connection master=local[8] app=sparklyr local=TRUE

  user_id  item_id rating timestamp      category
   <chr>   <chr>  <dbl>    <int>    <chr>
1 A1EE2E3N7PW666 B000GFDAUG      5 1202256000 Amazon Instant Video
2 AGZ8SM1BGK3CK B000GFDAUG      5 1198195200 Amazon Instant Video
3 A2VHZ21245KBT7 B000GIOPK2      4 1215388800 Amazon Instant Video
4 ACX8YW2D5EGP6 B000GIOPK2      4 1185840000 Amazon Instant Video
5 A9RNM09MUSMTJ B000GIOPK2      2 1281052800 Amazon Instant Video
6 A3STFVPM8NHJ7B B000GIOPK2      5 1203897600 Amazon Instant Video
7 A2582KMXLK2P06 B000GIOPK2      5 1205884800 Amazon Instant Video
8 A1TZCLCW9QGBH B000GIOPK2      4 1209427200 Amazon Instant Video
9 A2E2I6B878CRMA B000GIOPK2      5 1378684800 Amazon Instant Video
10 AD5MZA8SOVMPJ B000GIOPK2      5 1218240000 Amazon Instant Video
# ... with 8.074e+07 more rows
```



Data are not perishable as services

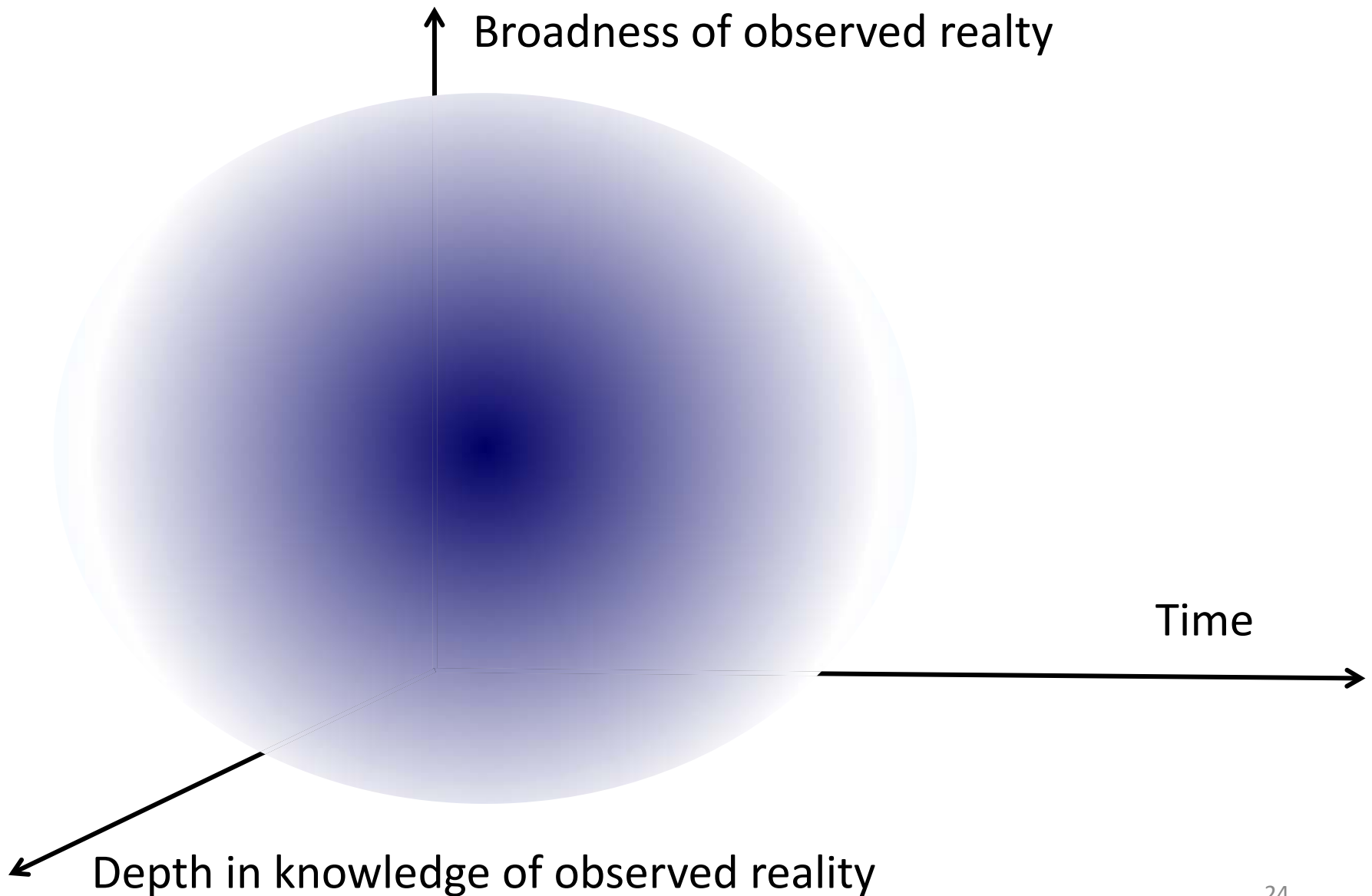


```
Source: query [8.074e+07 x 5]
Database: spark connection master=local[8] app=sparklyr local=TRUE

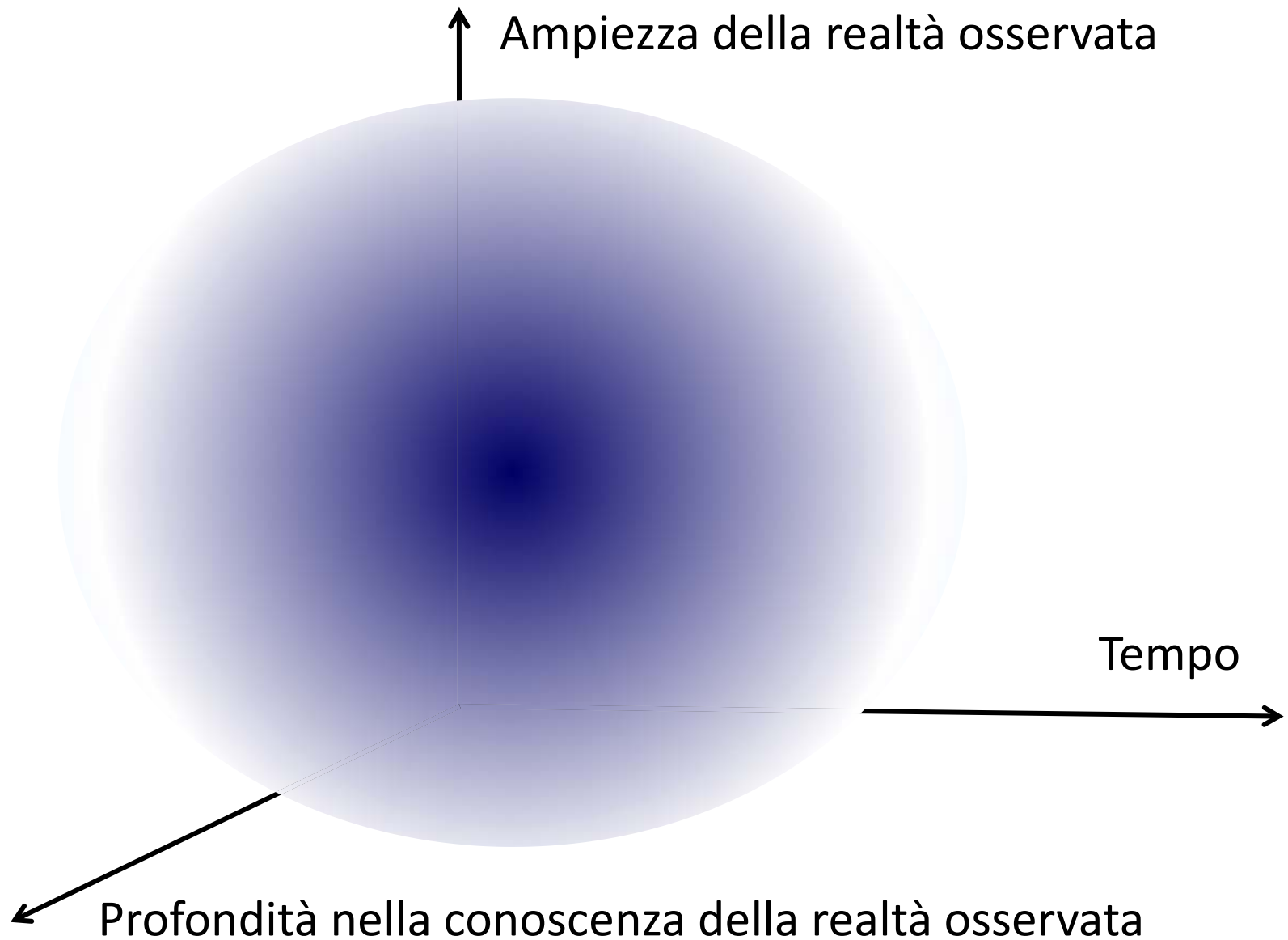
user_id  item_id  rating  timestamp  category
<chr>   <chr>   <dbl>   <int>     <chr>
1  A1EE2E3N7PW666  B000GFDAUG  5 1202256000 Amazon Instant Video
2  AGZ8SM1BGK3CK  B000GFDAUG  5 1198195200 Amazon Instant Video
3  A2VHZ21245KBT7  B000GIOPK2  4 1215388800 Amazon Instant Video
4  ACX8YW2D5EGP6  B000GIOPK2  4 1185840000 Amazon Instant Video
5  A9RNM09MUSMTJ  B000GIOPK2  2 1281052800 Amazon Instant Video
6  A35TFVPM8NHJ7B  B000GIOPK2  5 1203897600 Amazon Instant Video
7  A2582KMXLK2P06  B000GIOPK2  5 1205884800 Amazon Instant Video
8  A1TZCLCW9QGBH  B000GIOPK2  4 1209427200 Amazon Instant Video
9  A2E2IGB878CRMA  B000GIOPK2  5 1378684800 Amazon Instant Video
10 AD5MZA85OVMPJ  B000GIOPK2  5 1218240000 Amazon Instant Video
# ... with 8.074e+07 more rows
# ... with 8.074e+07 more rows
# ... with 8.074e+07 more rows
# ... with 8.074e+07 more rows
# ... with 8.074e+07 more rows
# ... with 8.074e+07 more rows
# ... with 8.074e+07 more rows
# ... with 8.074e+07 more rows
# ... with 8.074e+07 more rows
# ... with 8.074e+07 more rows
```



A blurred and unequal reality



A blurred and unequal reality

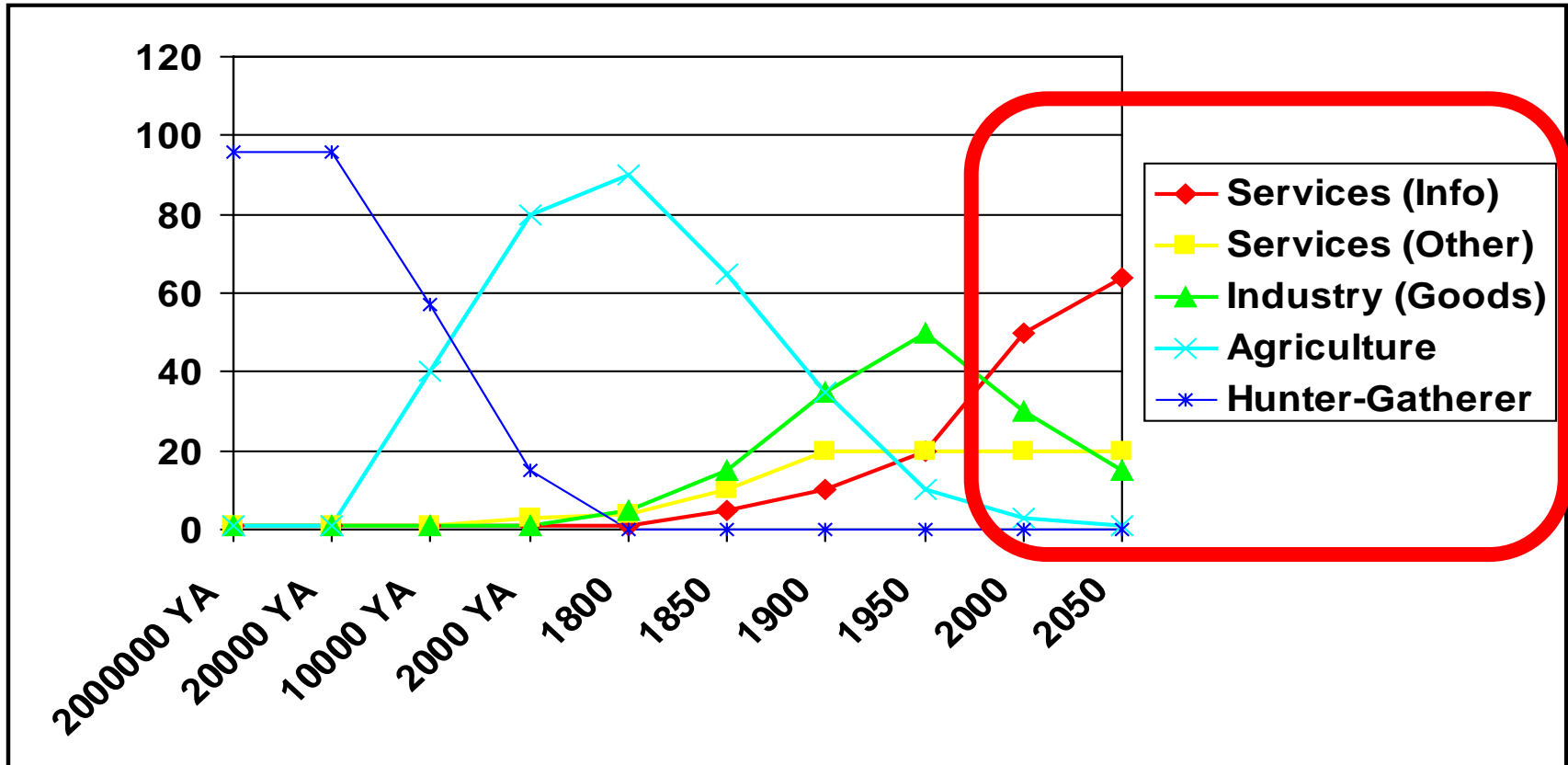


Blurred and unequal Modena vs Nairobi Kibera



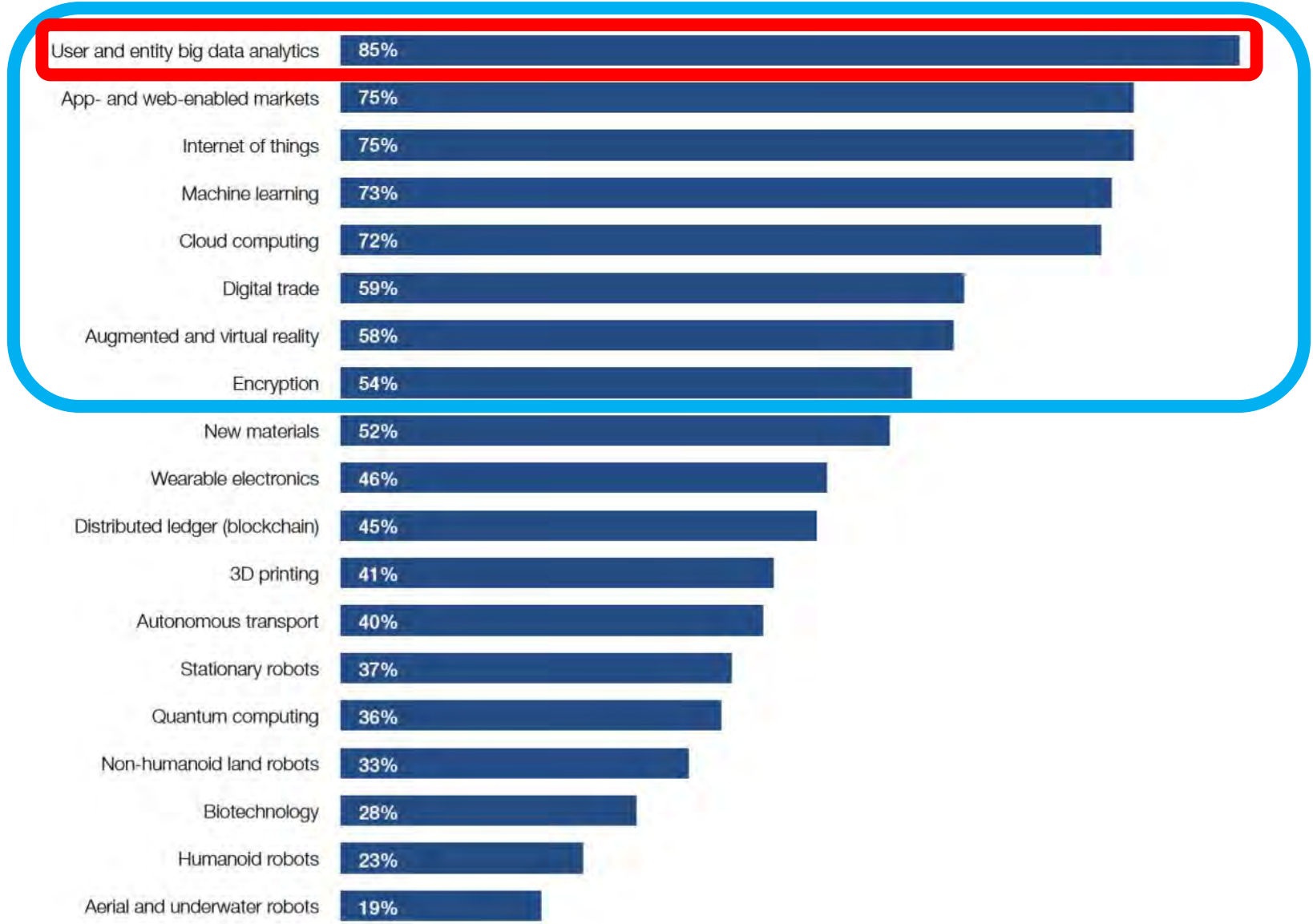
I lavori nel tempo della datacy

Cambiano i lavori...



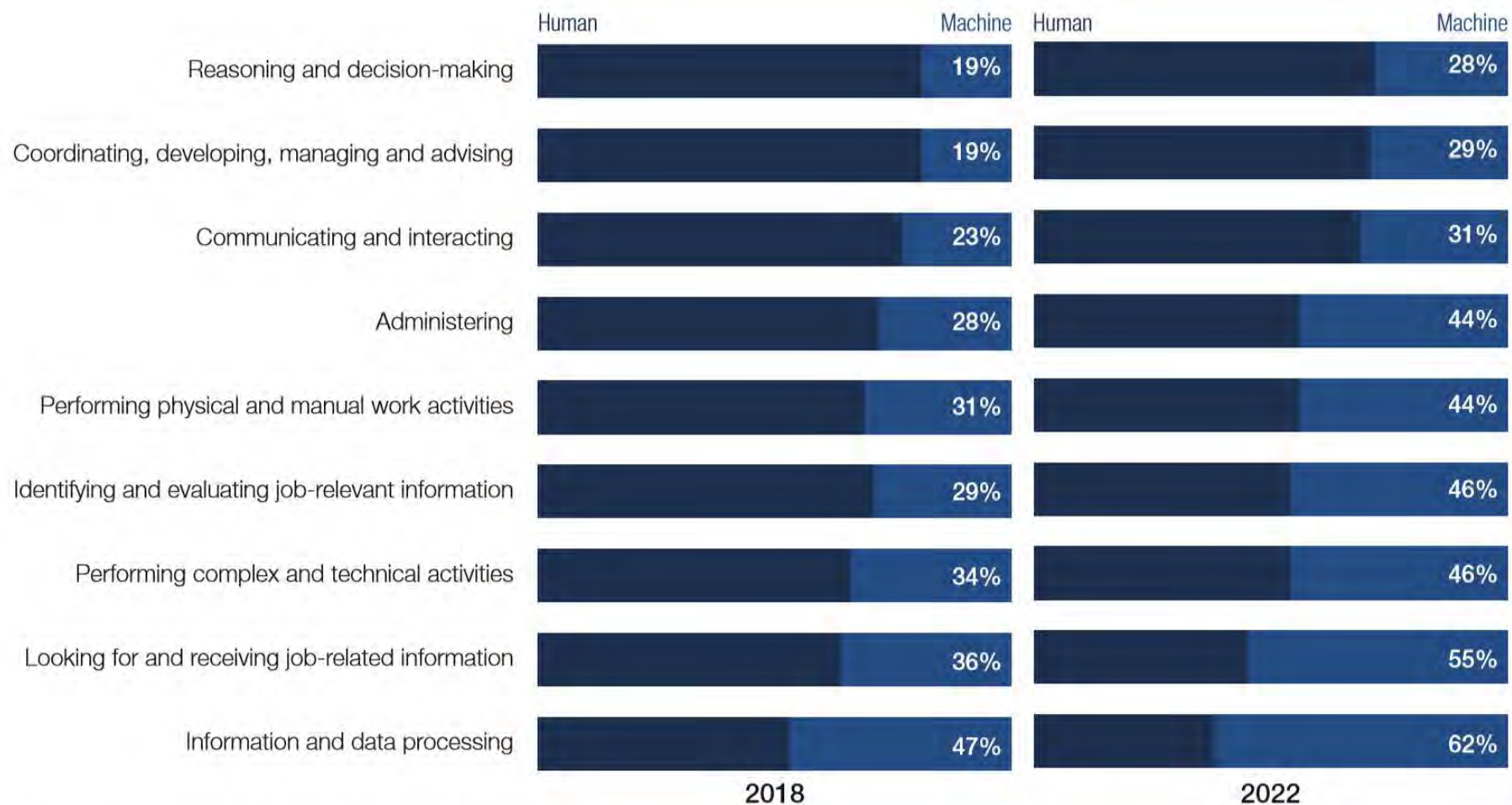
Estimations based on Porat, M. (1977) Info Economy: Definitions and Measurement

Technologies by proportion of companies likely to adopt them by 2022 (projected)



Source: Future of Jobs Survey 2018, World Economic Forum.

Ratio of human-machine working hours, 2018 vs. 2022 (projected)



Source: Future of Jobs Survey 2018, World Economic Forum.

Dall' Economist, Nov. 10th, 2016

Le tecnologie costano molto meno delle persone...

- The agencies not only do more, they also spend less.
- To **deploy agents on a tail** costs \$ 175.000 a month because it takes a lot of manpower.
- To **put a GPS receiver in a someone's car** takes \$ 150 a month.
- To **tag a target's mobile phone**, with the help of a phone company, costs only \$ 30 a month.

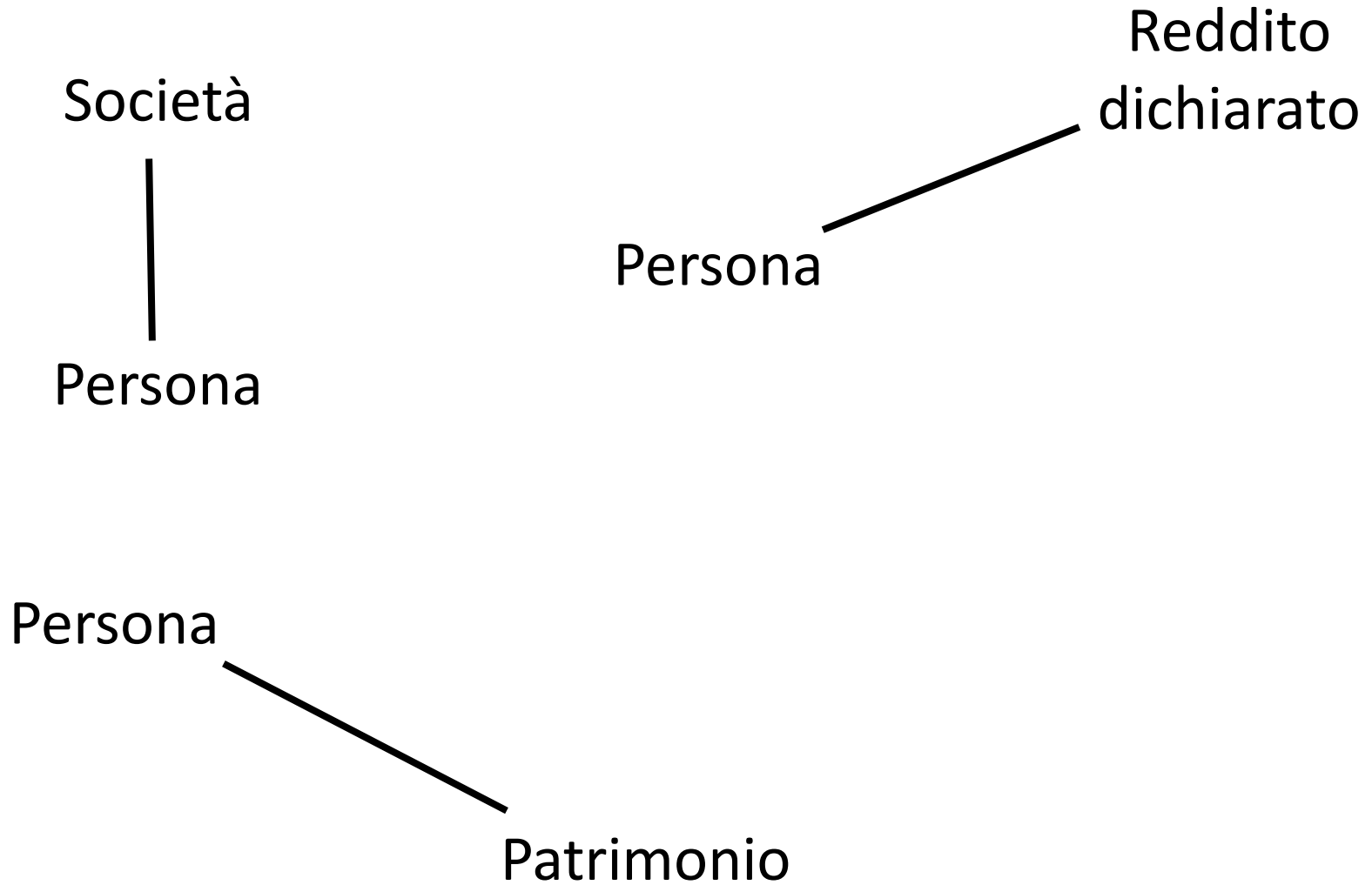
Il presupposto per capire:
Dare ai dati un significato

Il presupposto per convivere:
Integrare i dati aumenta il loro valore

Esempio 1

Scoprire gli evasori

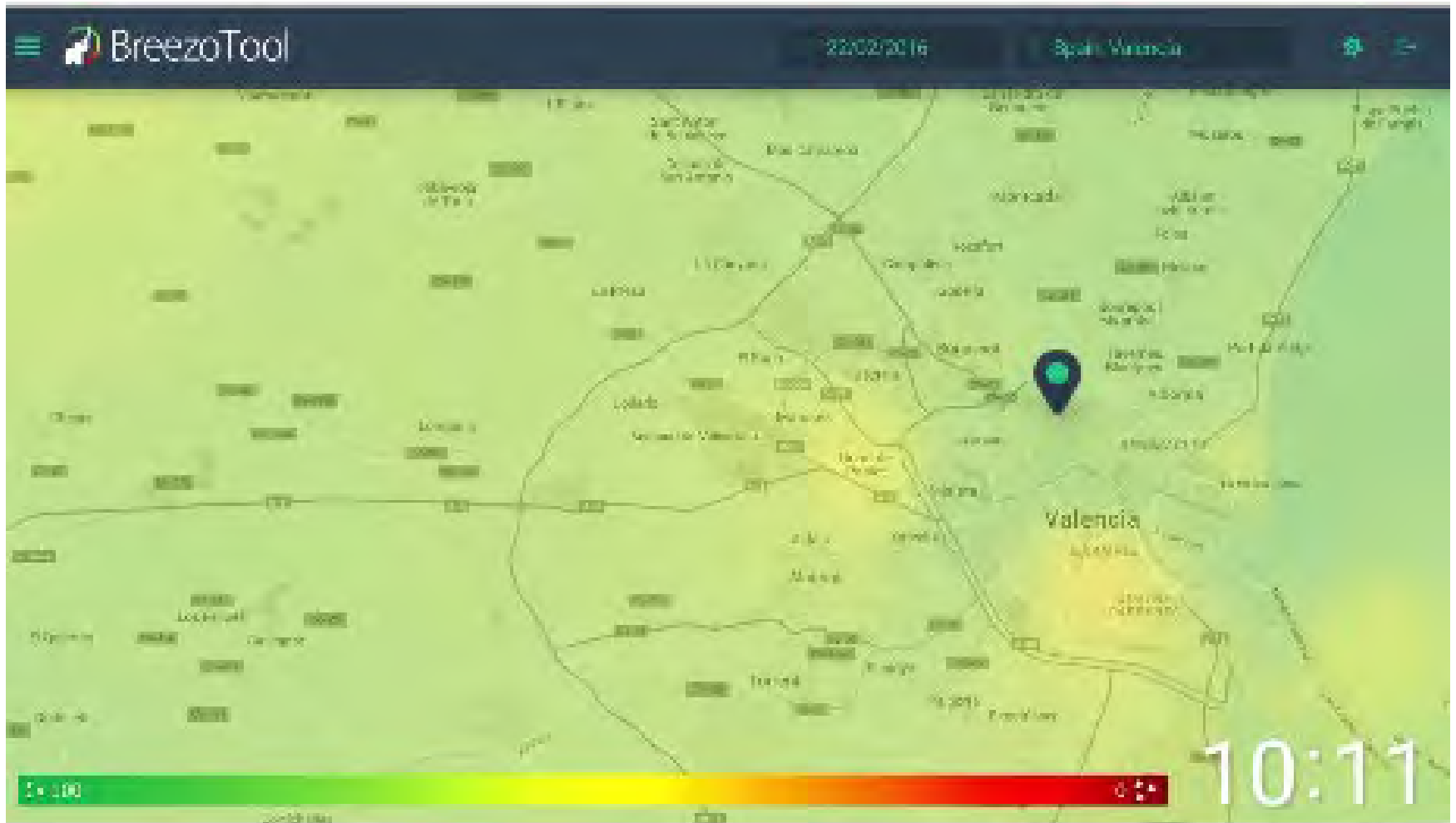
Integrare porta valore



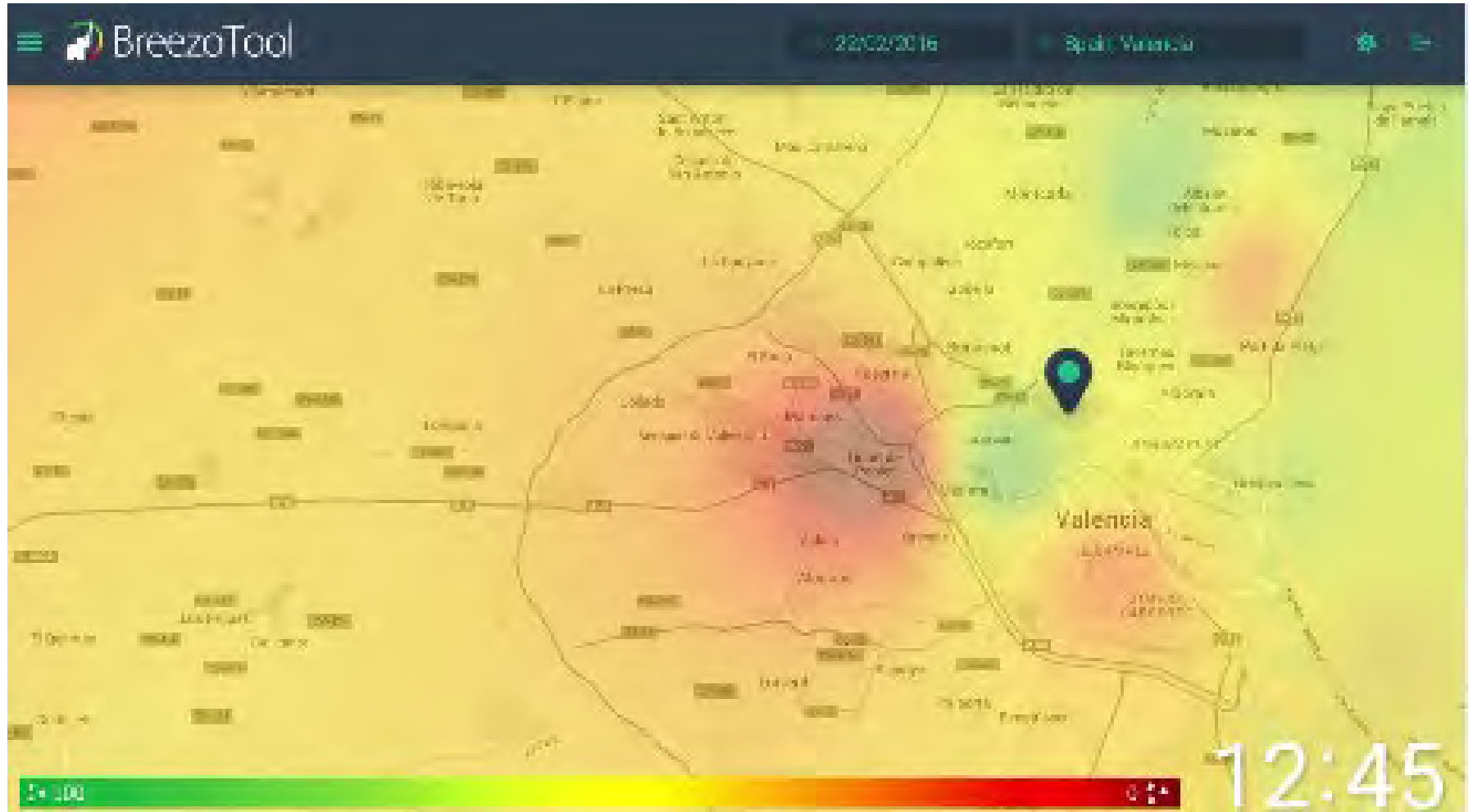
Esempio 2

Prevedere i livelli di inquinamento

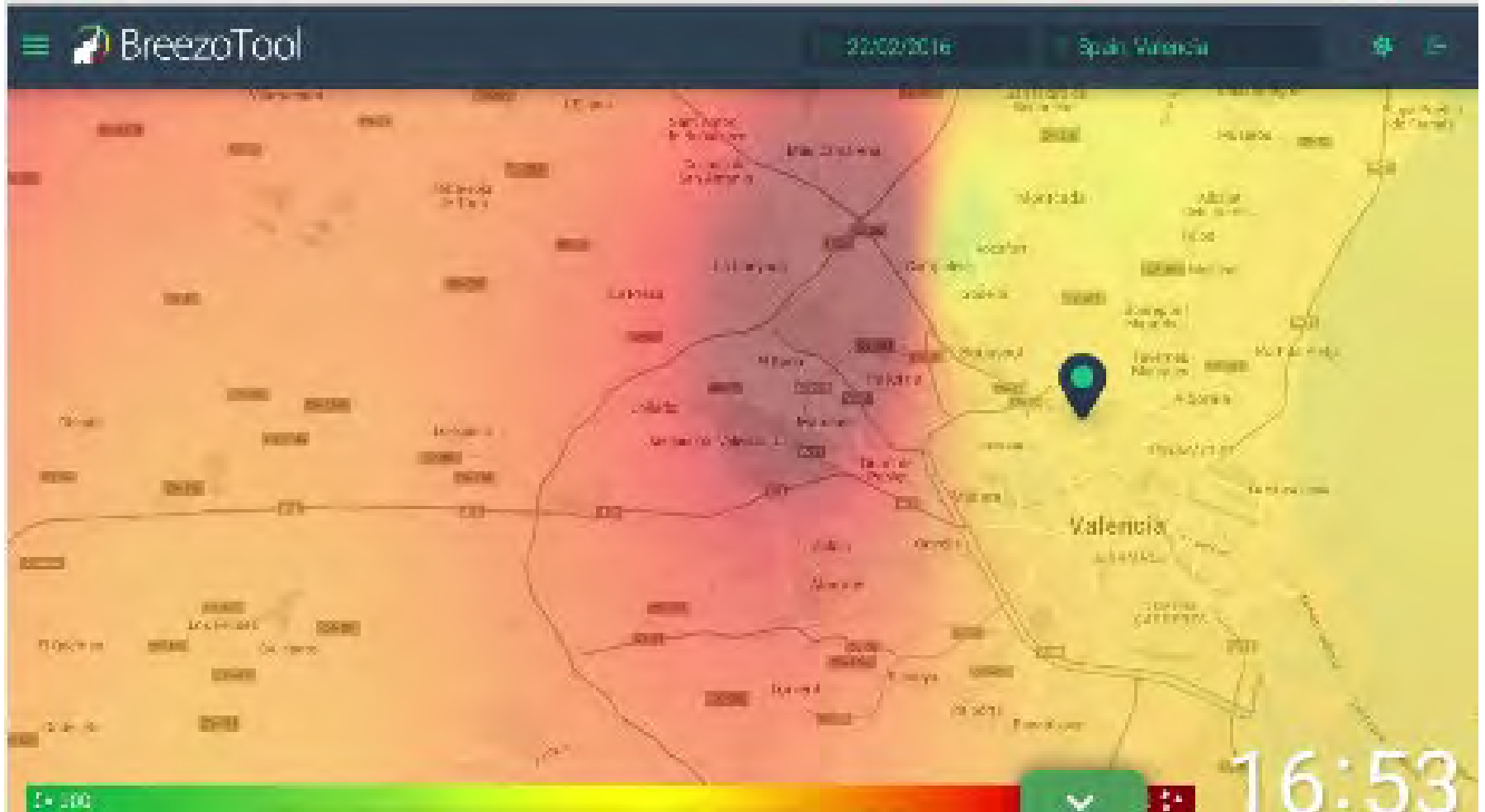
10:11



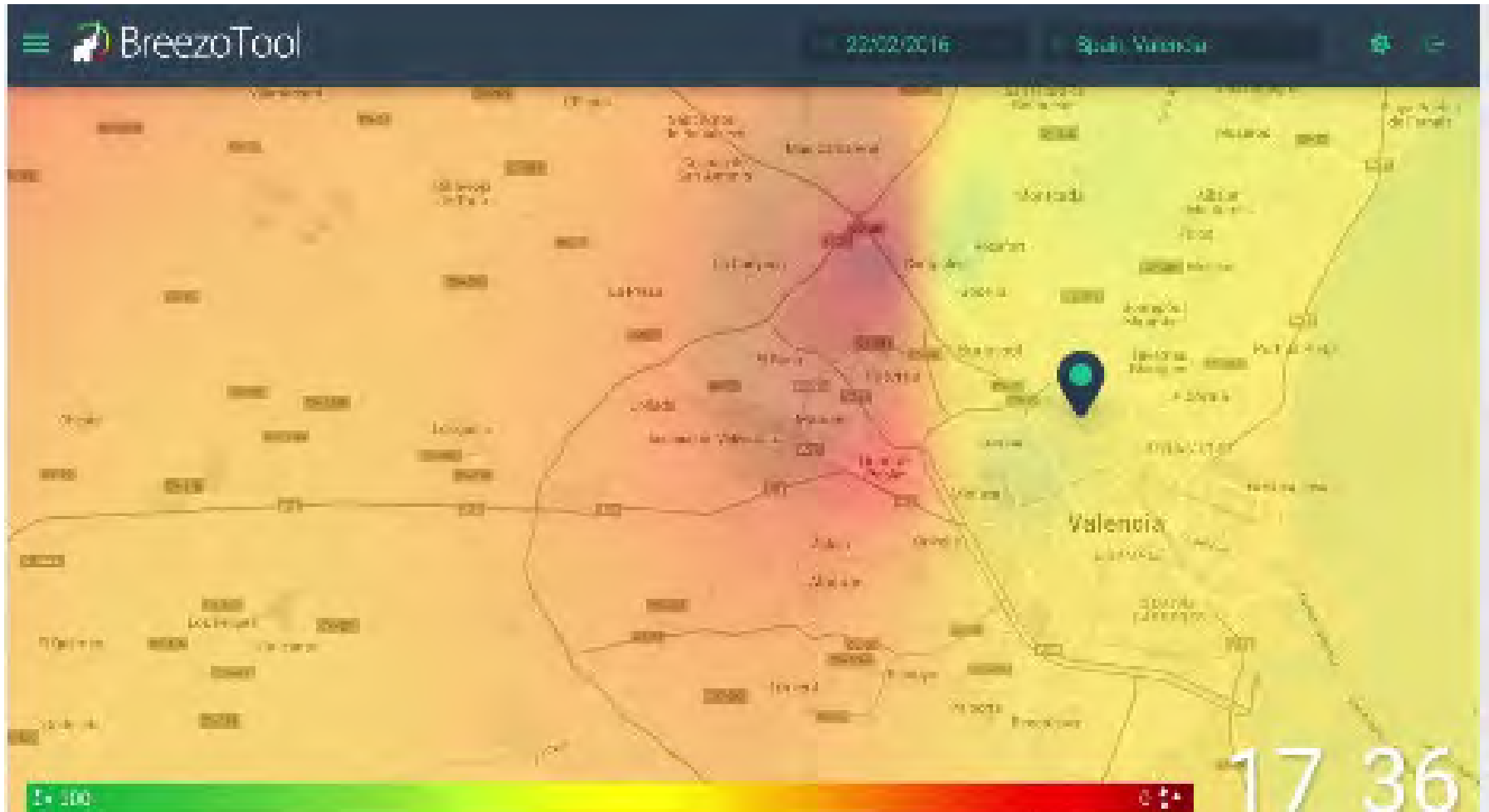
12.45



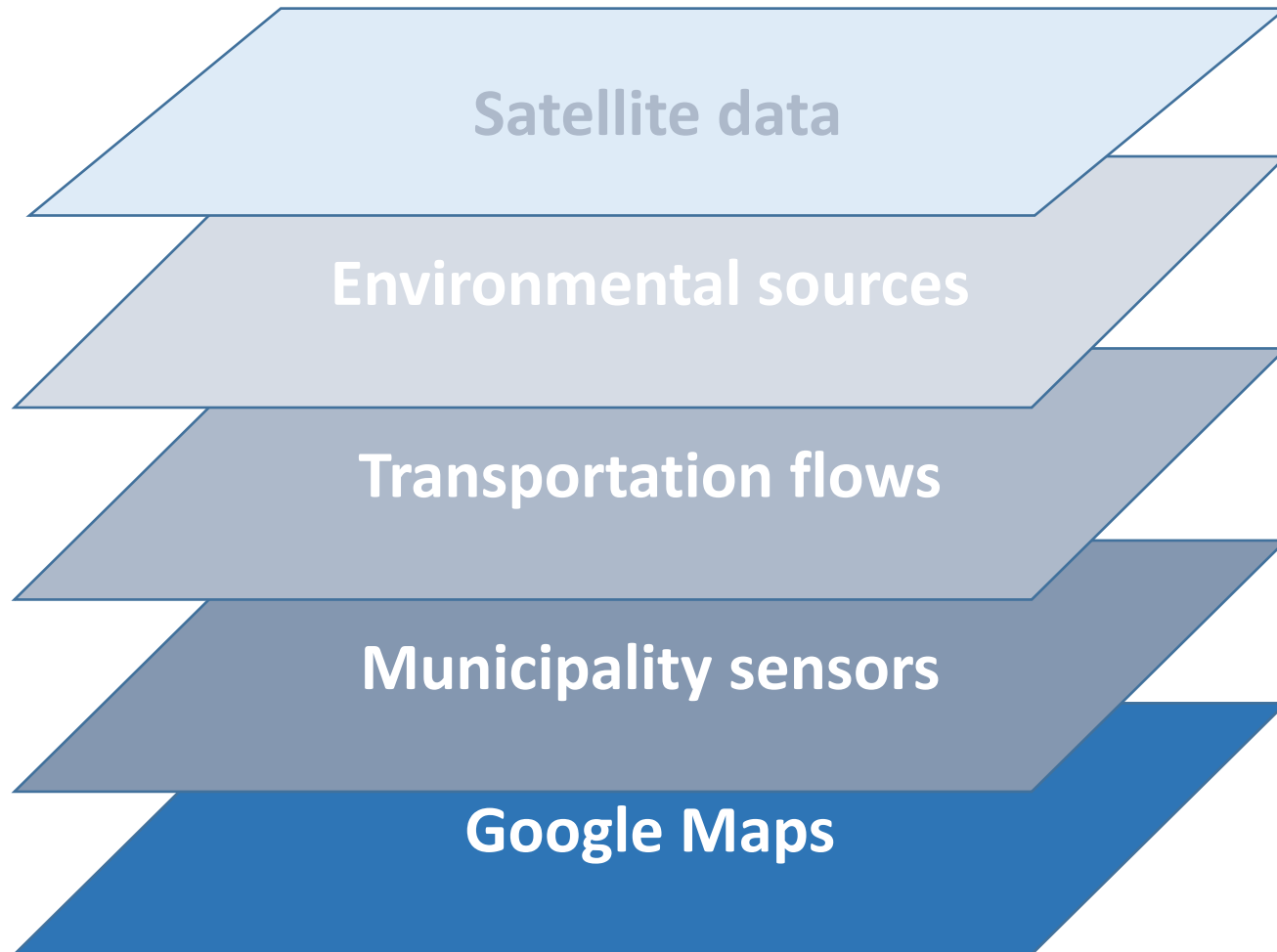
16.53



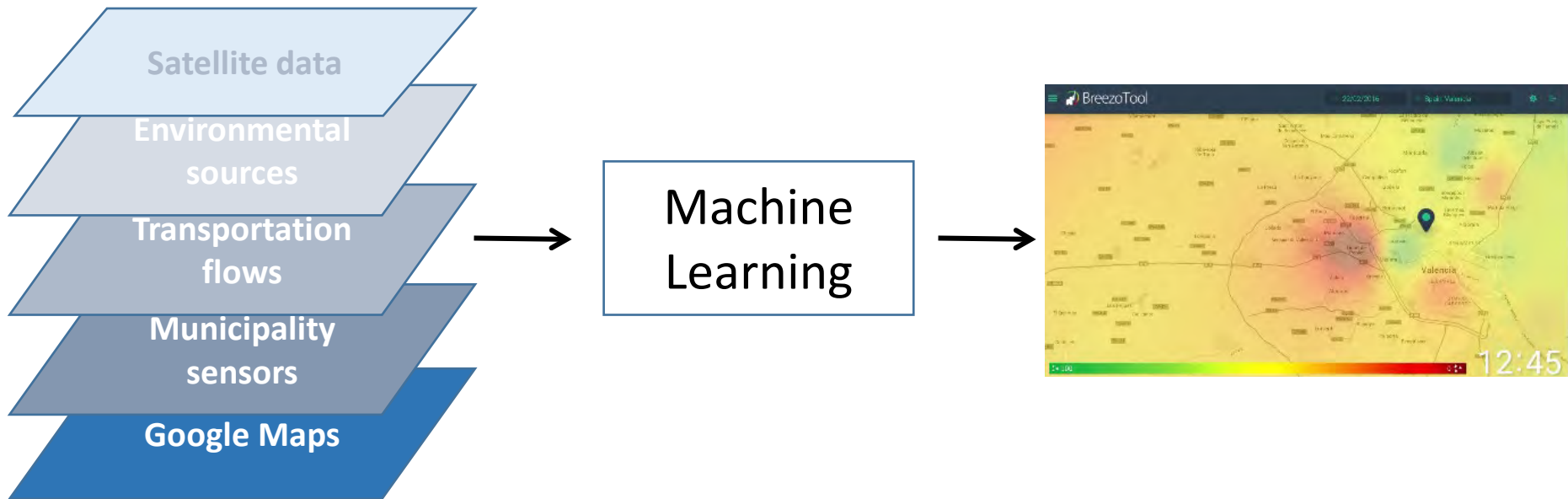
17:36



How data are produced in the app Breezometer

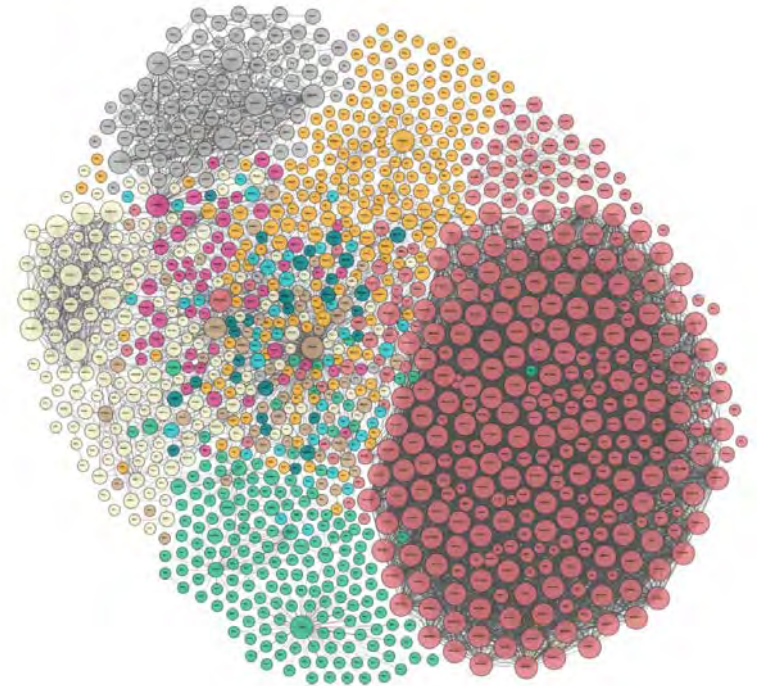
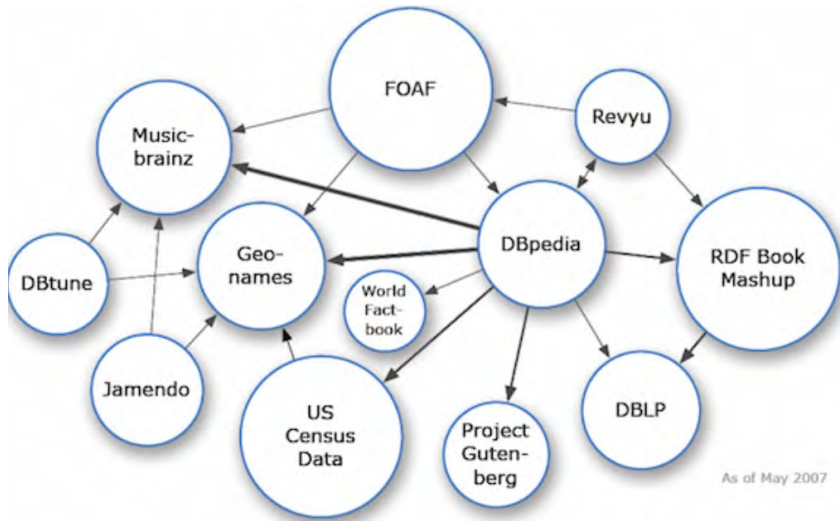


How forecasts are produced in Breezometer



LOD Cloud evolution

2007 **one billion triples** → 2017 > **200 billion triples**

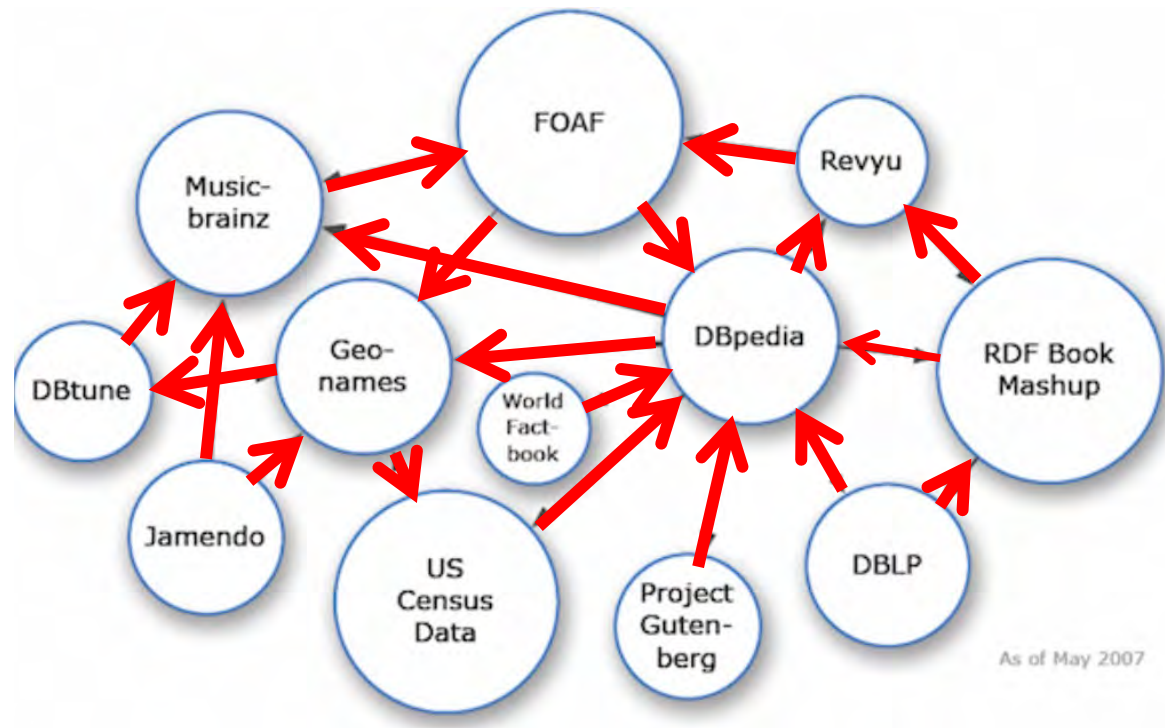


2007

2017

Il mondo è la totalità dei fatti, non delle cose

Collegare, integrare, riconciliare




I dati come bene economico

I dati sono il nuovo petrolio



Prosperità e Problemi



*Credo che i dati siano il nuovo petrolio.
E così come il petrolio ha portato prosperità
ma anche problemi, lo stesso faranno i dati*

Mikko Hypponen
Chief Research Officer, F-Secure

I dati come bene economico:
scarsità, codifica, astrazione

The I-Space (Figure from Boisot 1995)

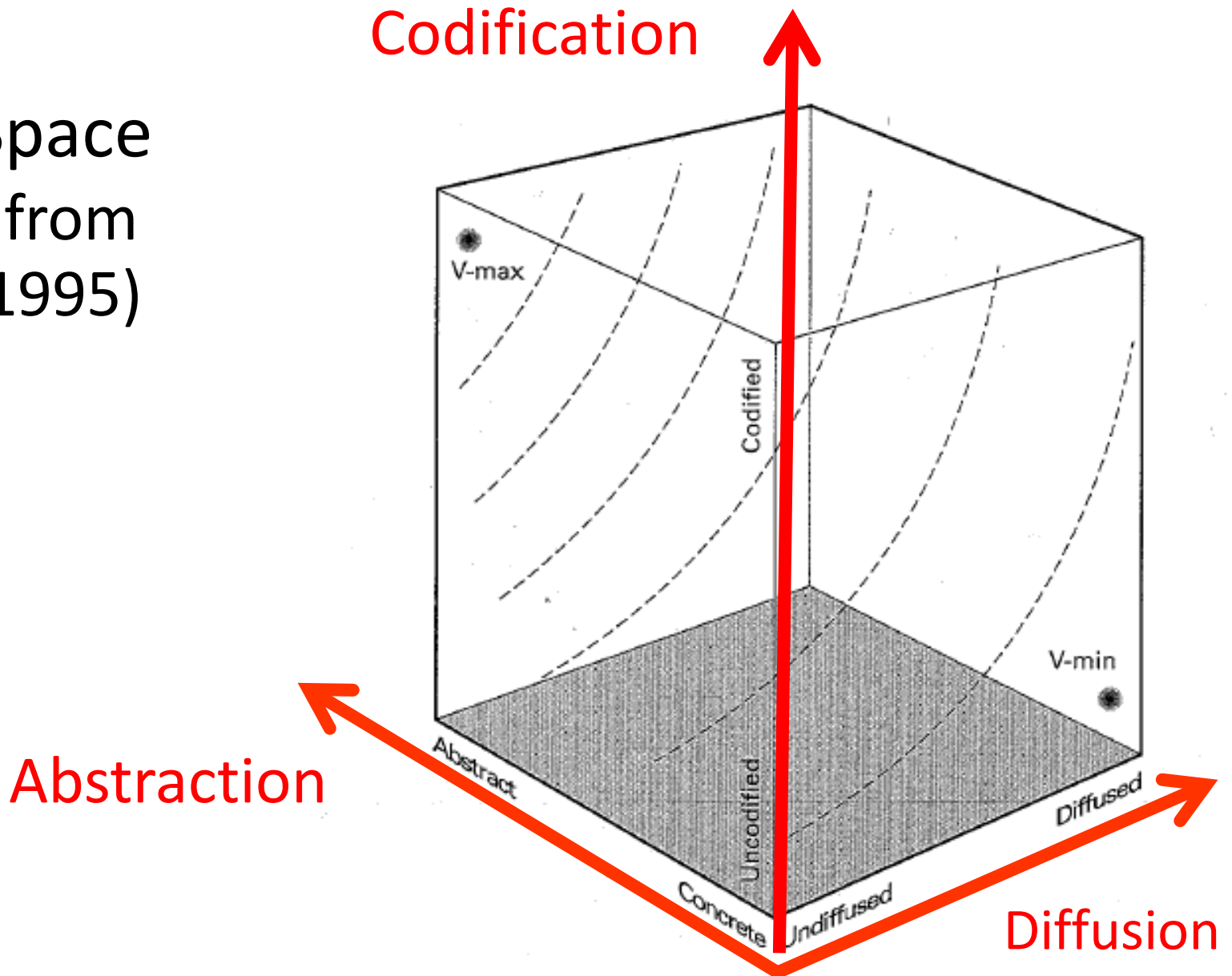


FIG. 4.1 Maximum and Minimum Value in the I-Space: V-max and V-min

Abstractions from data to big data

Abstractions Enable Change of paradigm in Machine Learning Techniques

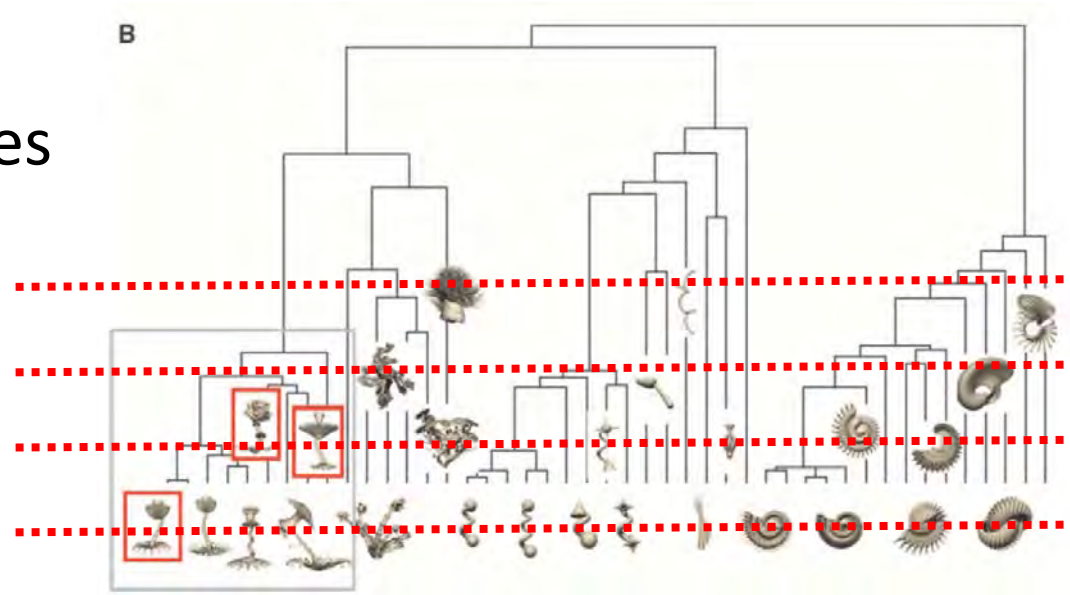
Volume

Big
Data

Small
Data

Deep
learning

Probabilistic Generative
models: **Bayes rule**



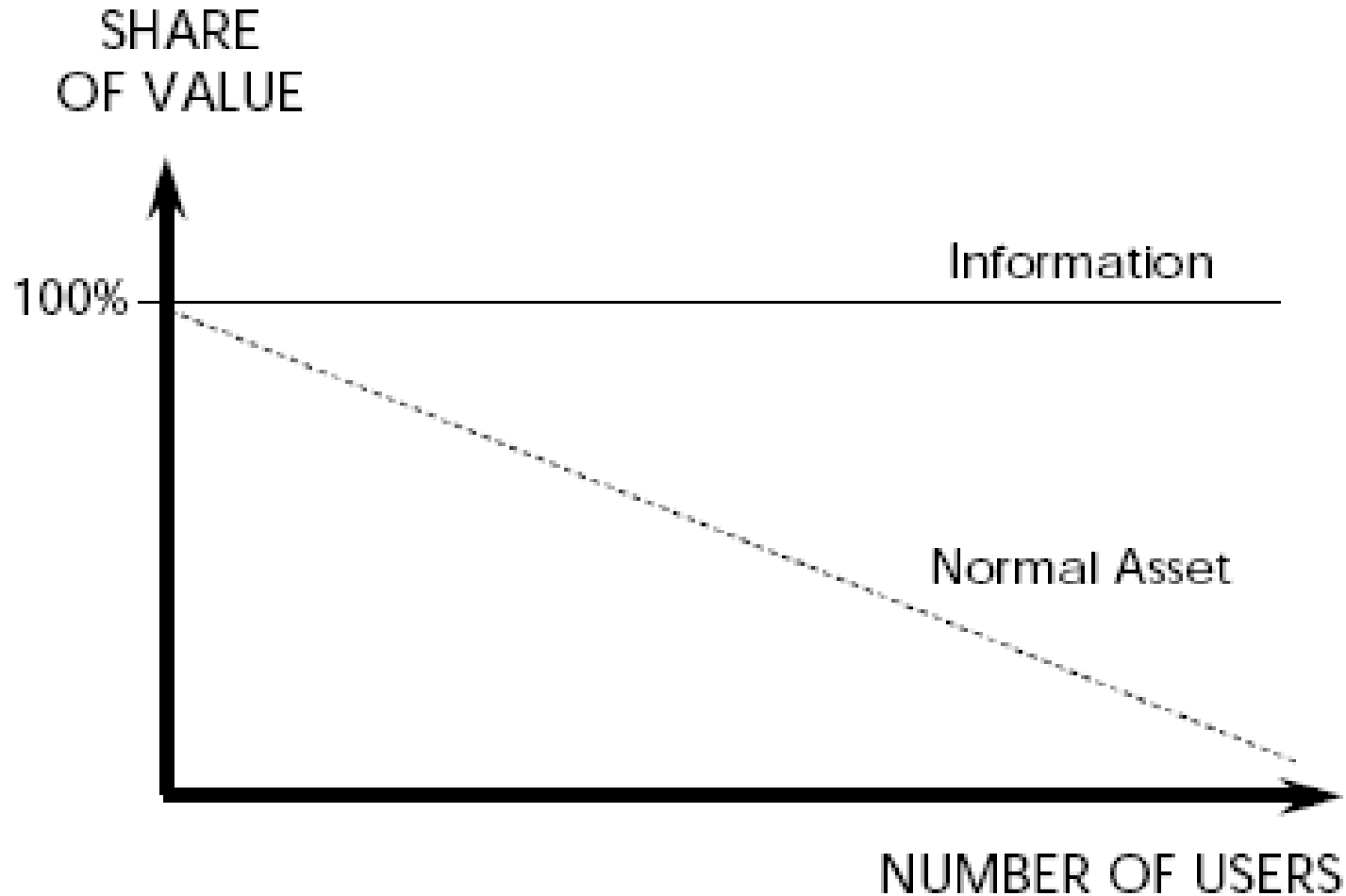
Da una economia goods & services -driven
a una economia data-driven

Pricing information

- Information is costly to produce (and may have high sunk cost too) but cheap to reproduce and to deliver (think to a paper book vs an ebook)
- There are no natural capacity limits for additional copies.
- Information can be flexibly tailored to the needs of different customers

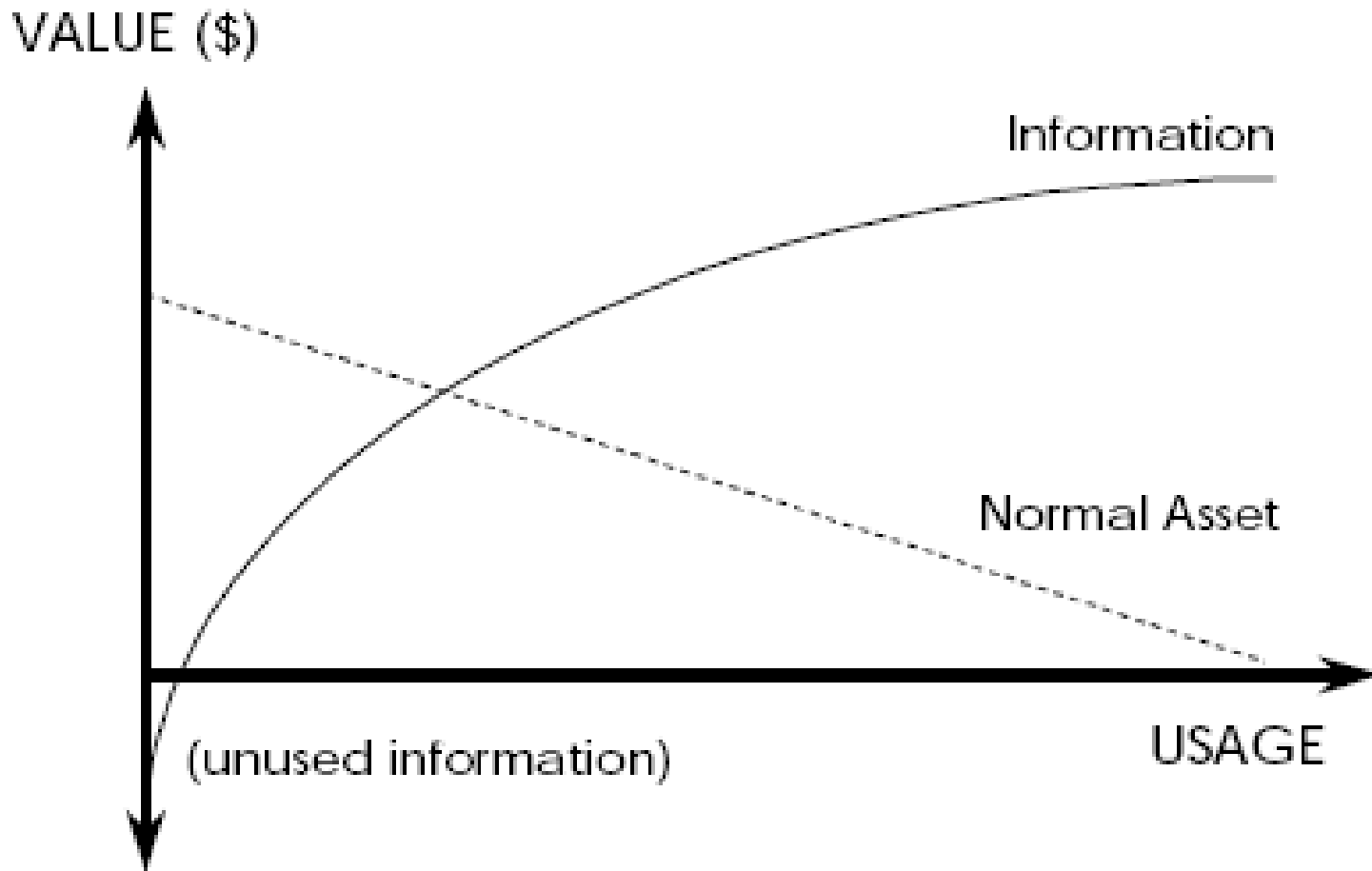
Moody & Walsh, 1999 Legge 1

L'informazione e' infinitamente divisibile

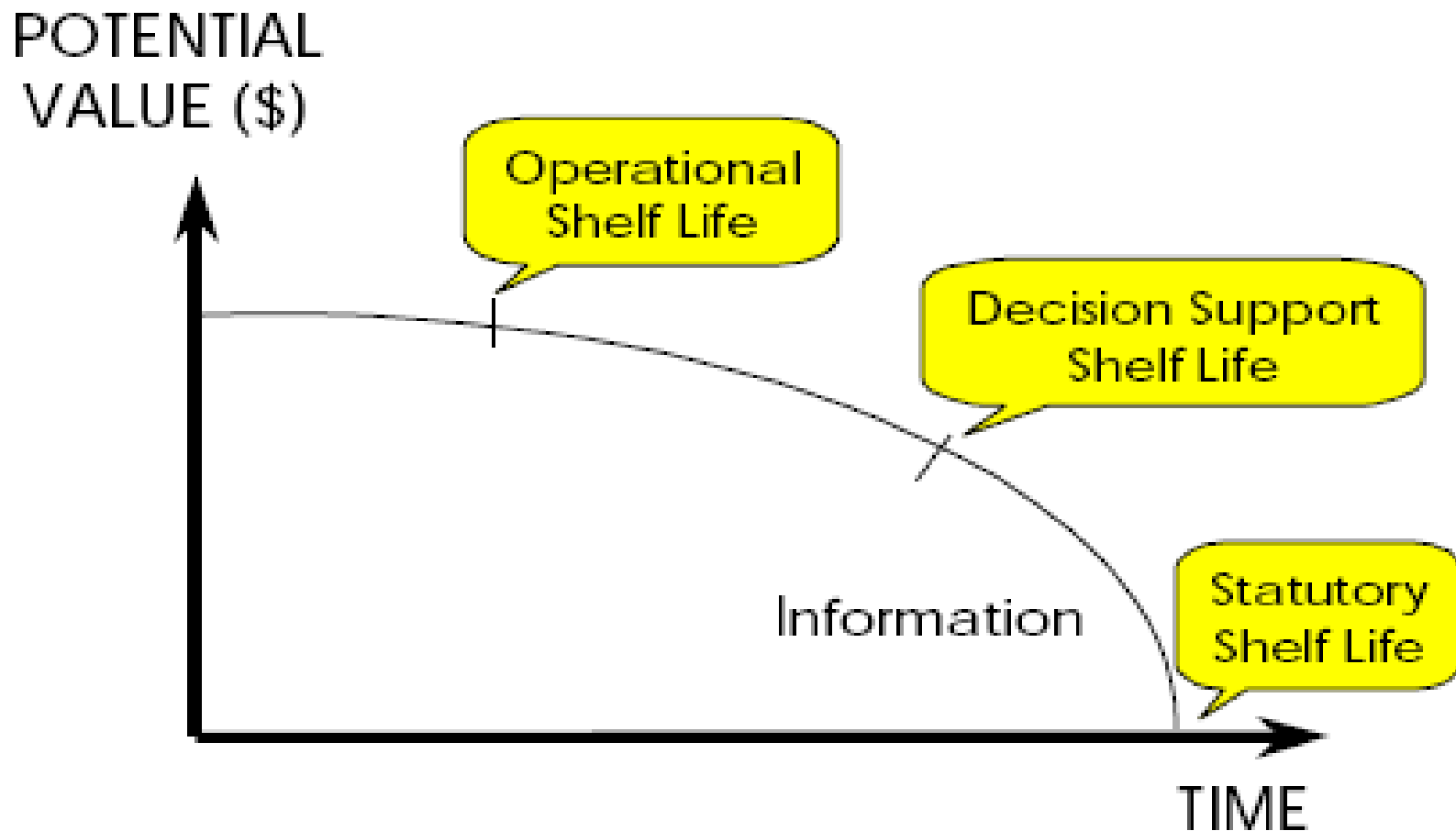


Moody & Walsh, 1999 Legge 2

Il valore della informazione cresce con l'uso

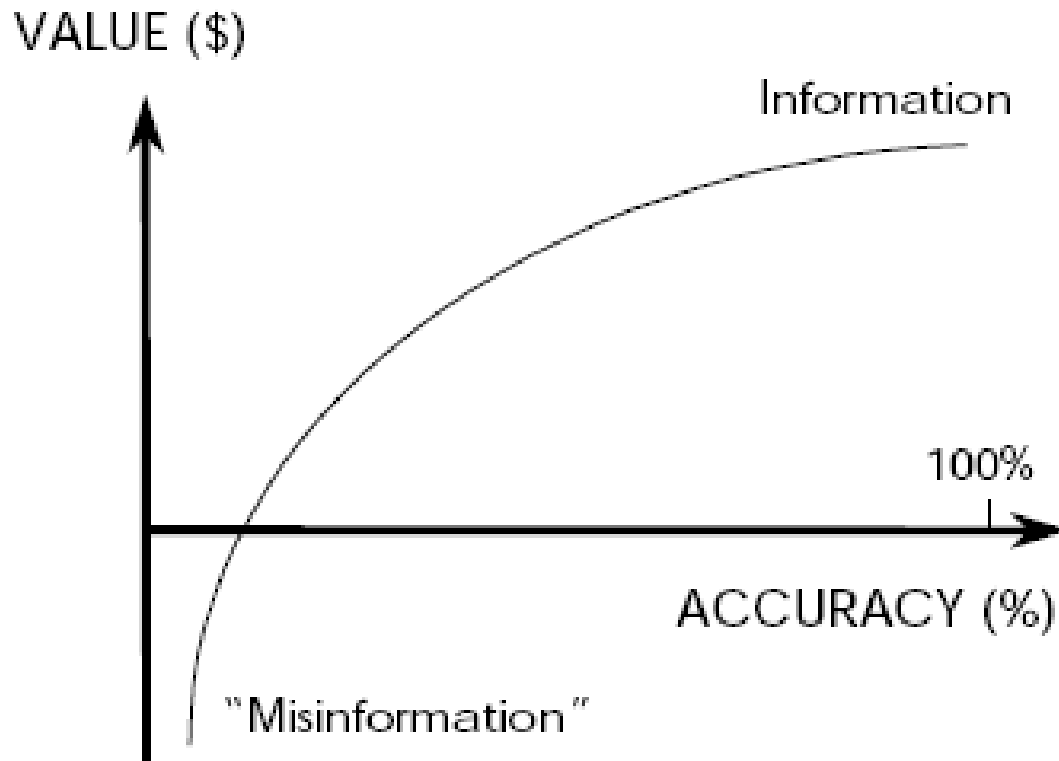


Moody & Walsh, 1999 Legge 3
L'informazione non e' scarsa...ma e' deperibile



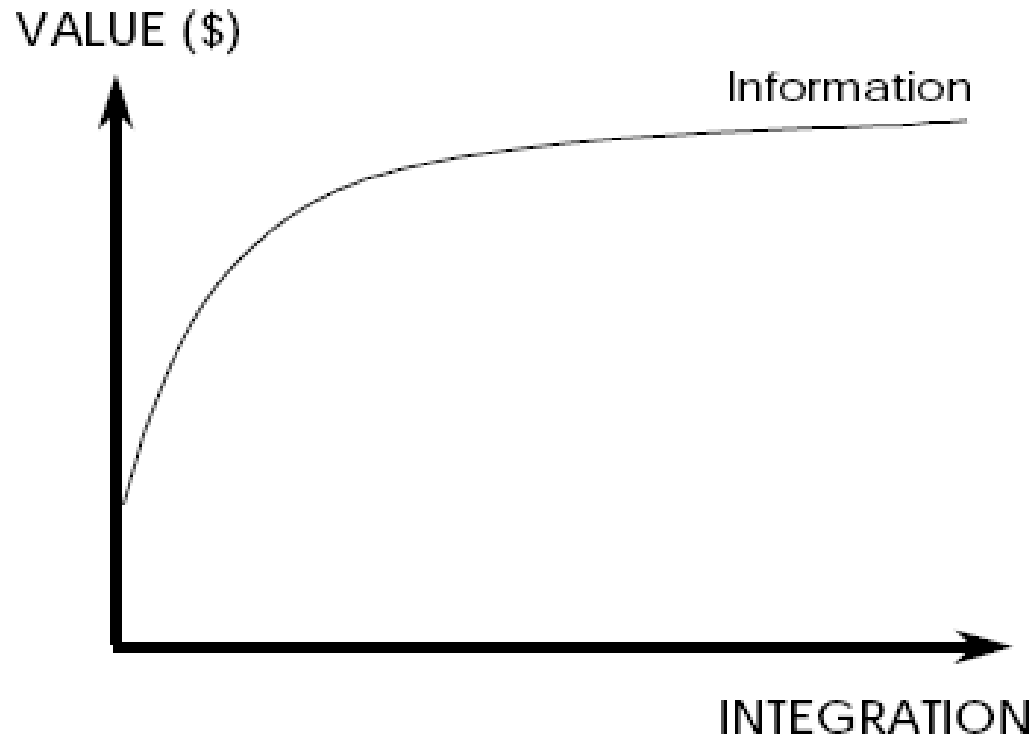
Law 4: The value of information increases with accuracy

The more accurate information is, the more useful and therefore valuable it is. Inaccurate information can be very costly to an organisation in terms of both operational errors and incorrect decision making.



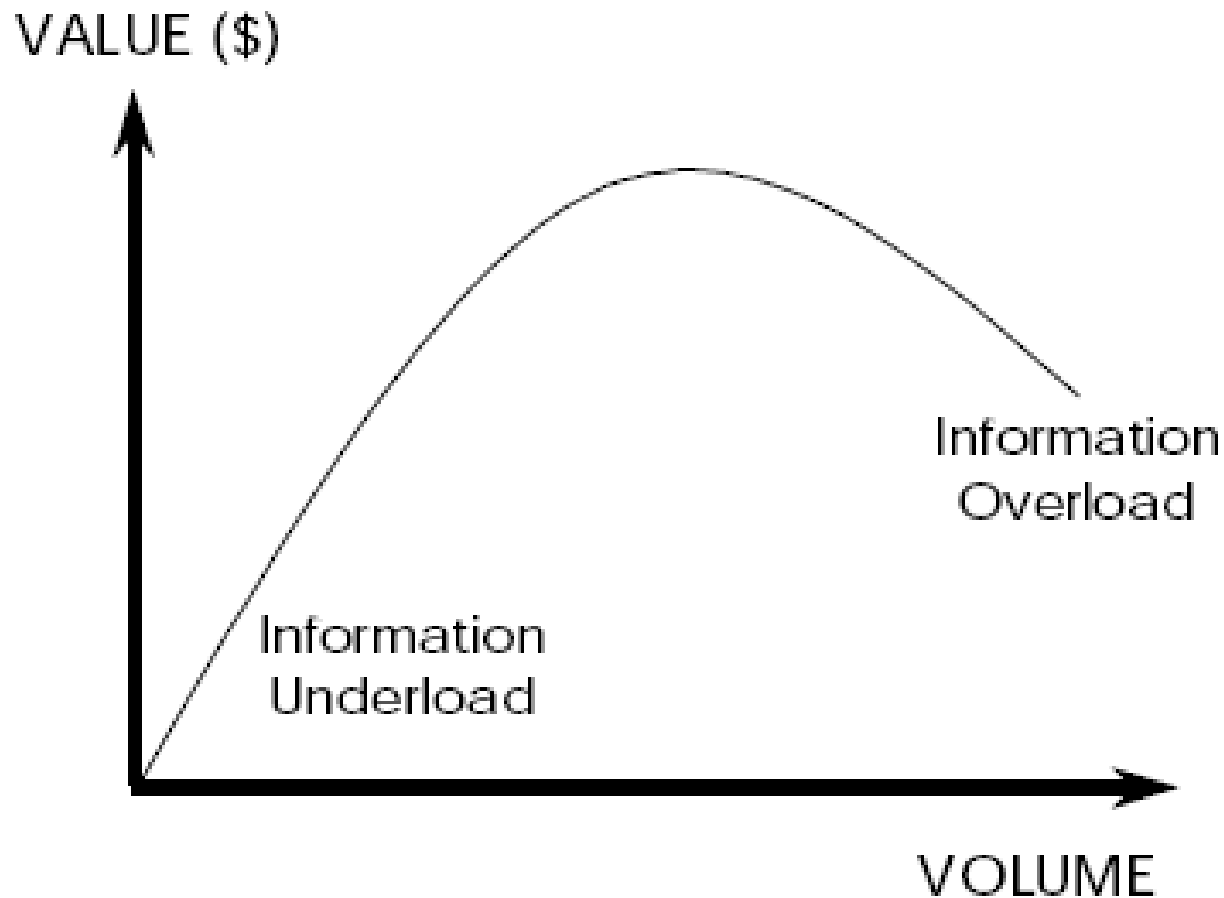
Law 5: The value of information increases when combined with other information

Information generally becomes more valuable when it can be combined with other information. For example, customer info and sales info are each valuable info sources in their own right, but being able to relate the two sets of info. Together is infinitely more valuable from a business viewpoint.



Law 6: More is not necessarily better

Information overload



Reinventing Capitalism in data-rich markets


1. Il prezzo «aumentato»


- The key difference between conventional markets and data-rich ones is the role of information flowing through them and how it gets translated into decisions.
- In data-rich markets, we no longer have to condense our preferences into **price** and can abandon the oversimplification that was necessary because of **communicative and cognitive limits**.
- This makes possible to pair **decentralized** decision making, with its valuable qualities of robustness and resilience, with much improved transactions **efficiency**.

Esempio

Confronta TUTTE le Tariffe e i Prezzi di TUTTI i gestori di Energia e Gas in Italia!

Di quale servizio hai bisogno?

 LUCE

 GAS

[AVANTI >](#)

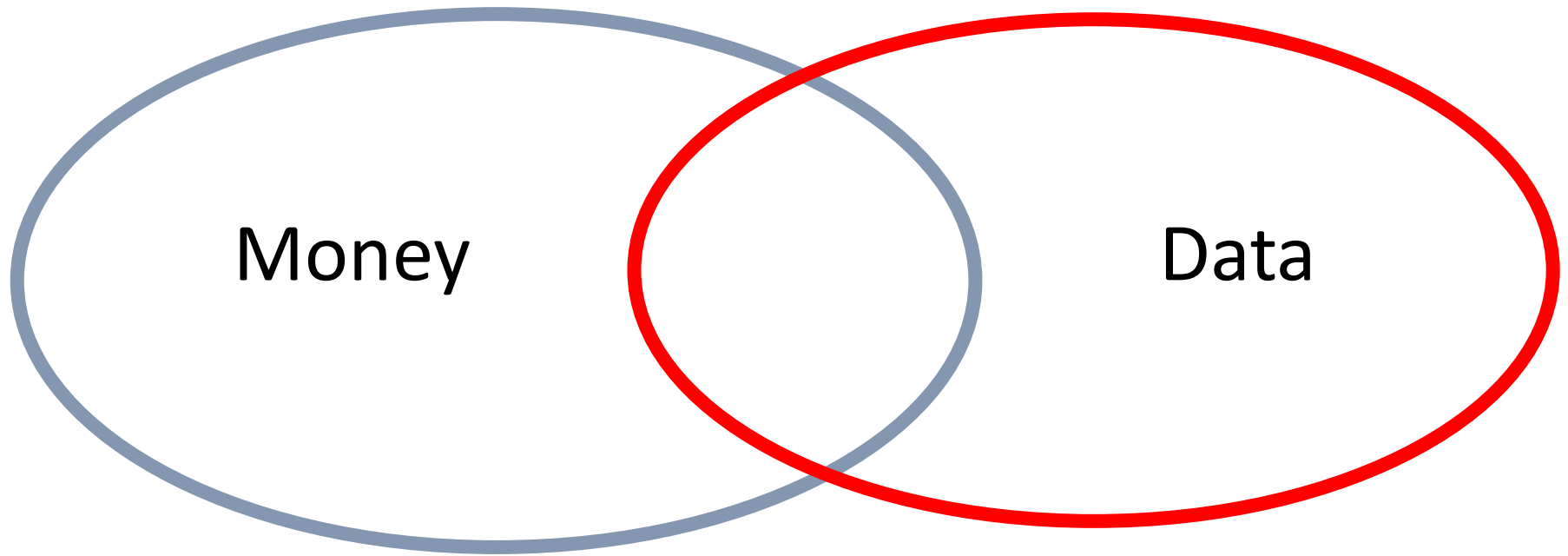
I VANTAGGI PER TE

- ✓ Servizio imparziale e personalizzato
- ✓ Preventivi e Consulenza gratuiti
- ✓ Le migliori offerte a confronto

ALCUNI GESTORI CONFRONTATI



Money and data



Data rich markets

- For hundred of years, Kerala fishermen, in absence of information on the market, were forced to make risky calculation when they had to steer boats to the market.
- When in 1997 mobile phone tower were installed in a series of coastal towns, the fisherman could transact with buyers while they were still out on the water.
- For Kerala fishermen, mobile phones become an empowering communication tool because they enabled one-to-one conversation with their potential buyers.
- Thanks to a better flow of information, the market became vastly more efficient.

Capital decline

- If markets teem with information that facilitates transactions, that information itself holds value.
- In the future we may see transactions paid in data rather than money.
- Infact, Google and Facebook would not be what they are without the billions of users who pay for their services with personal data.
- Similarly in an increasing number of cases, companies contract with outside services to have their data analyzed for them and they pay these services essentially with data, by letting them reuse the data for other purposes.

Market effects

Three distinct effects are often at play when markets become concentrated:

- Scale effects
- Network effects
- Feedback effects

The scale effect lower costs, the network effects expands utility, and the feedback effects improves the product.

Feedback effects - 1

- Online markets seem particularly vulnerable to concentration.
- Feedback effects occur whenever computer systems use feedback data to learn.
- Services based on machine learning systems fueled by feedback data «buy» innovation at diminishing cost as the user base grows.
- This has great implications for market competition.
- Start ups cannot compete with big companies as they lack the volume of feedback data that can drive product development.

Feedback effects - 2

A progressive **datasharing mandate** is needed

- Algorithmic transparency is useful, but not enough.
- Regulators wanting to ensure competitive markets should mandate the sharing of data.
- In such a way, the value derived from the data gets spread about.
- Furthermore, it is crucial that machine learning systems not only work well, but also work sufficiently independently of each other.
- If each is designed and developed independently of the others, it is less likely they will share common flaws.

Il valore sociale dei dati

The Economist 2011

Social value - Quality of health care in Uganda



World politics Business & finance Economics Science & technology Culture Blogs Debate & discuss Multimedia Print edition

Audio
Video
The Economist audio edition

The Open Government Partnership

The parting of the red tape

Is it just another global talking-shop—or a fresh approach to shaking out government secrecy?

Oct 8th 2011 | NEW YORK AND TALLINN | from the print edition

Like 151

0

UGANDA is not best known as a testbed for new ideas in governance. But research there by Jakob Svensson at the University of Stockholm and

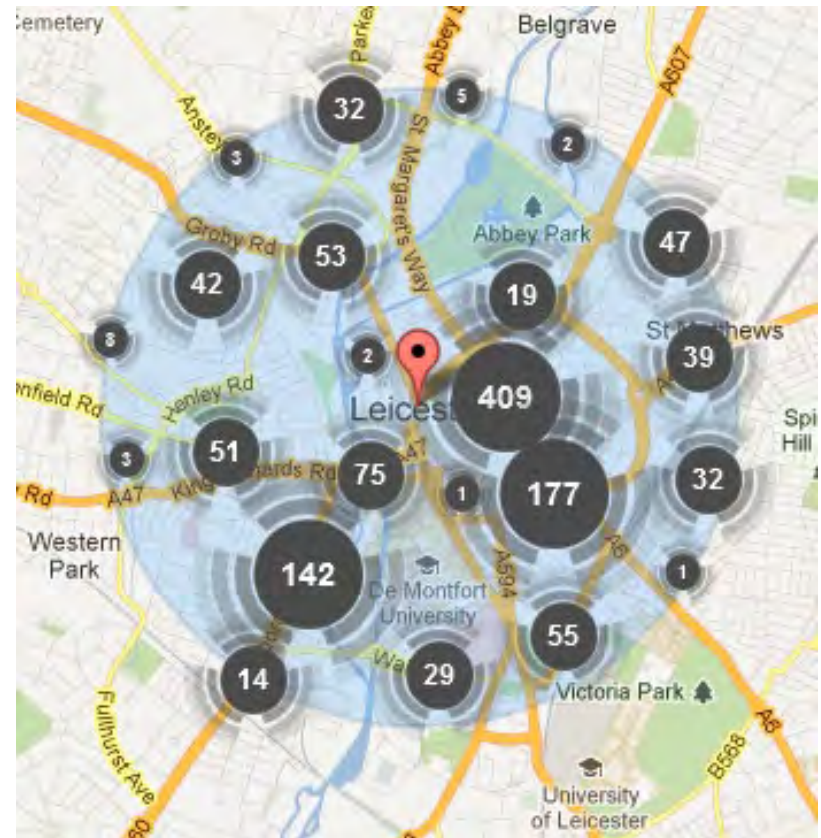
colleagues suggested that giving people health-care performance data and helping them organise to submit complaints cut the death rate in under-fives by a third.

Publishing data on school budgets reduced the misuse of funds and increased enrolment.



Crimes at Leicester, value for me...

All crime	1241
Burglary	75
Anti-social behaviour	317
Robbery	12
Vehicle crime	57
Violent crime	181
Public disorder and weapons	45
Shoplifting	172
Criminal damage and arson	102
Other theft	178
Drugs	48
Other crime	54



...and value for house landlords

All crime	1241
Burglary	75
Anti-social behaviour	317
Robbery	12
Vehicle crime	57
Violent crime	181
Public disorder and weapons	45
Shoplifting	172
Criminal damage and arson	102
Other theft	178
Drugs	48
Other crime	54



What the Leicester example shows: Tension between economic value and social utility

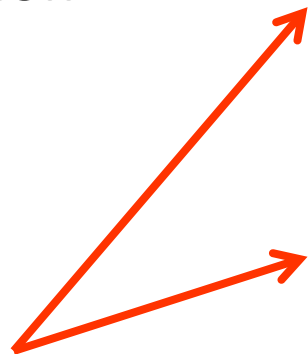
- Data can provide the user a **positive social value** or else a **negative economic utility**, resulting in a well known tension in the history of human mankind.

Tecnologici



Indicatori per misurare l'impatto dei siti di open data

European Commission 2011

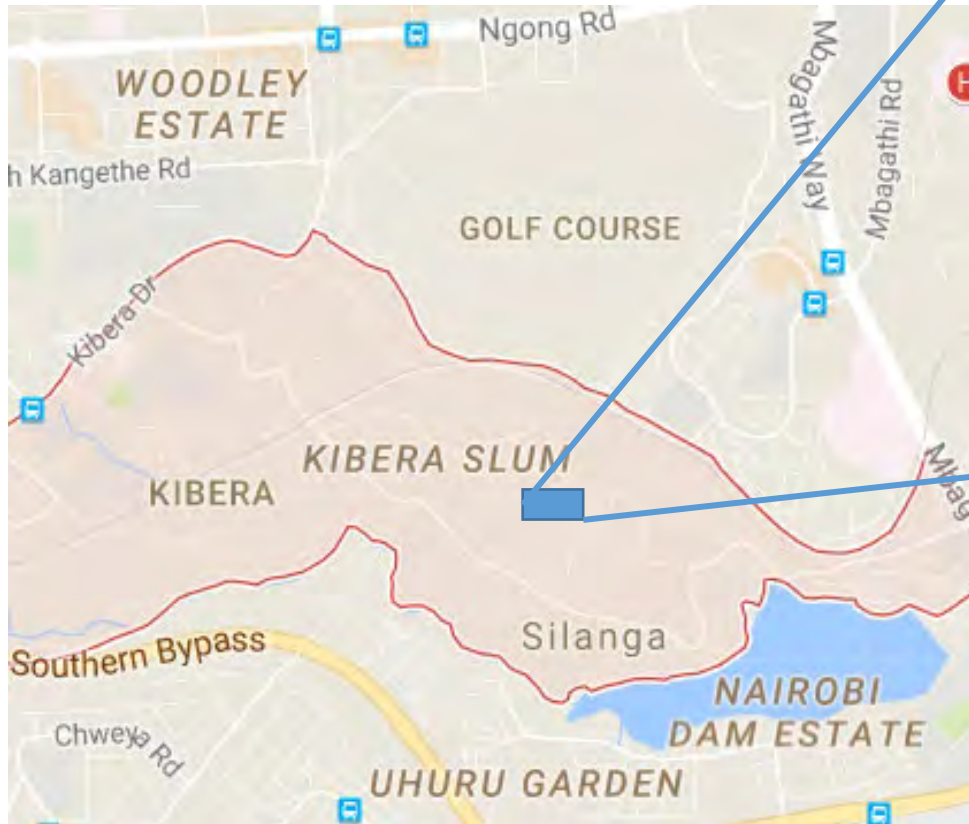


Sociali

1. Number of datasets available or number of dataset listing pages	
2. Implementation of five star linked data system of Tim Berners-Lee ¹	<ul style="list-style-type: none">- data availability (★make your stuff available on the web (whatever format))- machine-readable data (★★make it available as structured data (e.g. excel instead of image scan of a table))- non-proprietary formats (★★★ non-proprietary format (e.g. csv instead of excel))- linked data (★★★★use URLs to identify things, so that people can point at your stuff, ★★★★★ link your data to other people's data to provide context)
4. Clear re-use conditions (license-free or license bound)	
5. Position of the governing body (high-level, low-level)	
6. Accessibility – anonymous access	
7. Take-up by citizens	<ul style="list-style-type: none">• Number of visits and unique visitors• Number of data downloads• Conversion rate (click-throughs to data hosting websites) in case of web repositories
8. Take-up by application developers	<ul style="list-style-type: none">• Number of apps• App contests entries numbers
9. Availability of take-up data on the website	
10. Range of government agencies contributing dataset listings	
11. Availability of 'suggest a dataset' form	
11. Presence of guidelines, events and other capacity building strategies	
12. Community engagement indicators	<ul style="list-style-type: none">• Number of blog comments, tweets, activities on fora, ATOM feeds subscriptions, other feedback channels and metrics

Dal digital divide al data divide

Data divide in maps

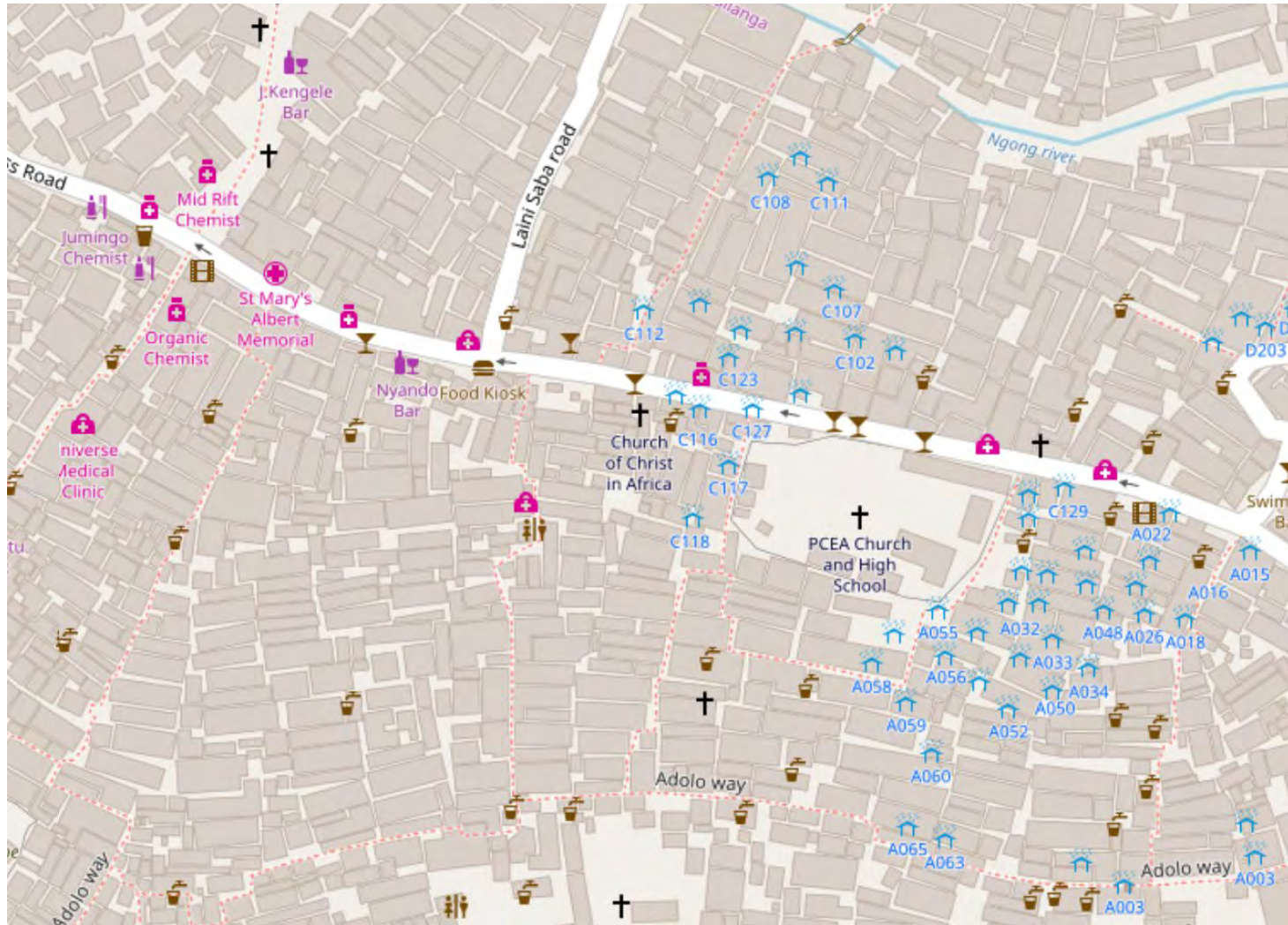


Kibera Slum

New York Times Square



Map Kibera Project

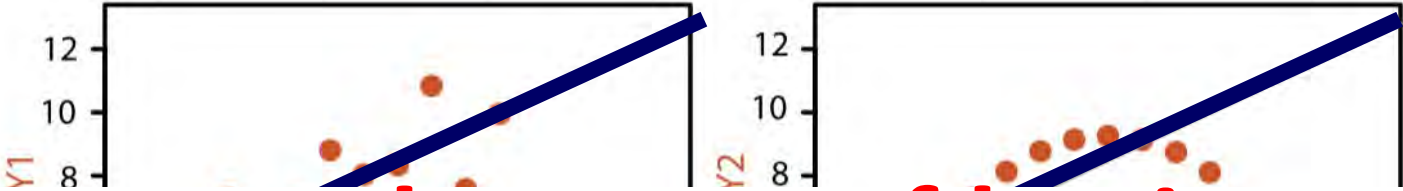


From Why to Correlation to What, or:
with enough data, “the data speak for
themselves” (the end of theory)

Example: when to buy a flight ticket
from **Why...**

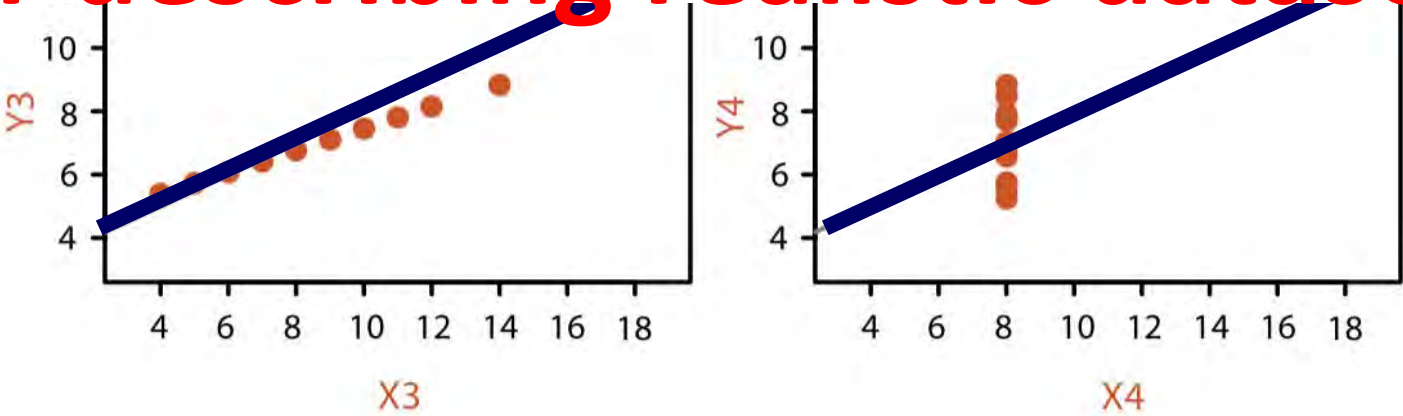
We can investigate to find on a sample
the law for pricing
applied by airline companies

Anscombe's Quartet is four datasets with identical simple statistical properties: mean, variance, correlation, and linear regression line. However, visual inspection immediately shows how their structures are quite different

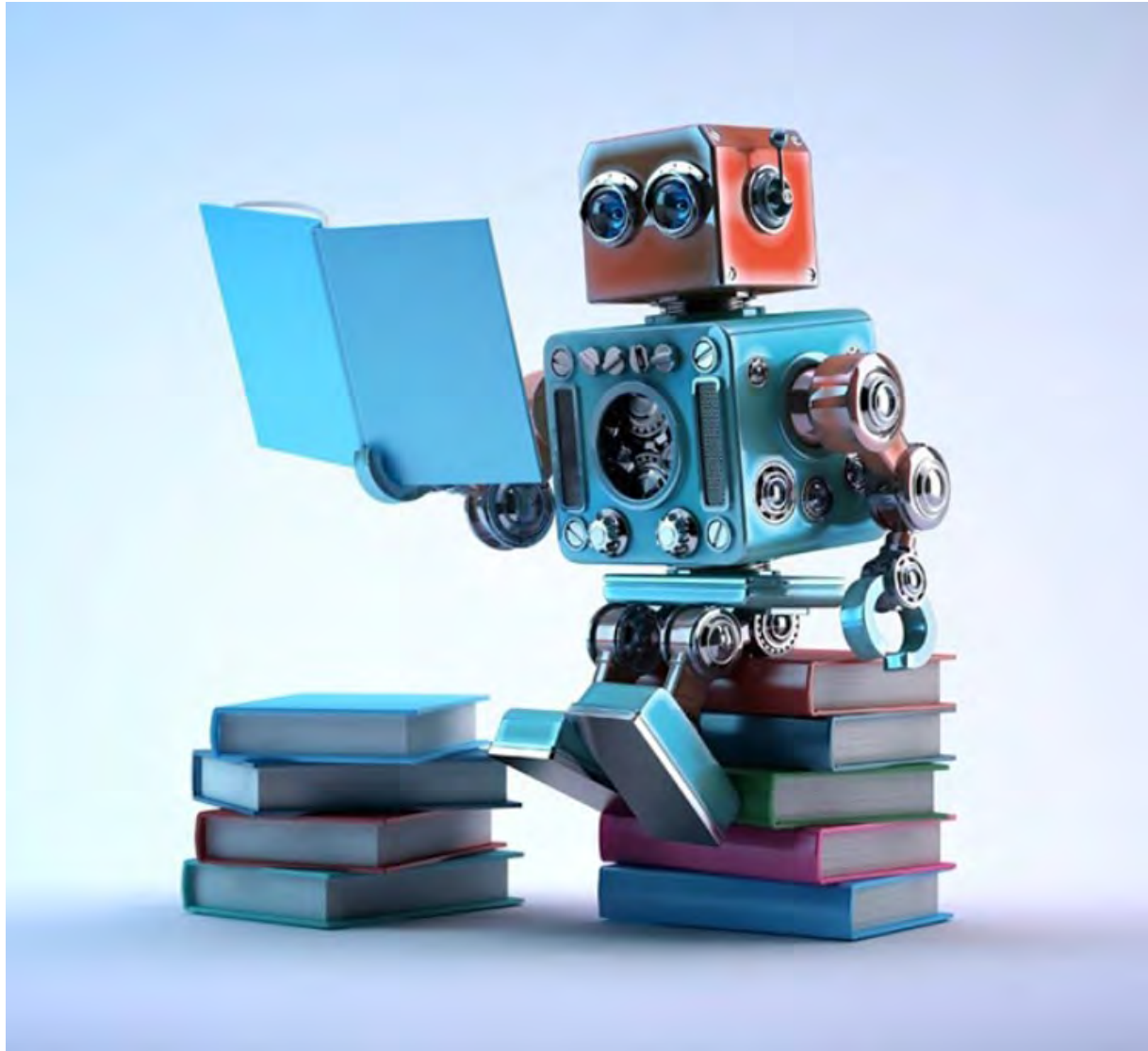


**Inadequacy of basic
statistical properties**

for describing realistic datasets



Machine learning



... to **What**
Oren Etzioni's Farecast

Sample of
12.000
tickets



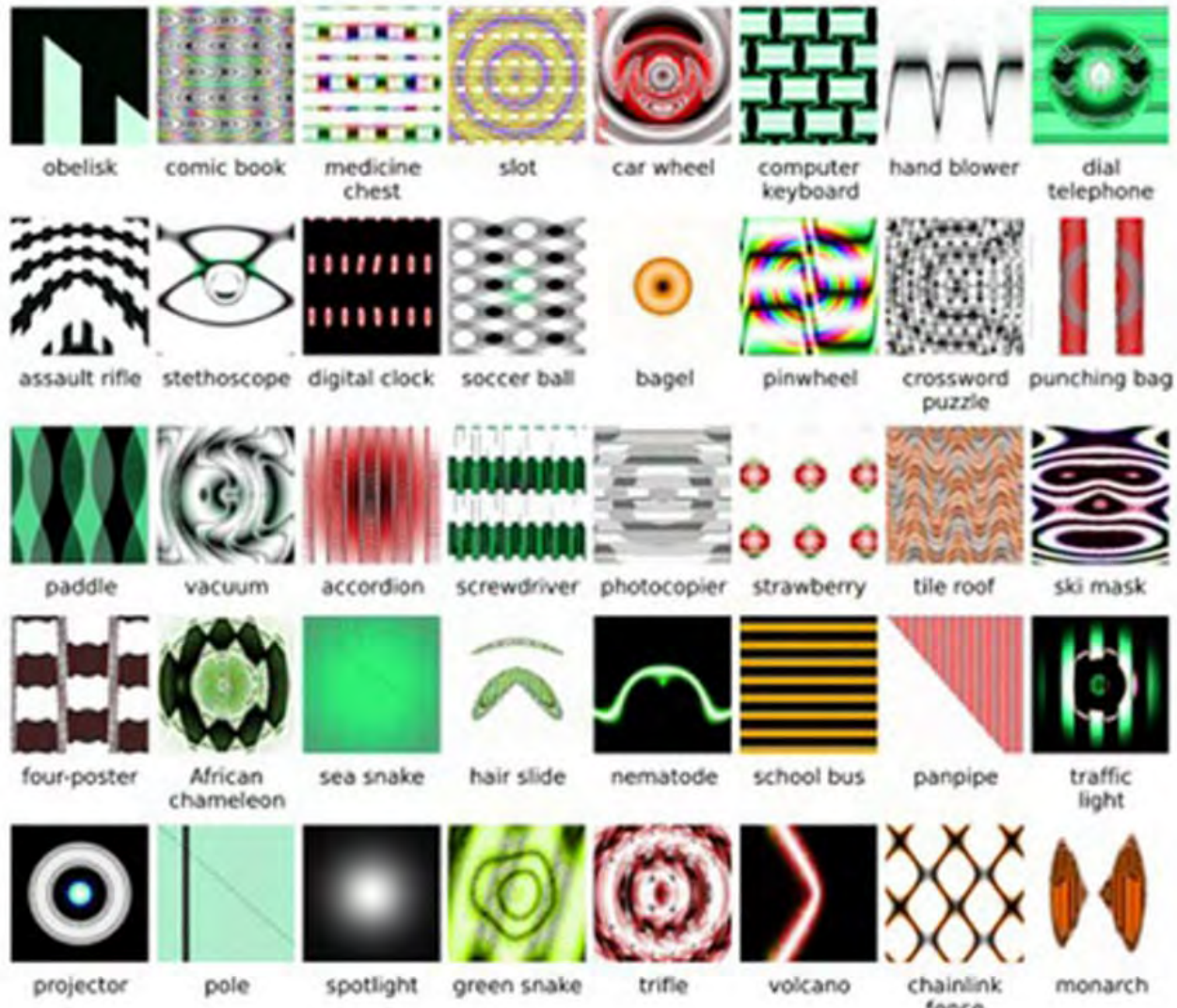
200 10⁹

50 \$ average savings per ticket
Farecast sold for 110 10⁶ \$

Chris Anderson - 'The End of Theory: The Data Deluge Makes the Scientific Method Obsolete', 2008

- Forget taxonomy, ontology, and psychology. Who knows **why people do what they do?** The point is **they do it**, and we can track and measure it with unprecedented fidelity. With enough data, **the numbers speak for themselves.'**

Apophenia in machine learning



Source: Anh Nguyen, Jason Yosinski, and Jeff Clune, "Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images," cv-foundation.org, 2015 →.

Fairness-aware learning techniques vs Self-fulfilling prophecies: the case of predictive policing

Example: ProPublica analysis of COMPAS

- Criminal justice: recidivism algorithms (COMPAS)
- Predicting if a defendant should receive bail
- Unbalanced false positive rates: more likely to wrongly deny a black person bail

Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Ethical Goals of Datacy

- Fairness, Accountability and Transparency in Machine Learning Models
- Right to explanation and to interpret the outcome

Very good news! General Data Protection Regulation

European General Data Protection Regulation (GDPR)



in 1 minute

Countdown to
compliance
deadline

25 May 2018



Does GDPR enforce a right to explanation?

- The idea that the GDPR mandates a “right to explanation” from machine learning models — meaning that those significantly affected by such models are due an accounting of how the model made a particular decision — has become a controversial subject.
- Some scholars, for example, have spoken out vehemently against the mere possibility that such a right exists.
- Others, such as the UK’s own Information Commissioner’s Office, seem to think the right is pretty clearly self-evident.

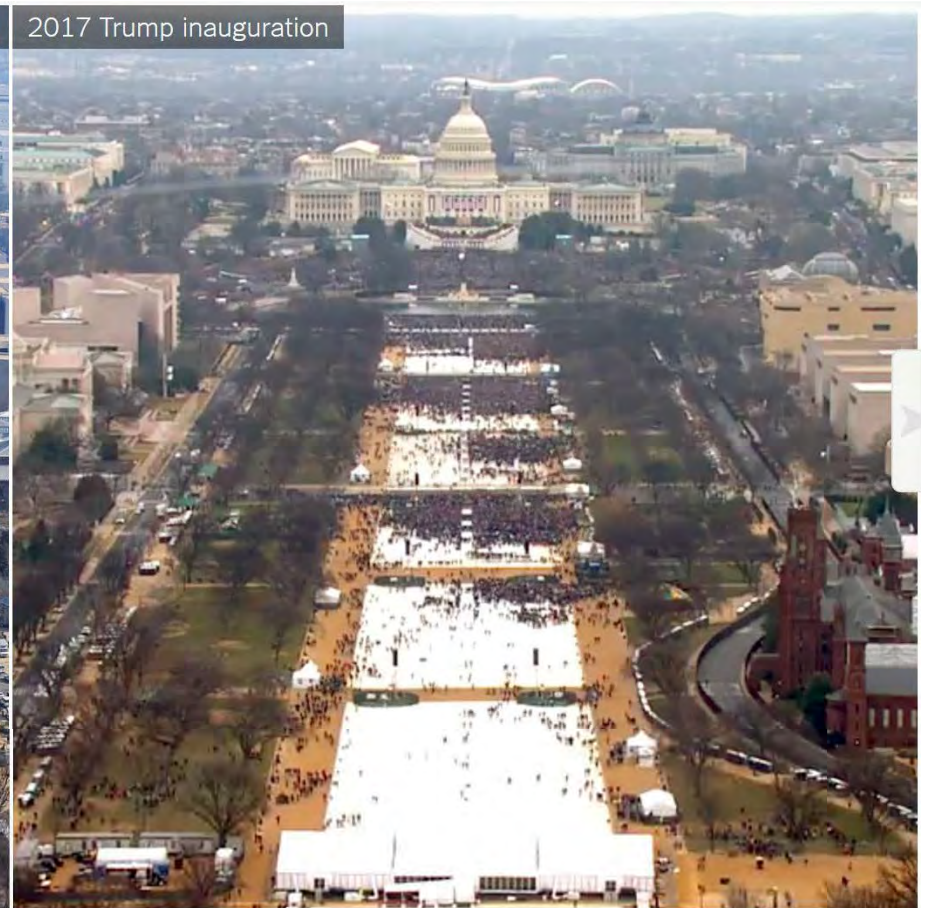
Dalla verità alla post-verità

festival *filosofia* verità

Modena Carpi Sassuolo 14_15_16 settembre 2018



I «fatti alternativi» dello staff di Trump



Alternative facts - 2



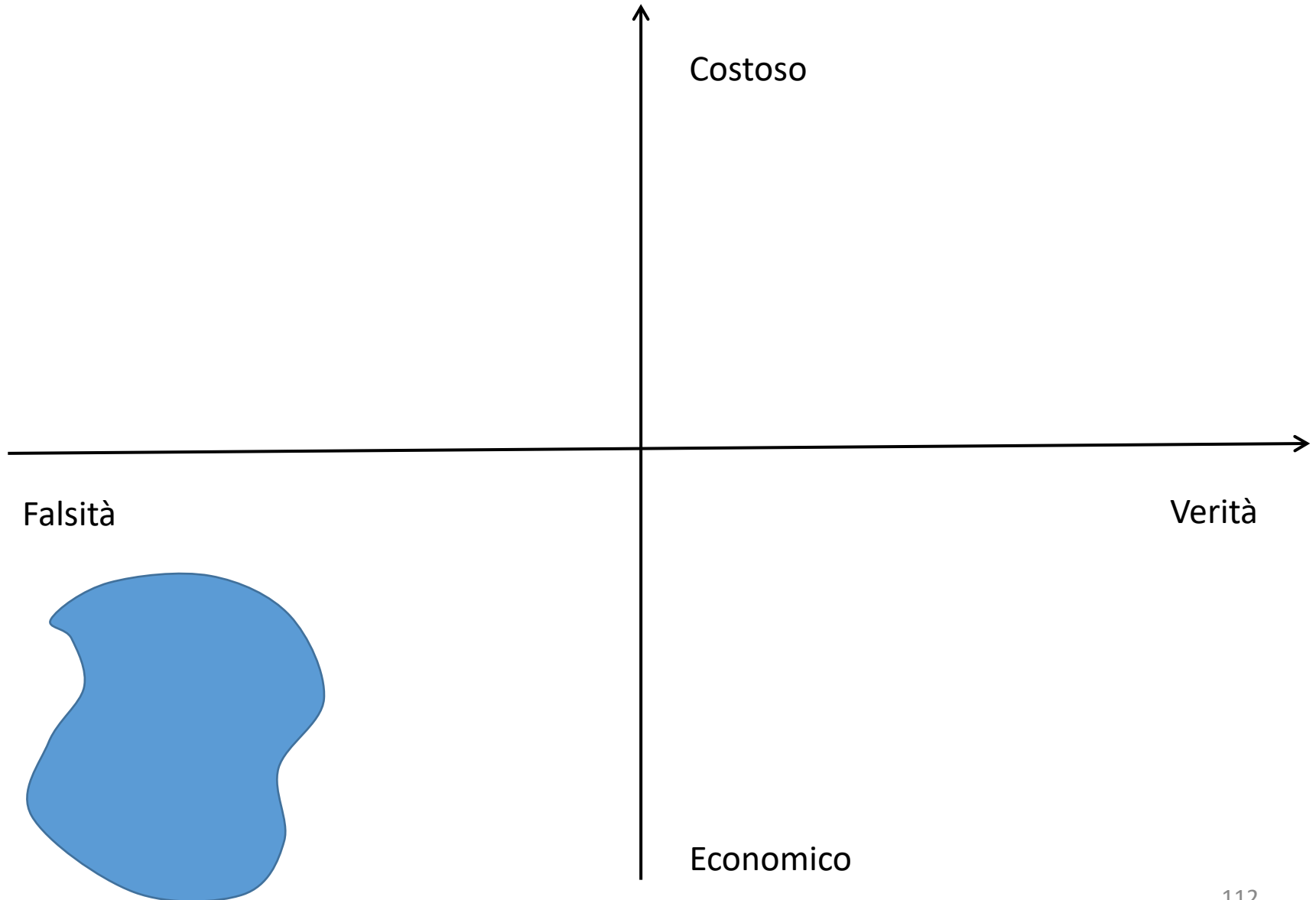
i Kellyanne Conway denies Trump press secretary lied: 'He offered alternative facts'

Stabilire la verità costa

Dal Guardian 1 settembre 2018

- Establishing the truth costs money – wages, travel, hotel bills.
- Journalists who learned their craft in the pre-digital age never knew too much about profit and loss.
- As Rusbridger writes, the phrase “business model” never came up in the pub (and it was only in the late 1980s that I first heard the word “budget” in an editorial meeting).
- The bad news is that digital technology and the web have created “the most prodigious capability for spreading lies the world has ever seen”, while the economics that support truth-seeking journalism have never looked feebler. To adapt the dictum of one of Rusbridger’s predecessors: **bad facts are free and good ones expensive.**

Stabilire la verità e' un processo molto faticoso...



Dalla psicologia cognitiva non arrivano buone notizie

**Non è tanto
rilevante ciò che
la gente pensa,
ma come pensa.
Riconoscere la
cattiva
informazione
richiede processi
cognitive
complessi**

**Un semplice
mito è più
attraattivo
cognitivamente di
una
complicata
correzione**

**Per coloro che
sono
fortemente
convinti delle
proprie idee,
gli argomenti
fortemente
contrari
possono
rafforzare le
loro convinzioni**

Misinformation

Fake news

Post truth

False amplifiers

Disinformation

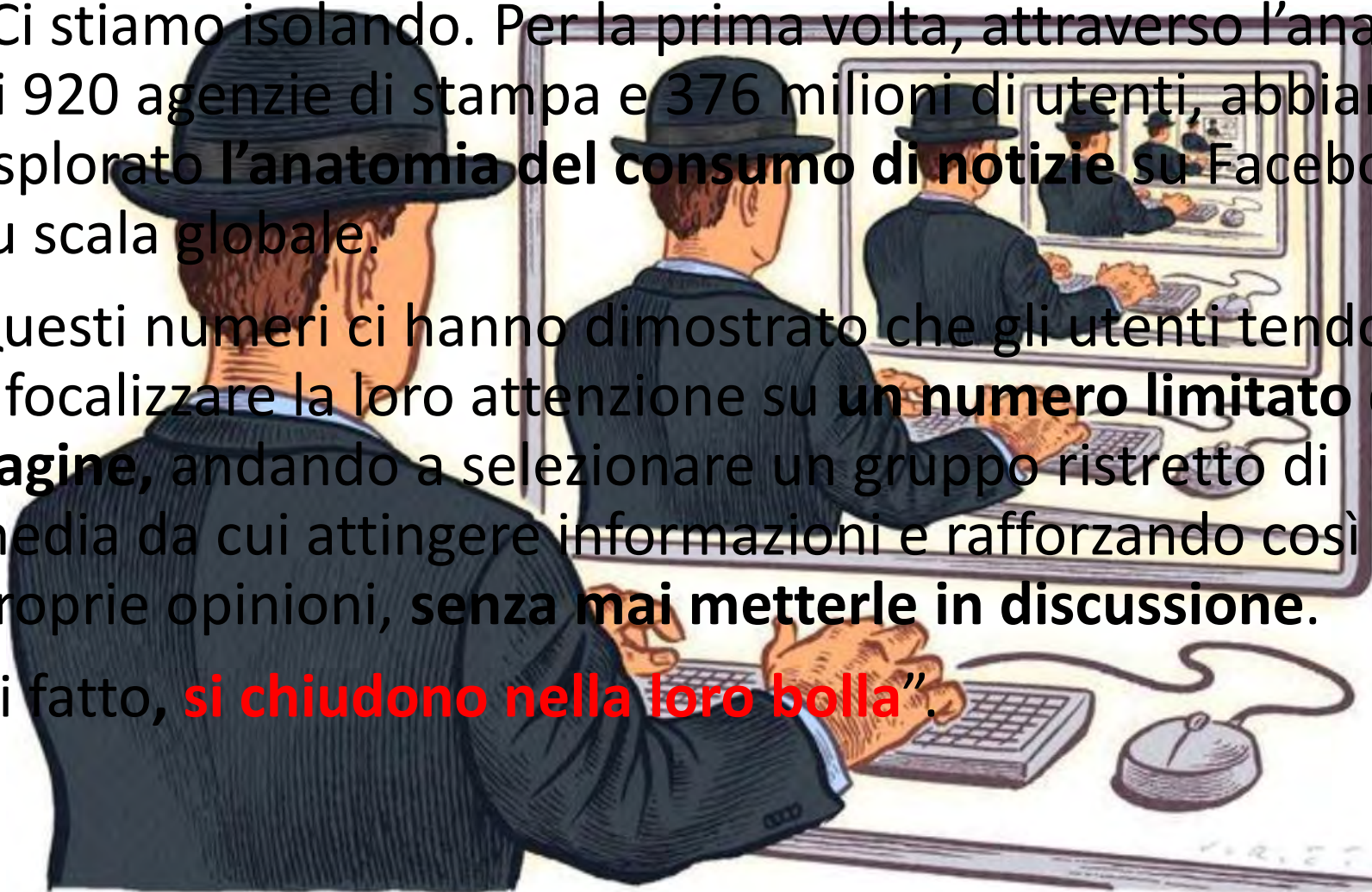
Oxford dictionary – Post-truth definition

Relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief

Indica perciò una strategia retorico-persuasiva in cui è prevalente la componente soggettiva e passionale su quella referenziale.

Walter Quattrobricchi

- “Ci stiamo isolando. Per la prima volta, attraverso l’analisi di 920 agenzie di stampa e 376 milioni di utenti, abbiamo esplorato l’anatomia del consumo di notizie su Facebook su scala globale.
- Questi numeri ci hanno dimostrato che gli utenti tendono a focalizzare la loro attenzione su un numero limitato di pagine, andando a selezionare un gruppo ristretto di media da cui attingere informazioni e rafforzando così le proprie opinioni, senza mai metterle in discussione.
- Di fatto, **si chiudono nella loro bolla”**.

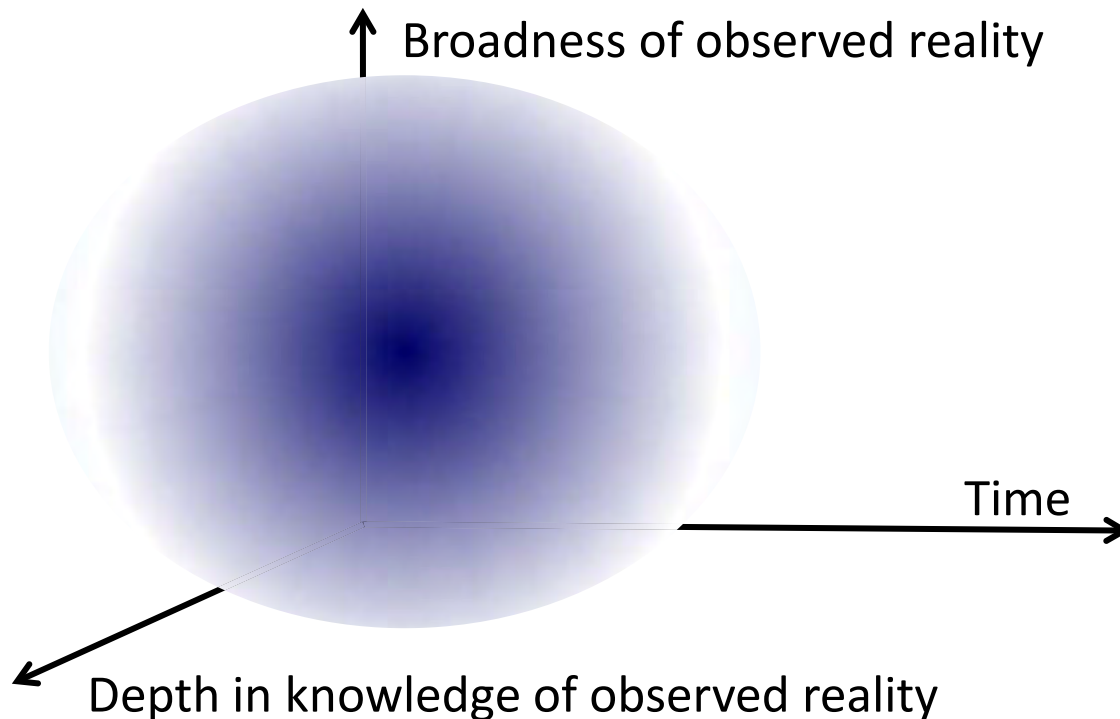


Lorusso, semiologa

- I media non rappresentano un reale già fatto, che sta da qualche parte là fuori
- I media **costruiscono** il reale, **lo modellano**.
- Gli spazi mediatici sono luoghi di costruzione del reale perché sono i luoghi in cui elaboriamo i modelli con cui poi classifichiamo il mondo e ci muoviamo in esso.
- «E' vero solo ciò che passa dalla televisione»

L'illusione del fact checking

- In un mondo di verità moltiplicate la possibilità statistica dell'errore cresce a dismisura.
- Se gli agenti della informazione si moltiplicano a dismisura, e possono essere ovunque, allora il controllo diventa infinitamente più difficile.



L'illusione del fact checking

- In un mondo di verità moltiplicate la possibilità statistica dell'errore cresce a dismisura.
 - Se gli agenti della informazione si moltiplicano a dismisura, e possono essere ovunque, allora il controllo diventa infinitamente più difficile.
 - Diventa paradossalmente più difficile il presupposto di una qualunque ricerca della verità, l'accesso alle fonti e la verifica di veridicità delle fonti
- da «l'ha detto il calcolatore» a «l'ha detto la tecnica di machine learning»)

I livelli di verità di Politifact.com



VIRAL IMAGE

Says CNN's Anderson Cooper stood in a ditch during Hurricane Florence to exaggerate water levels.

— *PolitiFact Facebook fact-checks on Tuesday, September 18th, 2018*



A flood of falsehoods



ANDREW GILLUM

Says there is a report "that showed that, should we move to cover more people to a Medicare for All system, we could actually save the system trillions over an extended period of time."

— *PolitiFact Florida on Tuesday, September 18th, 2018*



Could also cost a boatload of money



DONALD TRUMP

Says 3,000 people "did not die" in two hurricanes that hit Puerto Rico. That death toll estimate was "done by Democrats."

— *PolitiFact National on Tuesday, September 18th, 2018*



Not magic. It's science.

In politifact.com la verità è la risultante di

- **Completezza**, nei fatti ci devono essere tutti gli elementi fondamentali
- **Contestualizzazione**, i fatti devono essere messi in relazione con gli elementi cui sono connessi
- **Tenuta**, i fatti devono aver tenuto in conto i punti di vista critici che potrebbero dare un'altra lettura

La verità non è immagine della realtà, ma costruzione, è dunque un processo.

Ancora sulla tenuta

- **Problematizzare** la verità, **vederne** la molteplicità, **relativizzarne** la natura, non significa che tutto va bene, che tutto è verità, ma significa recuperare il concetto di prova: non tutte le verità hanno la stessa tenuta.
- I discorsi hanno carriere, formazioni, forza, autorevolezza, raggi di applicazione. Guardare a questi elementi significa «provare» la tenuta dei discorsi

La verità, i fatti, e l'etica

- La verità si avvicina mettendo alla prova le diverse verità espresse in ragione di interessi diversi, e via via facendo la realtà.
- Tenendo conto della intersoggettività, e in questo senso dell'etica della interpretazione e della comunicazione.

I fatti sono cocciuti

Caso Kroll, assolto Tronchetti Provera

Cade la ricettazione: «Non costituisce reato». L'ex n.1 Telecom: «Io cocciuto come i fatti»

Condanna in Tribunale a 20 mesi nel 2013, assoluzione in Appello nel 2015, assoluzione annullata in Cassazione nel 2016: nell'annoso ping-pong di sentenze sulla ricettazione collocata nel 2004 e imputata nel 2010 dal pm Alfredo Robledo all'ex presidente di Telecom Marco Tronchetti Provera, ieri il nuovo processo d'Appello torna a far cadere la palla nel campo dell'assoluzione di Tronchetti, «perché il fatto non costituisce reato».

«Dopo tanti anni credo sia

investigativa Kroll e compromessi lo spionaggio illegale di Kroll ai danni di Tronchetti per conto dei rivali brasiliani; la riunione tra Tavaroli, Tronchetti e gli avvocati Chiappetta e Mucciarelli su come utilizzare quei dati; «l'espressa autorizzazione di Tronchetti all'utilizzo» nel modo suggerito da Tavaroli (fingere di far arrivare dal Brasile alla segreteria di Tronchetti un cd rom anonimo); e l'avviso di Tronchetti alla segreteria circa l'arrivo dal

Brasile di un plico importante.

Tronchetti ha però sempre affermato di aver solo dato mandato ai legali (che l'hanno confermato) di esaminare il materiale e di andare a fare denuncia in Procura. Il pg Felice Isnardi e l'avvocato di parte civile Lucio Lucia obiettavano che in realtà nessuna denuncia fosse poi stata presentata né da Tronchetti né da Telecom, né a una Procura in Italia né in Brasile, dove un uomo di Tavaroli andò solo a consegna-

re alla polizia il cd anonimo. «Non vedo che differenza ci sia tra "denuncia" e "consegna alla polizia", ma comunque Tronchetti non c'entra», ribatte ieri l'avvocato Deluca in una arringa mai interrotta dai giudici Paparella-Cairati-Brambilla neanche di fronte all'attribuzione alle controparti di dire «fesserie», «scempiaggini» e cose «ai limiti della buona fede». Tronchetti, riassume Deluca, «non era un esperto di diritto ma la sua chiara indica-

zione fu fare denuncia: il prof. Mucciarelli ha rivendicato l'indicazione di farla in Brasile, e l'esecuzione di questo suo parere fu poi affidata all'ufficio legale Telecom» (il cui capo Chiappetta in aula ha detto di nulla sapere) «e alla Security che, come mostrato dalle sentenze su Tavaroli (3 anni e 8 mesi) «aveva autonomia».

Ma soprattutto, argomenta Deluca, «fare denuncia, cioè rivolgersi alla difesa pubblica per non arrendersi a una ag-

non costituisce reato».

«Dopo tanti anni credo sia un giorno molto importante di chiarezza, facile dire che avendo fiducia nella giustizia alla fine si ha giustizia», commenta Tronchetti accanto ai legali Marco Deluca e Giuseppe Lombardi, ringraziati «per la passione con cui hanno lavorato. Ho sempre avuto fiducia nella giustizia, e ho rinunciato alla prescrizione nella convinzione che questo è un Paese democratico in cui bisogna rispettare lo Stato e la giustizia, il pilastro su cui si fonda una società democratica. I fatti sono cocciuti e io sono cocciuto quanto i fatti».

In attesa delle motivazioni, la formula assolutoria sembra far ritenere decisiva la ritenuta assenza dell'elemento psicologico della ricettazione, cioè del dolo specifico del profitto. La Cassazione, annullando la prima assoluzione, aveva infatti ritenuto «incensurabilmente accertato la consapevo-

Umberto Eco



- Non c'è dunque una verifica che basti da sola ad assicurare la veridicità di qualcosa.
- Al massimo ci possono essere verifiche incrociate che possono escludere interpretazioni false o errate.

Ciò è curiosamente simile a ciò che accade in Informatica – la congettura di Dijkstra

- Eseguendo un programma, non si può mai provare che è **corretto**, si possono solo trovare nuovi dati in ingresso al programma che dimostrano **l'esistenza di errori**.

Ma sta accadendo qualcosa di più

- Sempre di più la costruzione del senso si dà per via **narrativa**.
- La notizia è sempre più pensata come documento e sempre più come **racconto**.
- A prevalere non è un criterio di **attendibilità**, ma di **efficacia narrativa**, di **credibilità**.

Fatti, favole, fole, bugie

- La forza dei nuovi soggetti di informazione si misura più sulla capacità di riuso di frame narrativi consolidati che sulla attendibilità della informazione.
- Parrà esagerato, ma la dinamica è la stessa. Quando leggo una favola, io ho delle esigenze che mi fanno apprezzare quella favola, e fanno sì che ci creda e mi appassioni.

La parcellizzazione della verità

- L'impressione è che le verità si siano più che altro parcellizzate, e nel parcellizzarsi si sono moltiplicate, in un gioco di specchi e rimbalzi.
- Attraverso la condivisione sempre più estesa hanno acquisito credibilità
- La logica di Facebook rafforza la idea che le condivisioni siano indice di identificazione e adesione, dunque di credibilità.

Repubblica,
22 settembre 2018

Altan

SE AVETE CREDUTO
ALLE PROMESSE, VUOL
DIRE CHE ERANO CREDIBILI.



I soggetti della informazione siamo noi

- Oggi i soggetti della informazione siamo noi, persone comuni, dalle competenze comuni, il più delle volte con una nostra esperienza da condividere, con un nodo di emozioni da esprimere.
- Umberto Eco – «I social media danno diritto di parola a legioni di imbecilli che prima parlavano solo al bar dopo un bicchiere di vino, senza danneggiare la collettività. Venivano subito messi a tacere, mentre ora hanno lo stesso diritto di parola di un Premio Nobel»

Il legame sociale

- In un mondo di verità moltiplicate, assolutizzate l'una rispetto all'altra, e isolate entro comunità chiuse, la cosa più seria che si perde non è il vero ma il legame sociale.
- Saper discriminare la verità significa condividere saperi.
- E condividere saperi significa essere parte della stessa comunità.

**Maurizio
Ferraris**

L'ideologia che anima
la postverità è
l'atomismo di milioni
di persone convinte di
aver ragione non insieme
(come credevano,
sbagliando, le chiese
ideologiche del secolo
scorso) ma da sole.

POSTVERITÀ E ALTRI ENIGMI



il Mulino

VOCI

Quanta verità c'è nella postverità?

Anche se è forte la tentazione di dire che le bufale sono sempre esistite, che la menzogna è un ingrediente imprescindibile della politica e della vita, e che dunque non c'è niente di nuovo sotto il sole, anche se viene voglia di tagliar corto dicendo che si tratta tutt'al più di fare attenzione a quel che si legge così come si fa attenzione a quel che si mangia e si beve, la postverità è un concetto filosoficamente rilevante e la sua emergenza definisce una caratteristica essenziale del mondo contemporaneo: l'alleanza tra la potenza modernissima del web e il più antico desiderio umano, quello di aver ragione a tutti i costi.

Le emozioni nel tempo della datacy

La rabbia nella storia

Peter Sloterdijk
Ira e tempo



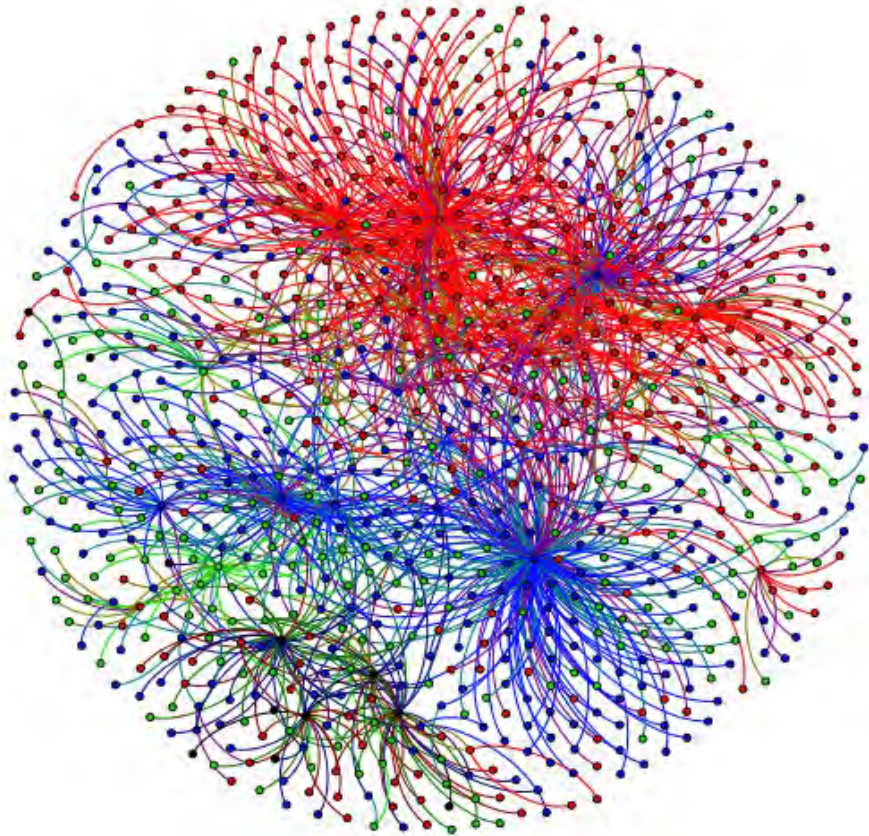
Biblioteca Meltemi

La rabbia è un sentimento insopprimibile, che attraversa tutte le società, alimentato da coloro i quali, a torto o a ragione, ritengono di essere esclusi, discriminati o poco ascoltati.

Storia della rabbia

- Secondo Sloterdijk, storicamente in occidente è stata la Chiesa a dare uno sbocco a questa **enorme accumulazione di rabbia** e poi, a partire dalla fine dell'ottocento, i partiti della sinistra.
- Che hanno svolto, secondo Sloterdijk, la funzione di **“banche della collera”**, **accumulando energie** che, anziché essere spese nel momento, potevano essere investite per costruire un progetto più grande.

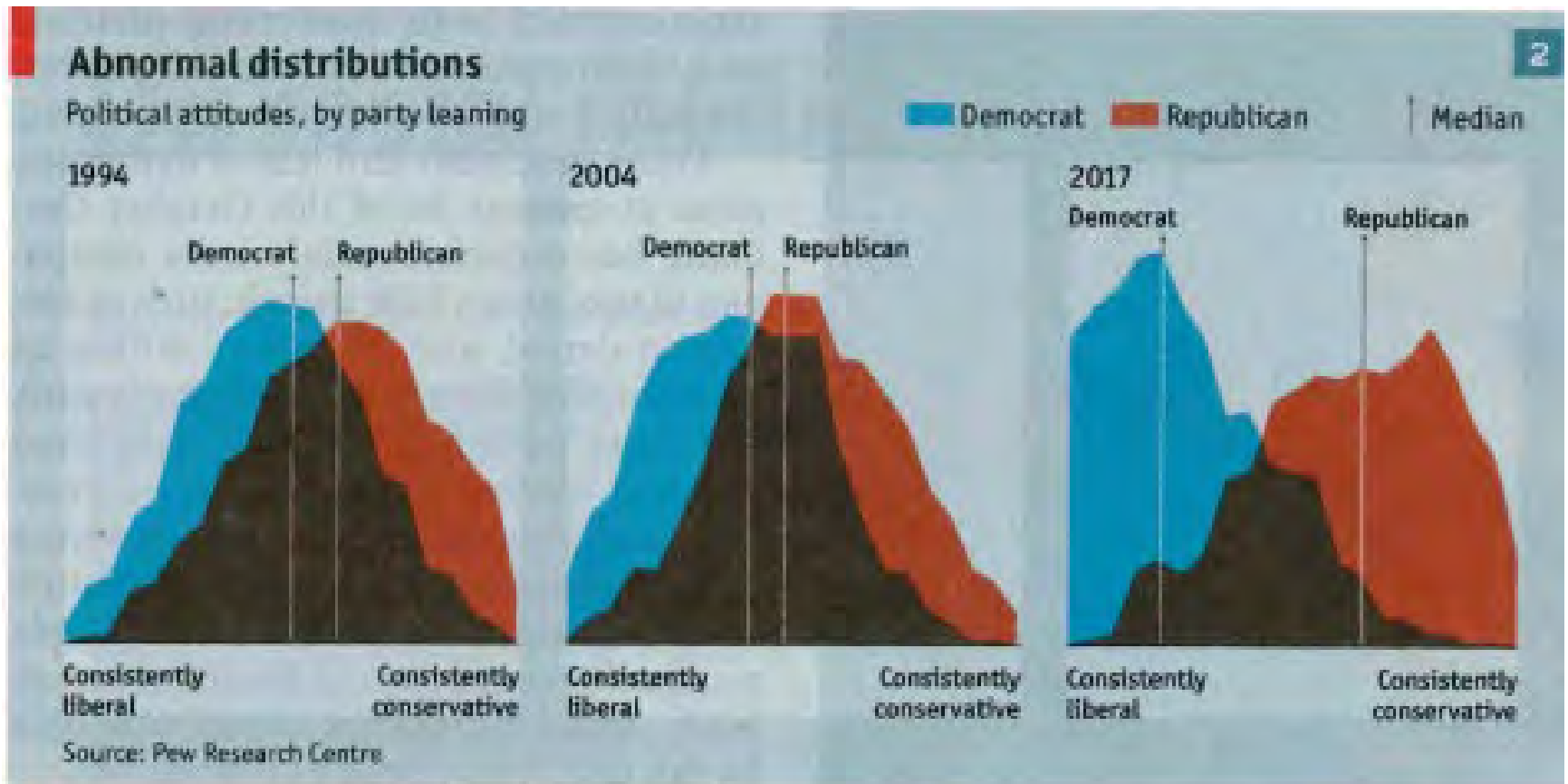
Anger is more popular than joy...



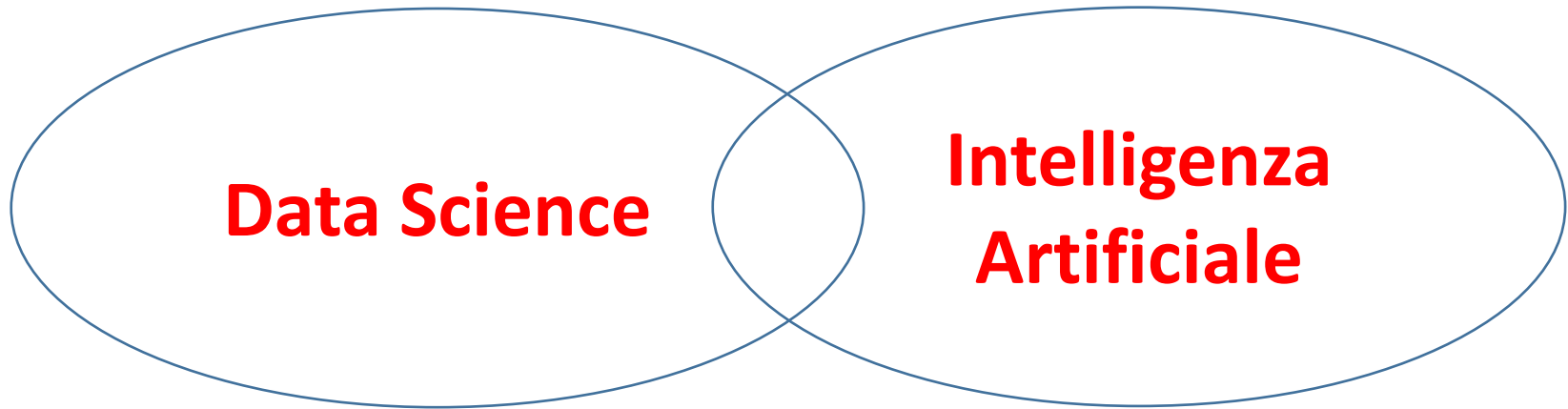
- **red** stands for anger,
- **green** represents joy,
- **blue** stands for sadness
- **black** represents disgust.

The regions of same color indicate that closely connected nodes share the same sentiment.

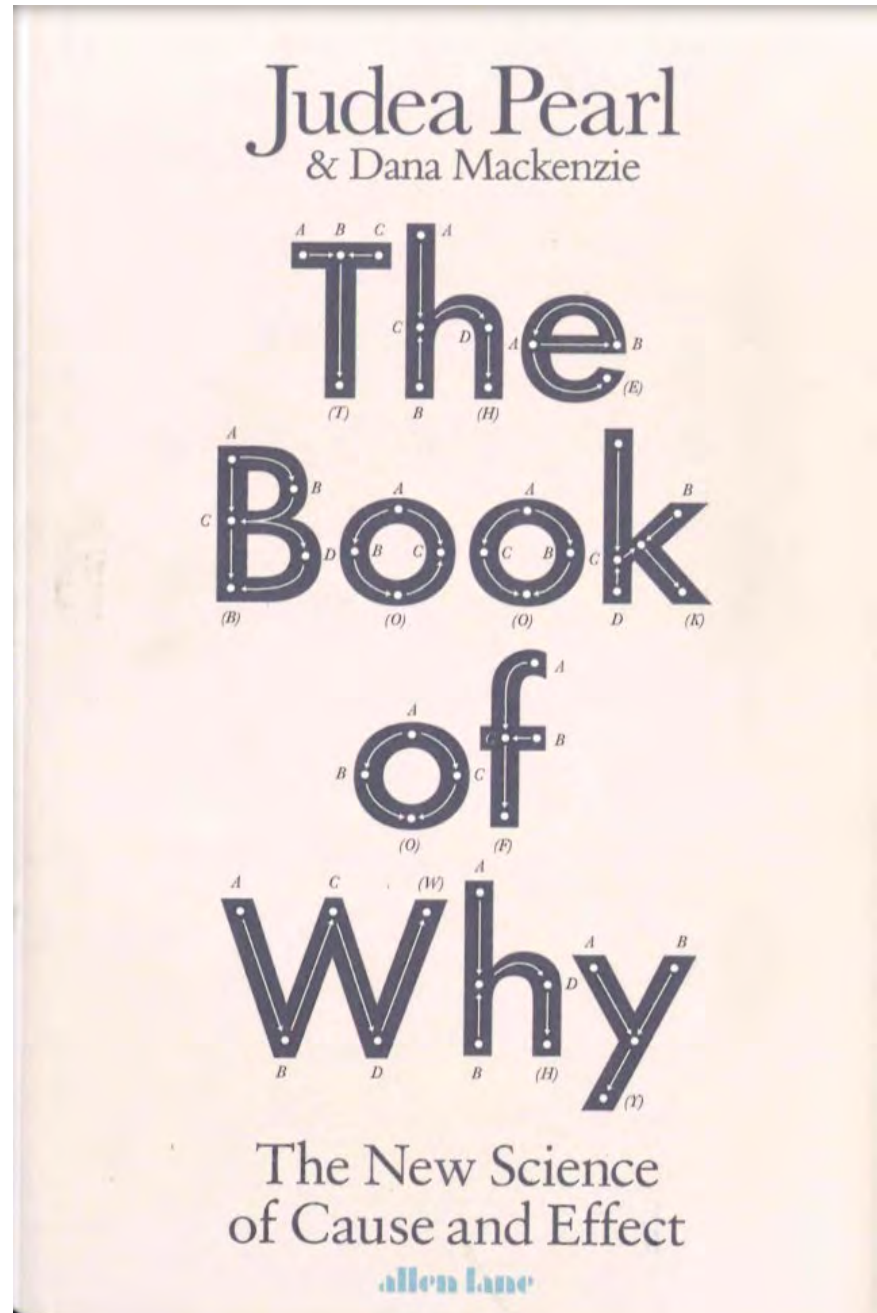
Polarizzazione delle opinioni politiche nella popolazione USA - Economist 4/11/2017



Cosa **non e'** la datacy?



Exploring the World of causal inference



The difference between What and Why in the Genesis

- God asks Eve: **What** is this you have done?
- She replies: (**Because**) the serpent deceived me, and I ate

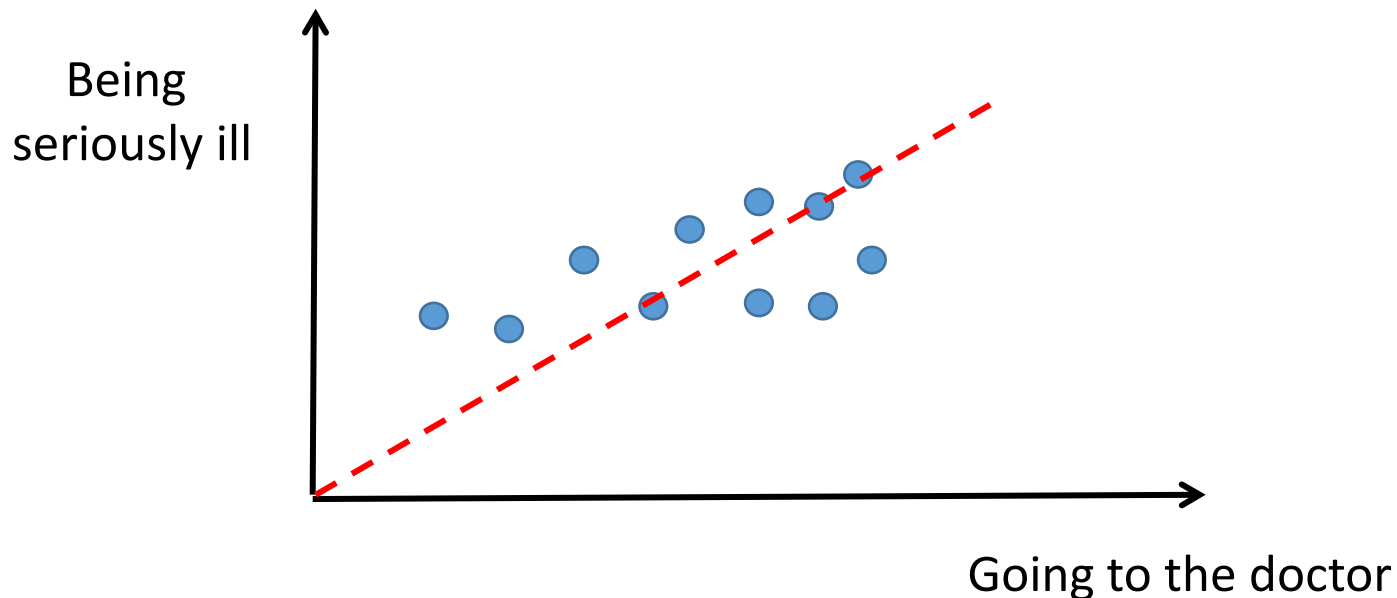
Theory is alive!

- It is a false believe that the answers to all scientific questions reside in the data, to be unveiled through clever data-mining tricks.
- Data can tell you that the people who took the medicine recovered faster than those who did not take, but they cannot tell you **why**.

Let us look in another way

Conditional probability evaluated on past observations

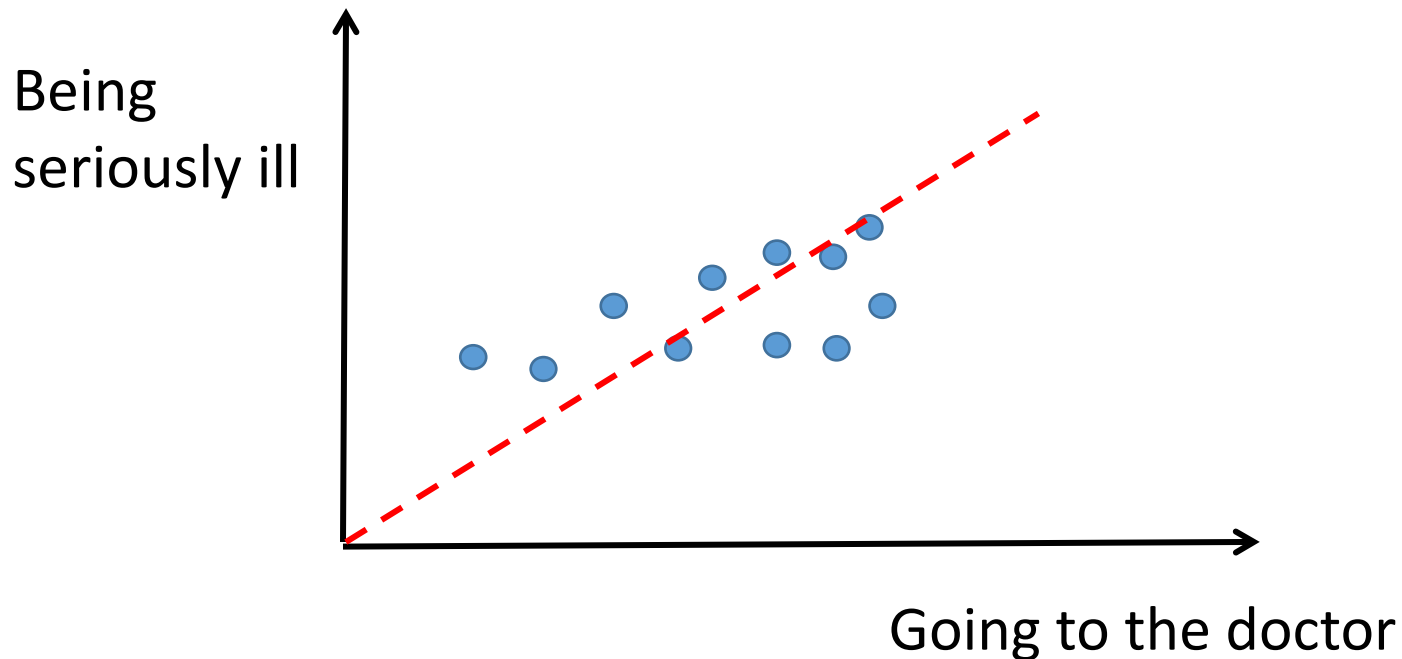
P (being seriously ill/going to the doctor) is measured in statistics through correlation.



Let us look in another way

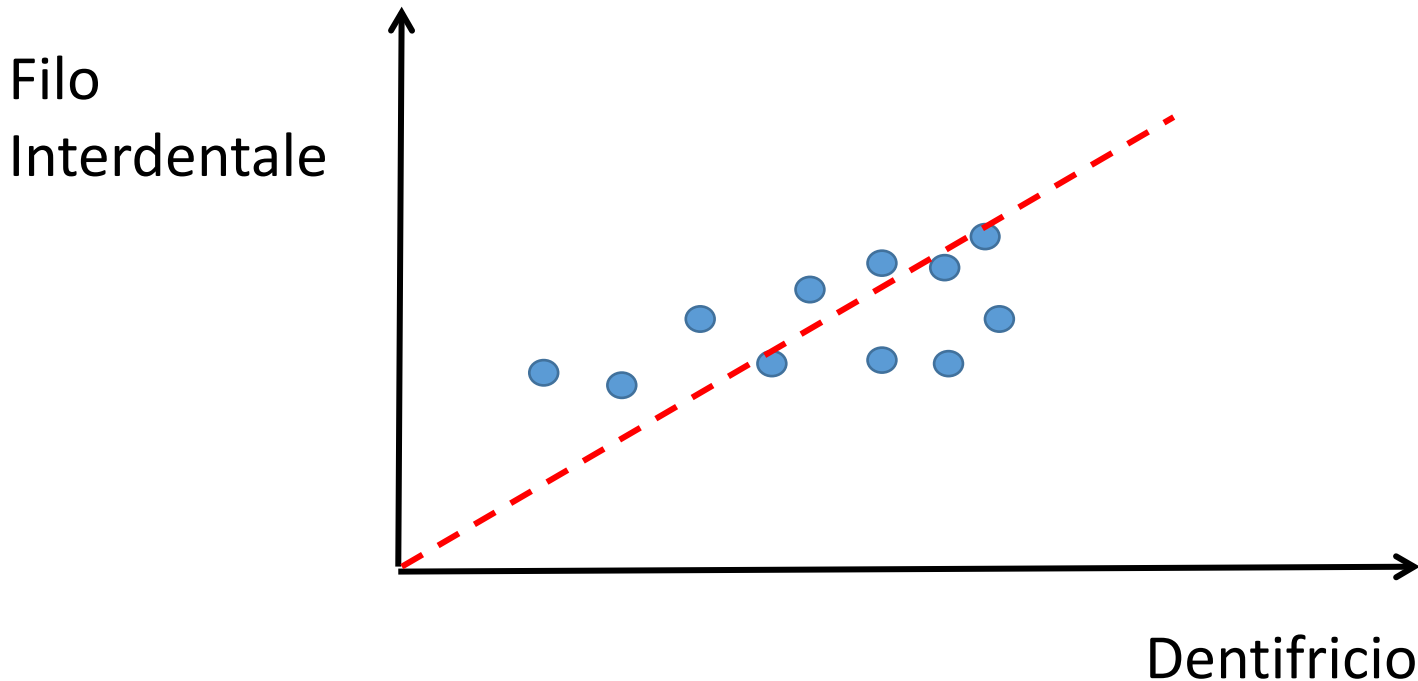
But this correlation has no direction!

So, paradoxically, patients would avoid going to the doctor to reduce the probability of being seriously ill



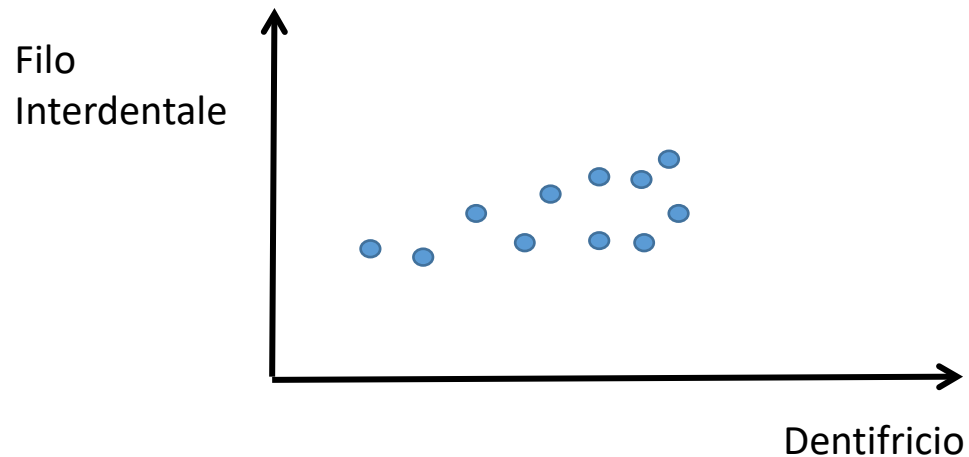
Floss vst toothpaste

Some associations may have causal interpretations, other may not. But statistics alone cannot tell which is the cause and which is the effect, toothpaste or floss



Floss vst toothpaste

Some associations may have causal interpretations, other may not. But statistics alone cannot tell which is the cause and which is the effect, toothpaste or floss



Seeing vs doing

In machine learning we make use of the standard conditional probability used in statistics $P(L/D)$, where

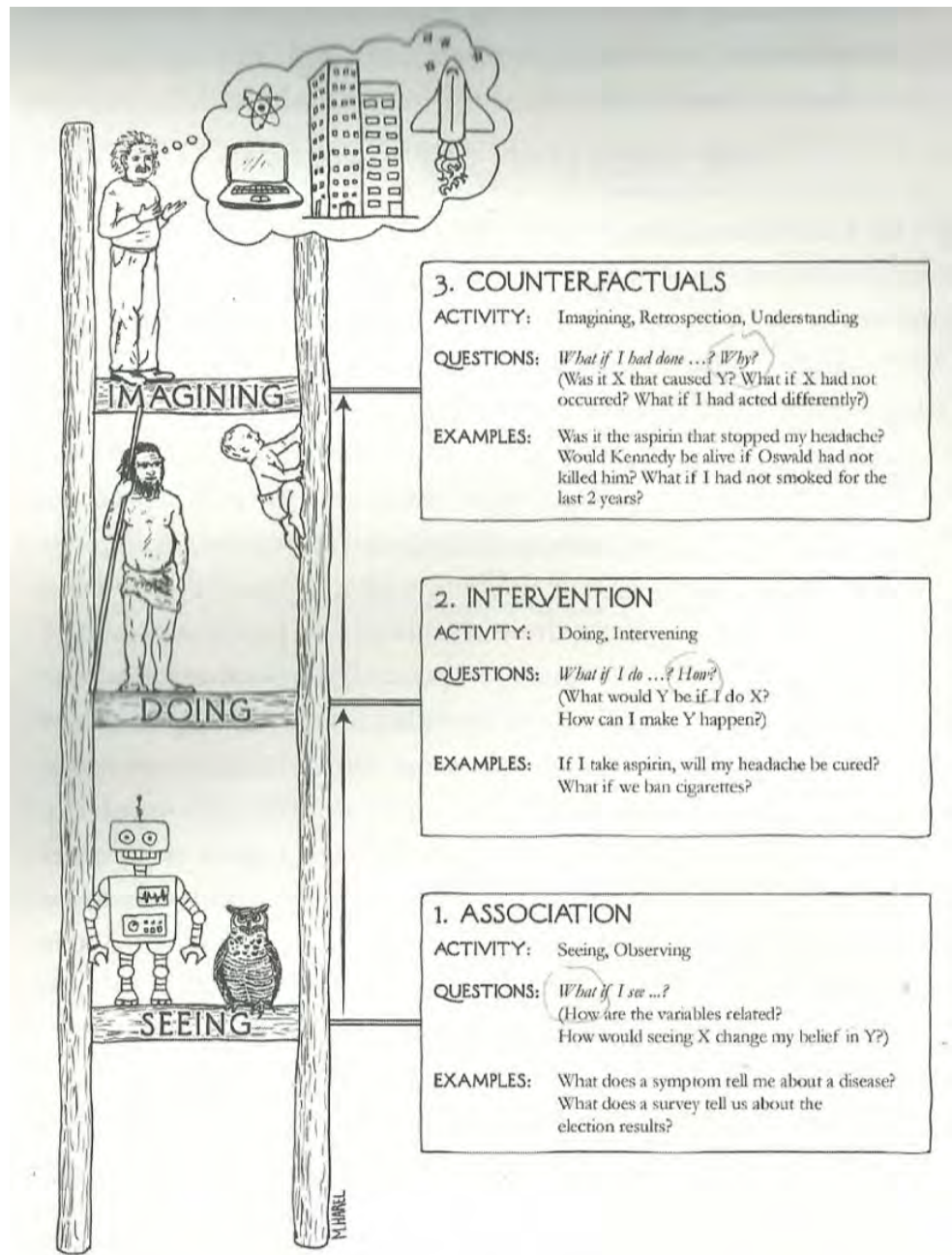
- L is e.g. the lifespan of patients
- D is the event of taking a drug

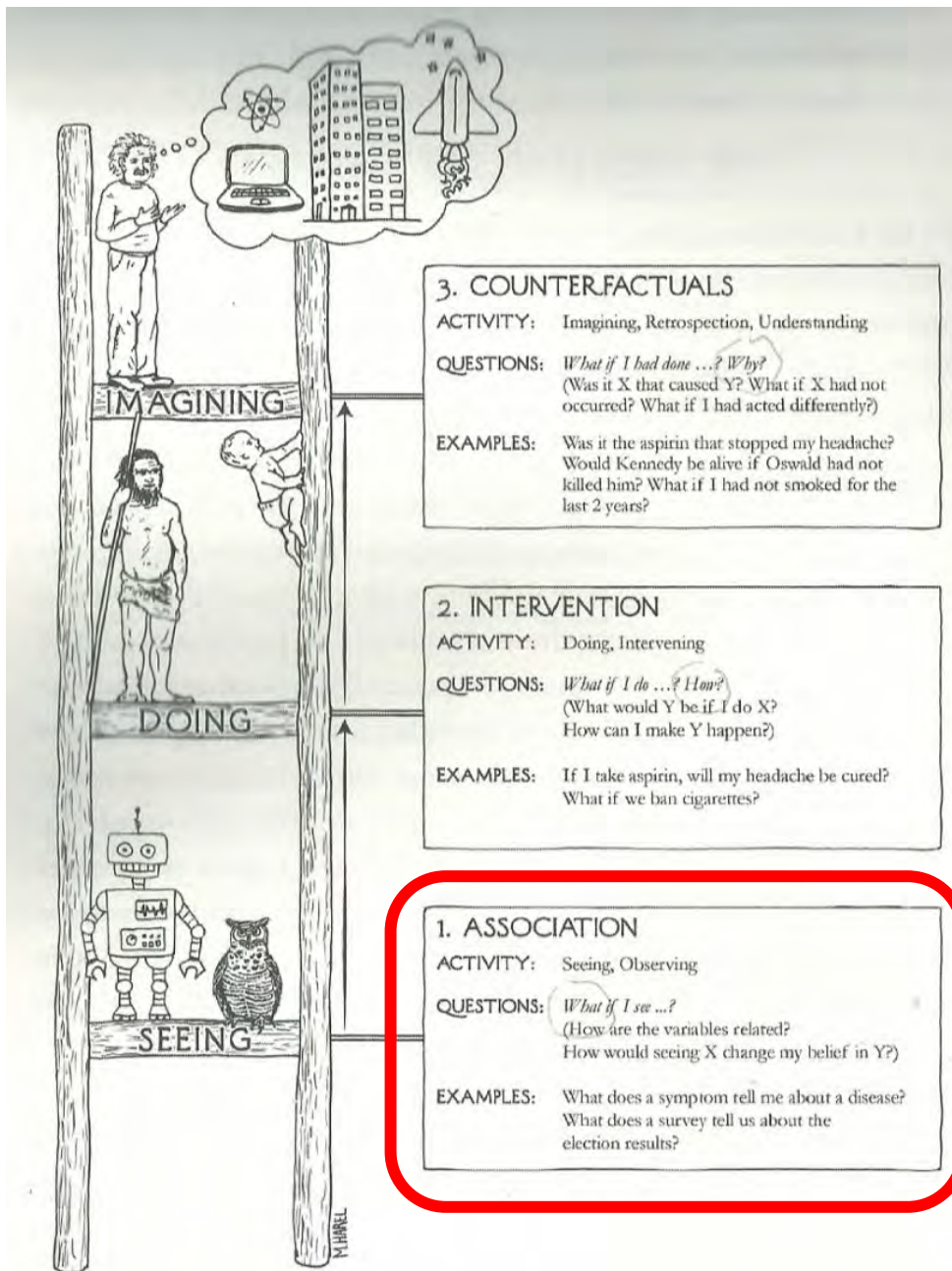
This expression stands for the probability P of lifespan, conditional on **seeing** the patient take drug D.

But $P(L/D)$ may be totally different from $P(L/\text{do}(D))$, the probability that a typical patient would survive L years if made to take the drug D.

This corresponds to what we measure in a clinical trial.

The ladder of Causation and the World of Causal inference





The world of **what**

Predictions based on passive observations

Activity: **seeing**, observing

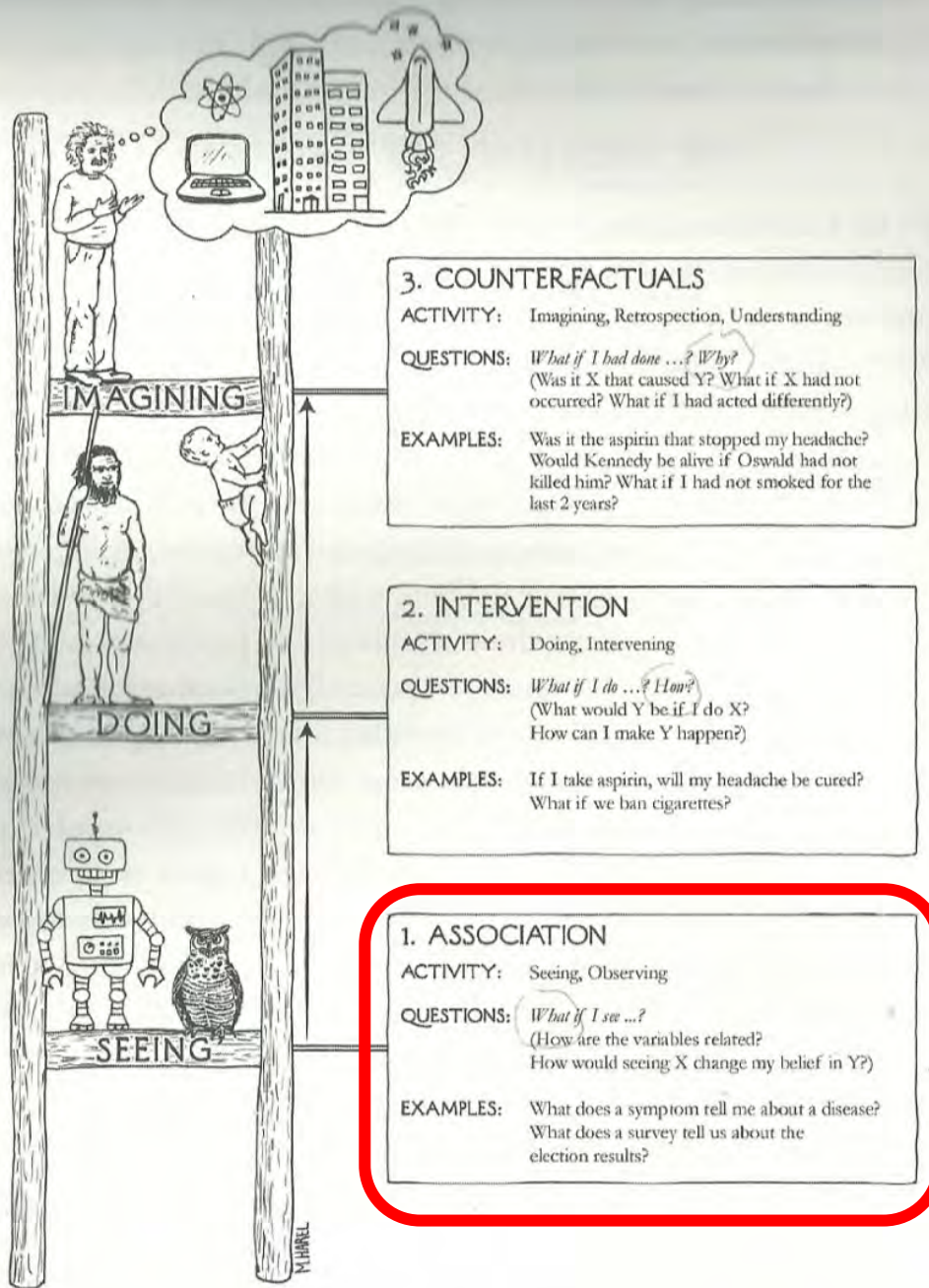
Question: what if I see ..

Examples

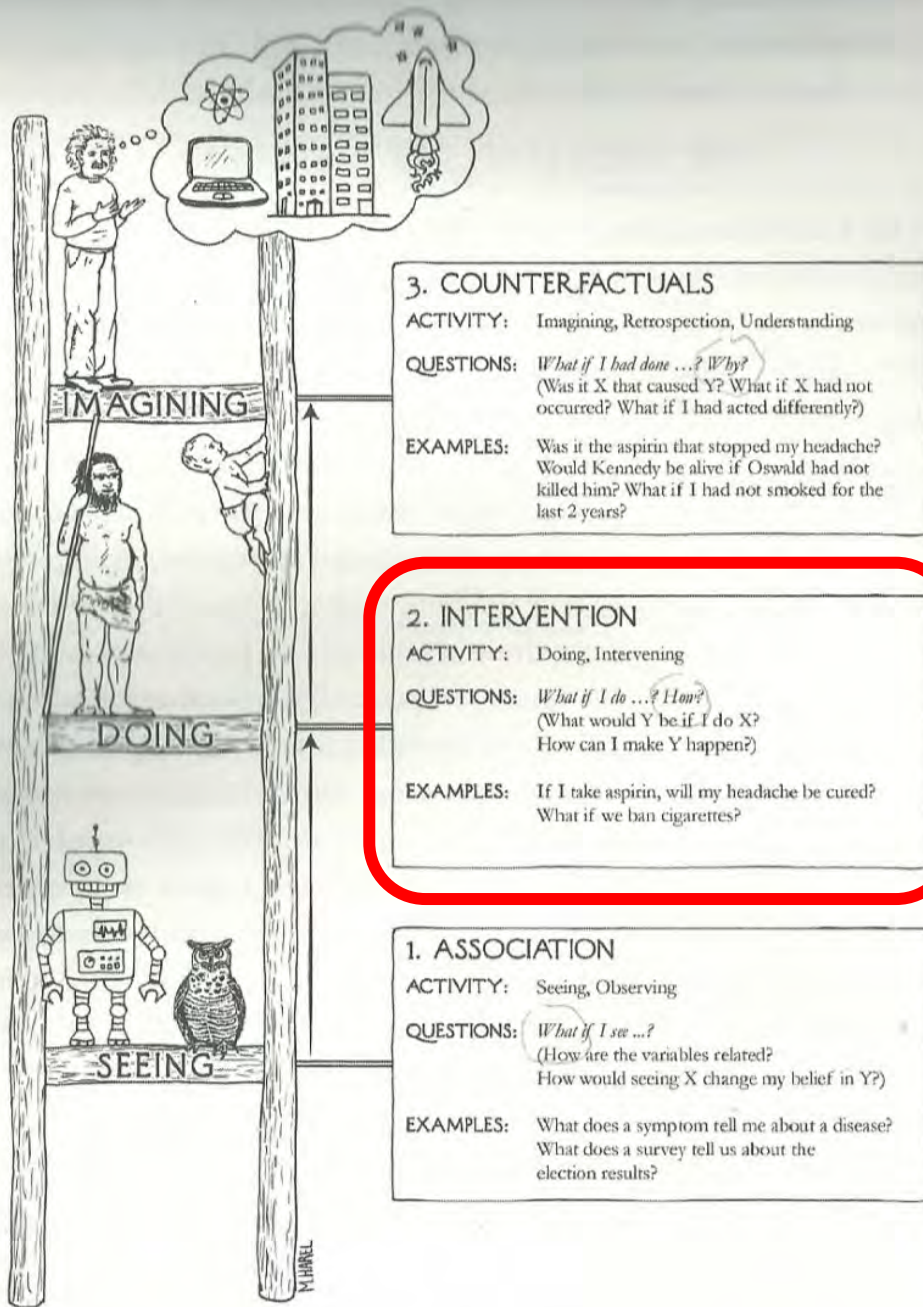
- What does symptom tell me about a disease?
- What does a survey tell me about the election results?

Limitation: absence of a model of reality

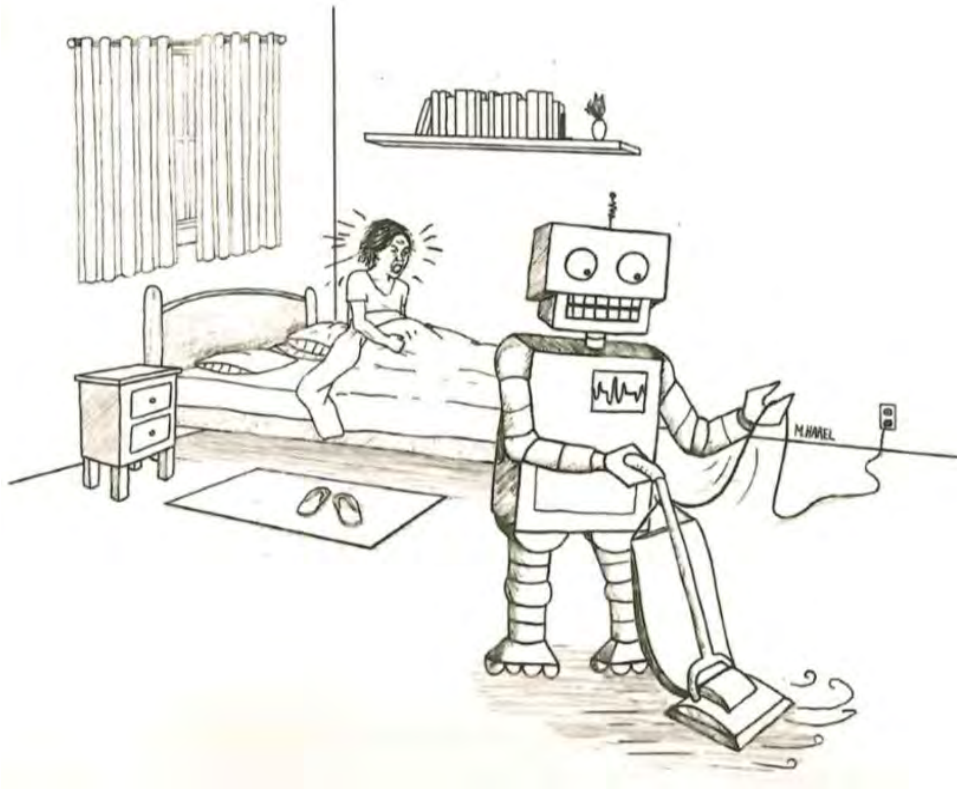
Cultural area: Data Science



- The world of **how**
 - Predictions based on **experiments and a model** of the world
 - Activity: **doing**, intervening
 - Question: what if I see ..
Examples
 - What will happen to our floss sales if we double the price of toothpaste?
 - How I can sell it?
- Cultural area: Strong AI



- The world of **why**
 - Explanations based on **reasoning and a model** of the world
 - Activity: imagining, retrospecting, **reasoning**
 - Question: what if I had done? Why
- Examples
- My headache is gone now, but why?
 - **Counterfactual**: What would have happened if I had not taken the aspirine?
- Cultural area: Strong AI



Example of
counterfactual
question:

You shouldn't have
woken me up!

Questions to be answered

- Have we already made machines that think?
- Can we make machines that think?
- Will we make machines that think?
- Should we make machines that think?
- Can we make machines that are capable of distinguishing good from evil?

L'etica dei dati

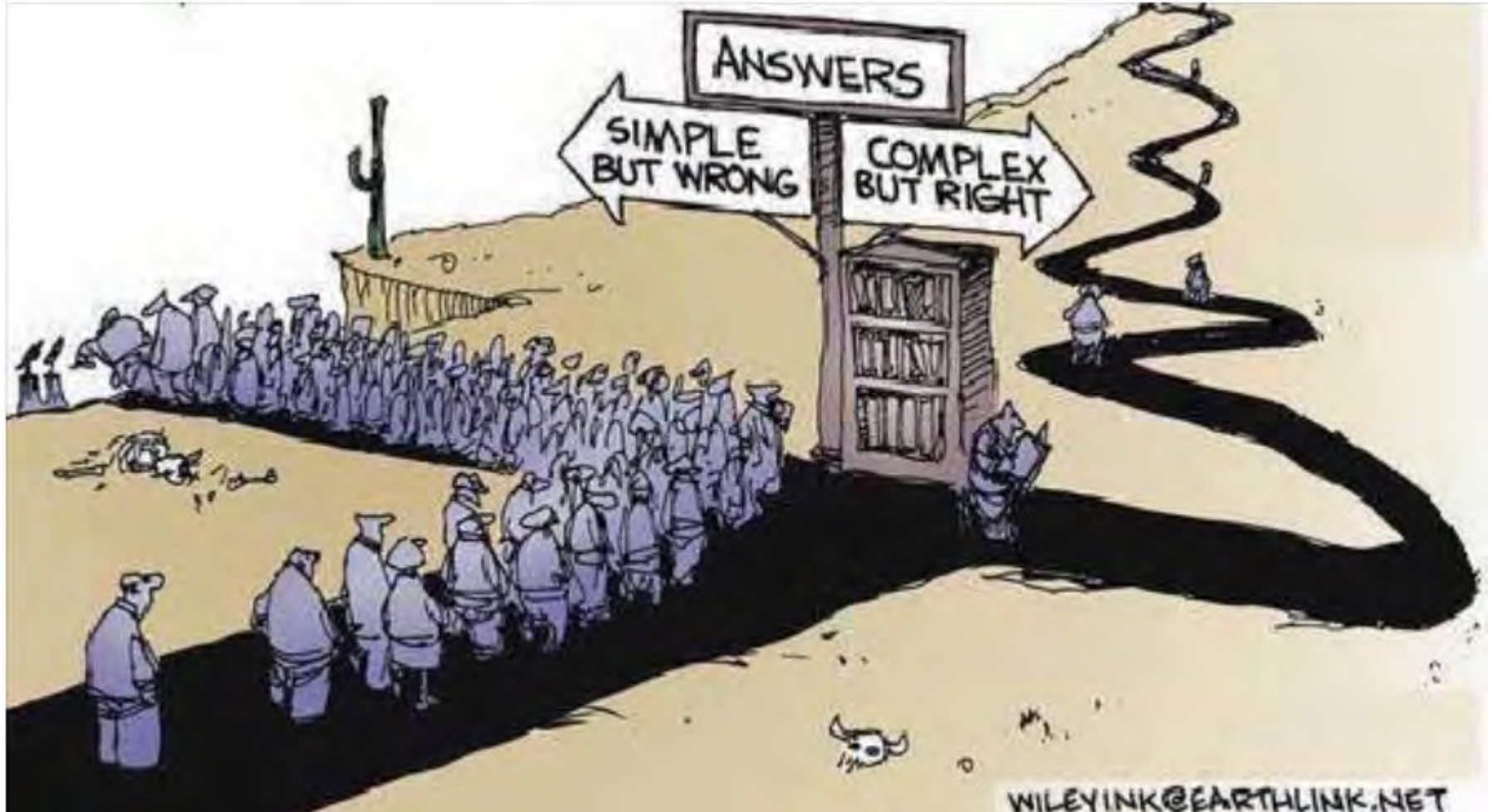
Determinants of Ethics in Digital Data Science

1. **Transparency vs Opacity**, it is easy for others to see from data what actions are performed.
2. **Accountability**, mechanisms are in place to determine who took a responsible action.
3. **Attribution of responsibility**, a process of cognition based on data by which moral account-tability is assigned to a person believed to have produced a disapproved behavior or effect
4. **Auditability**, ability of examining or evaluating (something) thoroughly
5. **Awareness**, knowledge and understanding that something is happening or exists
6. **Data Divide**, economic and social inequality with regard to use of or impact of data
7. **Equality of Opportunity** (Non Discrimination, Egalitarianism), data should enable people to compete on equal terms
8. **Explanation**, a statement made to clarify something and make it understandable
9. **Fairness**, the state, condition, or quality of being fair, or free from bias or injustice
10. **Fairness-aware (learning techniques)**, classifiers to be independent of sensitive features.
11. **Generalization vs Personalization**, data for all vs tailoring data to accommodate specific individuals, sometimes tied to groups or segments of individuals
12. **Liability**, extends responsibility further to the area of laws.
13. **Objectivity**, state of being true even outside a subject's individual biases, interpretations and feelings
14. **Quality of information, (Accuracy, Veridicity etc.)** property of information to be adherent to reality
15. **Responsibility**, you accept the potential costs, and duties for the decision you made
16. **Privacy**, the state of personal data of being free from public attention
17. **Sharing**, ~~dividing~~ having data in common with others
18. **Statistical discrimination**, racial or gender inequality based on stereotypes

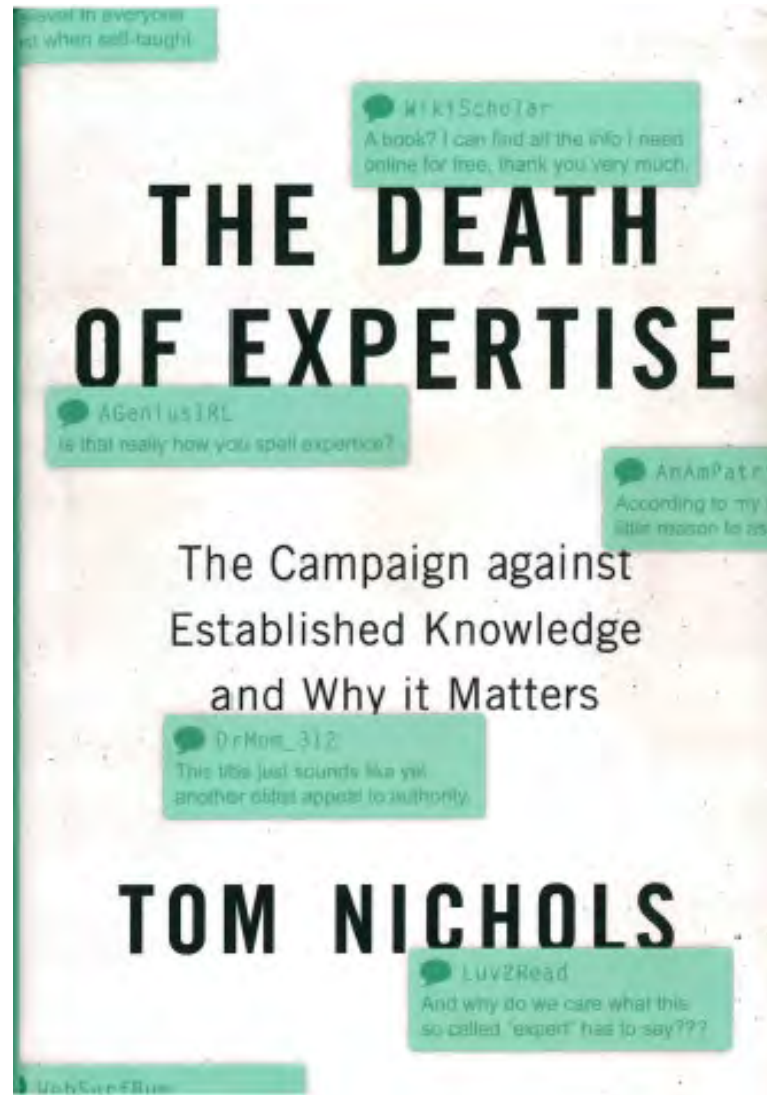
My Mantra...

Datacy! Datacy! Datacy!

No simple answer to complicate questions



La datacy richiede una nuova alfabetizzazione



Da Repubblica dell' 11 settembre 2018

L'INCONTRO



R: Intervista al Dalai Lama: "L'educazione è la cura per la rabbia globale. Rapporti con la Cina? Vedo progressi"

Il leader spirituale tibetano dissipa ogni dubbio sulla sua salute: "Nostra responsabilità è iniziare un lavoro etico, io penso di poterci lavorare altri 10, 20, 30 anni"

Una Definizione tentativa di Datacy

L'insieme dei modelli, metodologie, linguaggi, tecniche e tecnologie e delle loro applicazioni, che permettono di ragionare su ed elaborare un vasto insieme di tipologie di dati digitali, come i dati tabellari, i testi, le immagini, tratti da fonti eterogenee quali il Web, le reti sociali, l'Internet delle cose, essendo in grado di:

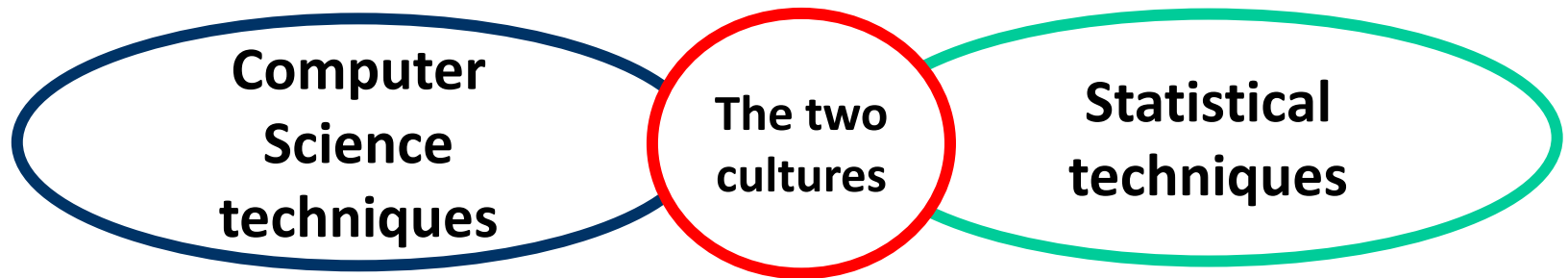
- ricostruirne e modellarne il **significato** attraverso linguaggi concettuali,
- applicare tecniche per la valutazione del livello di **veridicità e qualità**,
- **integrare** dati eterogenei riconciliandone le differenze,
- **interrogare e analizzare** i dati mediante linguaggi e ambienti per estrarne conoscenza

Una Definizione tentativa di Datacy

- applicare tecniche e modelli statistici e basati su **apprendimento** per costruire modelli decisionali, interpretativi e predittivi.
- **risolvere problemi** con il supporto dei dati e prendere decisioni complesse.
- comprendere l'impatto sulla **economia** e sulla **società** del fenomeno dei dati digitali.
- analizzare i **corpi giuridici** sviluppati dalle istituzioni pubbliche in tema di dati digitali
- affrontare i nuovi temi **etici** che nascono dall'uso dei dati digitali.

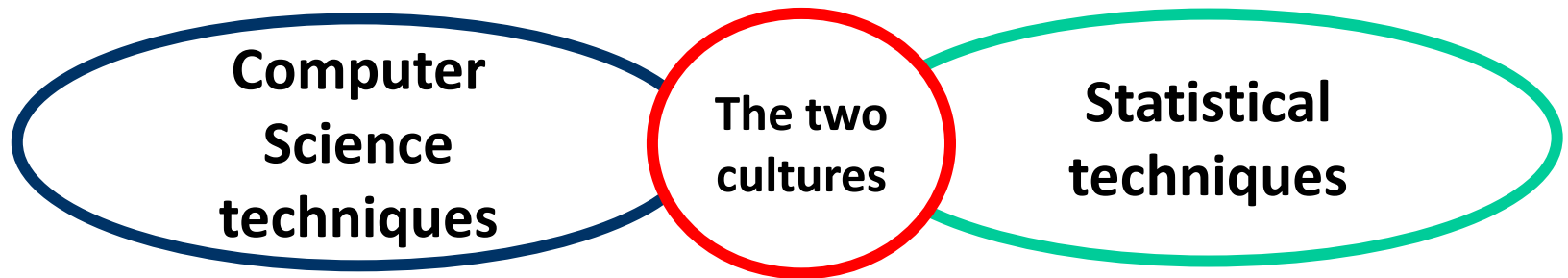
The two cultures

Datascy needs for new statistical
and algorithmic techniques

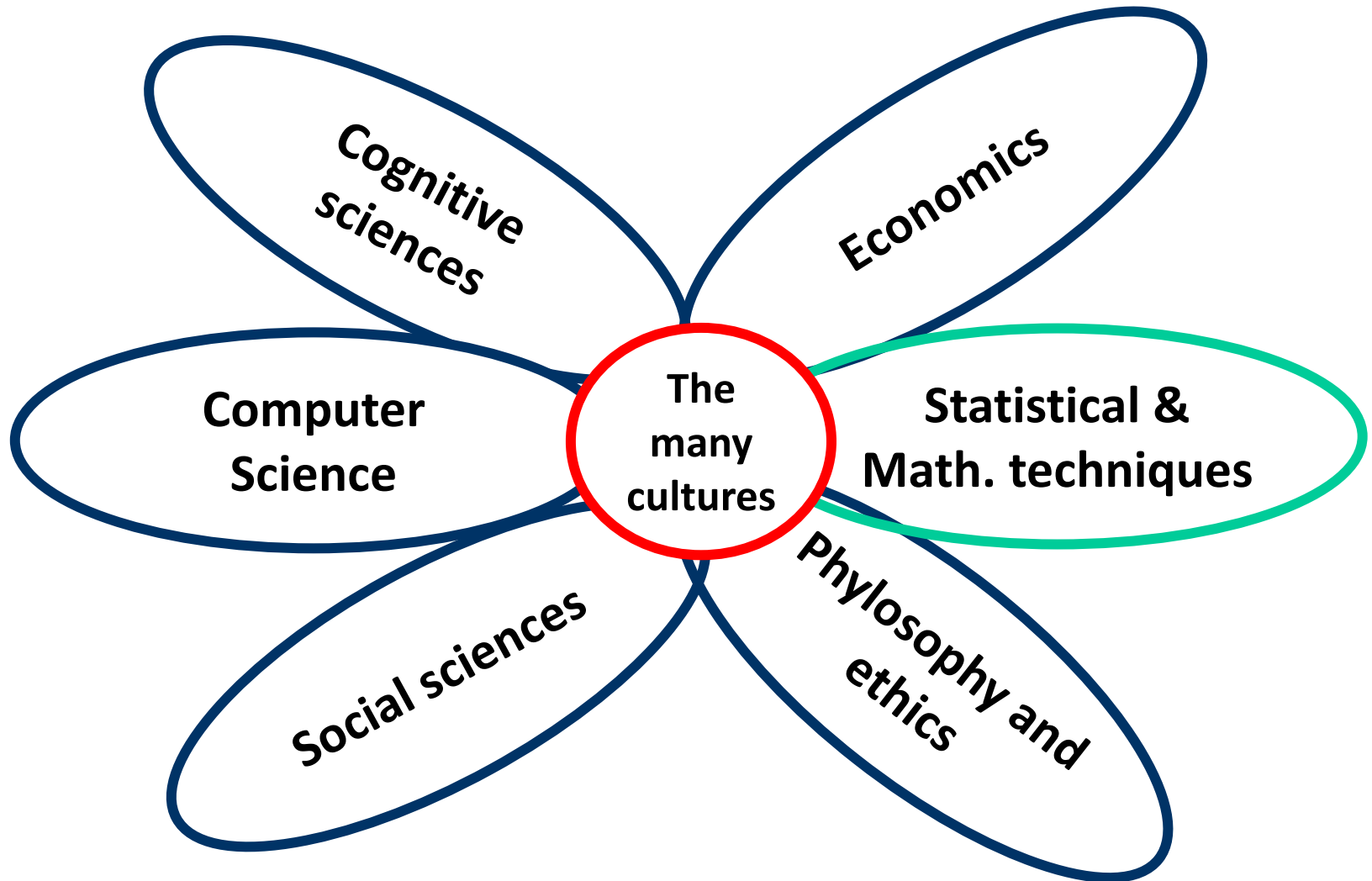


The two cultures

Datascy needs for new statistical
and algorithmic techniques



The many cultures



Riferimenti

- Batini, Carlo – Le Basi dell'Informatica, Editori Riuniti, 1984.
- Bertin, Jacques – Semiology of Graphics, diagrams, networks, maps, Esri Press, 1983.
- Boisot, Max – Information Space, a framework for learning in organizations, institutions and culture, Rutledge 1955.
- Borgman, Christine – Big Data, Little Data, No Data, scholarship in the networked world, the MIT Press, 2015.
- Ferraris, Maurizio – Postverità e altri enigmi, Il Mulino, 2017
- Lorusso, Anna Maria – Postverità, Edizioni Laterza, 2018
- Floridi, L – La quarta rivoluzione, come l'infosfera sta cambiando il mondo, Raffaello Cortina Editore, 2017
- Grefen P., Beyond eBusiness, 2016
- Mayer-Schonberger, Viktor and K. Kukier, Big data, a revolution that will transform how we live, Work and Think, John Murray, 2013
- Mayer-Schonberger, Viktor and Thomas Ramge – Reinventing Capitalism in the age of big data, 2018.
- Mommonier, Mark, How to lie with maps, second edition, The University of Chicago Press
- Moody, D – Measuring the value of information, Proceedings of the European Conference on Information Systems, 1999.
- Nichols, Tom – The death of expertise, Oxford University Press, 2016.
- O'Neill, C. Principles of mass destruction, 2017.
- Pearl, Judea and Dana Mackenzie – The Book of Why, the new Science of Cause and Effect, Penguin Random House, 2018.

Riferimenti

- Pasquale, Frank – The Black Box Society – Harvard University Press, 2015
- Quattrocioni W., Misinformation, Franco Angeli, 2016
- Rifkin, J. – The Zero Marginal Cost Society – Palgrave Macmillan, 2014.
- Sadin, Eric – La siliconizzazione del mondo, Einaudi 2018.
- Searle, John – La costruzione della realtà sociale, Einaudi 2006.
- Searle, John – Mente, Linguaggio, Società, Raffaello Cortina Editore 2000.
- Shapiro, C and H. Varian – Information Rules, Harvard Business Review Press, 1999.
- Sloterdijk, Peter – Ira e tempo, saggio politico psicologico, Molteni Editore, 2007.
- Steyn, Jacques – A critique of the claims about mobile phones and Kerala fisherman: the importance of the context of complex social systems, EJISDC, 2016.
- Tufte, Edward – The Visual Display of Quantitative information, Graphic Press, 2001.
- Waackman, Judy - Pressed for time: the acceleration of life in digital capitalism, Univ. of Chicago Press, 2016.
- World Economic Forum: Future of jobs Survey, 2018
- Porche III, I.R. Emerging Cyber Threats and Implications, RAND Office of External Affairs
Tinker, tailor, hacker, spy – Economist, Novembre 2016.

Fine

Da sistemare

Contents

Perché serve una nuova categoria, la datacy

Cosa è la datacy?

- I dati (e i concerns che generano) crescono esponenzialmente, le soluzioni della ricerca e della regolazione linearmente....
- Le tecnologie ICT
- Beni, dati, servizi
- I lavori della datacy
- Un framework per i big data
- La linea di confine tra analogico e digitale si sposta continuamente: numerazione, simbolizzazione, codifica, astrazione
- Dare ai dati un significato
- Integrare i dati aumenta il loro valore
- Codifica, astrazione, scarsità
- Da una economia money-driven a una economia data-driven
- Il valore sociale dei dati
- Dal digital divide al data divide
- Dati, simboli, numeri, visualizzazioni: il quantitativo/oggettivo e la percezione soggettiva
- Dalla correlazione, ai sistemi basati su regole, ai modelli probabilistici di apprendimento: il cosa e il perché
- Dalla verità alla post-verità
- Le emozioni nel tempo dei dati digitali
- Il diritto alla spiegazione nel GDPR
- La fine della expertise
- L'etica dei dati

Cosa non è la datacy

- Il cosa, il come, il perchè

Una definizione di datacy

- Non ci sono risposte semplici a domande complesse
- La datacy richiede una nuova alfabetizzazione

Riferimenti

Cittadini ai tempi di Internet

Per una matura cittadinanza nell'era digitale

Alfonso Fuggetta

Tuesday, 11 September 2018

Contents

Perché serve una nuova categoria, la datacy (o cultura del dato)?

- I dati (e i concerns che generano) crescono esponenzialmente, le soluzioni della ricerca e della regolazione linearmente....
- Le tecnologie ICT
- Beni, dati, servizi e lavori
- Un framework per i big data
- La linea di confine tra analogico e digitale si sposta continuamente: numerazione, simbolizzazione, codifica, astrazione
- Un mondo di informazioni: codifica, astrazione, scarsità
- Da una economia money-driven a una economia data-driven
- Il valore sociale dei dati
- Dal digital divide al data divide
- Dati, simboli, numeri, visualizzazioni: il quantitativo/oggettivo e la percezione soggettiva
- Dalla correlazione, ai sistemi basati su regole, ai modelli probabilistici di apprendimento: il cosa e il perché
- Dalla verità alla post-verità
- Il diritto alla spiegazione nel GDPR
- La fine della expertise
- L'etica dei dati

Cosa è dunque la datacy?

- Una definizione provvisoria di datacy
- Non ci sono risposte semplici a domande complesse
- La datacy richiede una nuova alfabetizzazione

Riferimenti

La linea di confine tra analogico e digitale
si sposta continuamente: numerazione, simbolizzazione,
codifica, astrazione

I dati (e i concerns che generano) crescono esponenzialmente,
le soluzioni della ricerca e della regolazione linearmente....

NYT - 4 settembre 2018

- Facebook, Twitter and Google executives have been invited to testify in Washington on Wednesday about foreign influence campaigns and disinformation online ahead of the midterm elections in November.
- The problem has been far reaching. Over the summer, Facebook announced the discovery of hundreds of fake pages and user accounts on its site. Some pages appeared to specifically target Americans with divisive messages, using the same tactics that Russian operatives did to influence voters during the 2016 presidential campaigns.
- The latest influence campaigns also imitated posts by legitimate pages and groups on Facebook that advocate political beliefs, making it difficult to tell what was a genuine post and what was not. Let's compare some to see if you can tell the difference. Spoiler: It isn't easy.

Left one

- Correct. “Resisters” was a fake account that Facebook removed in July. “Feminist News” is a real Facebook page, as far as we can tell. (We asked Facebook to confirm the examples of real accounts used here are in fact authentic, and were told there was no apparent reason to suspect otherwise.)
- The Resisters page existed for more than a year and described itself as “online and offline feminist activism against fascism,” according to Facebook and researchers at the Atlantic Council’s Digital Forensic Research Lab, which analyzes disinformation online and works with Facebook.
- **You may have guessed the post was from a fake page if** you looked closely at the language: “Girls make rules and you follow them. If you don’t like it you *live*. End of *the* story.”
- Broken English is not a sure sign that a post is part of an influence operation. But grammatical errors were a common trait among the Russian ads [Facebook disclosed in 2017](#) — particularly the misuse of “a” and “the,” which don’t exist in the Russian language

- Literacy is acquiring, creating, connecting and communicating meaning in a wide variety of contexts.
- Literacy fiert: a plan to action – Alberta Educartion 2010.

Trasparenza

Byung-Chul Han



Byung-Chul Han La società della trasparenza

figure nottetempo



«La domanda di trasparenza diventa forte proprio quando non c'è più fiducia. In una società che si fonda sulla fiducia, non esiste una forte richiesta di trasparenza. La società della trasparenza è una società della sfiducia che [...] si sottopone al controllo. La forte richiesta di trasparenza rinvia proprio al fatto che il fondamento morale della società è diventato fragile, che i valori morali come la sincerità o l'onestà divengono sempre più insignificanti. Al posto dell'istanza morale caduta in disgrazia, compare la trasparenza come nuovo imperativo morale».

L'analisi *Le agenzie e i conflitti di interesse*

Tornano le pagelle ma non la trasparenza

Dopo il crac Lehman Brothers Obama rafforzò i controlli
ma con la deregulation di Trump non cambierà nulla

Fake news e post-verità

Spunta un patto Google-MasterCard per monitorare gli acquisti dei clienti

La rivelazione di «Bloomberg». Ma Mountain View: nessuna condivisione di dati

NEW YORK Un accordo segreto tra Google e MasterCard consente al gigante digitale di accertare se i suoi utenti sono andati a comprare nei negozi prodotti dei quali hanno visto la pubblicità in rete, senza poi acquistarli online. La notizia pubblicata da *Bloomberg*, sconcerta, ma stavolta non c'è da scandalizzarsi più di tanto. Siamo alle solite: più svolgiamo attività sul web e usiamo strumenti elettronici di pagamento, più cresce la scia digitale che lasciamo dietro di noi. E che le aziende di big tech sanno sfruttare in modo magistrale e capil-



2

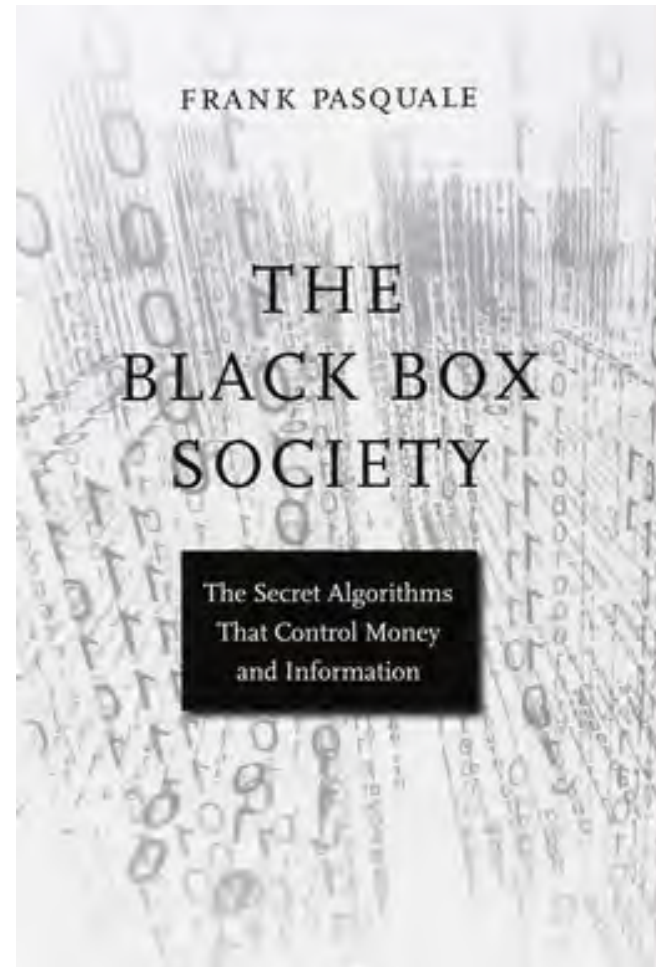
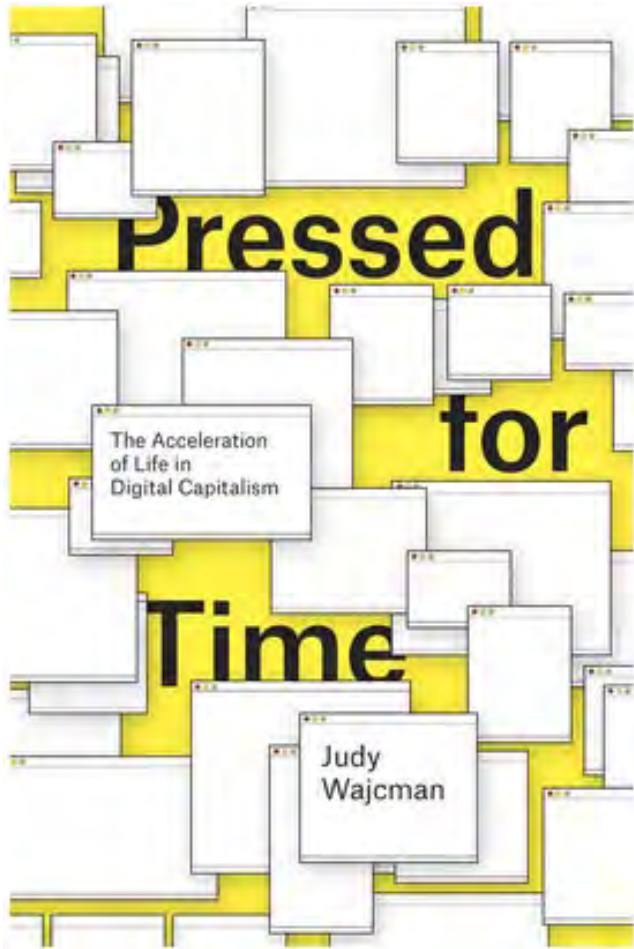
miliardi le carte distribuite da MasterCard nel mondo, uno dei tre circuiti più importanti al pari di Visa ed AmericanExpress

25

per cento la quota di mercato di MasterCard negli Stati Uniti, dove avrebbe fatto un accordo con Google per i dati degli utenti



Books



ÉRIC SADIN

LA SILICOLONIZZAZIONE DEL MONDO

L'IRRESISTIBILE ESPANSIONE
DEL LIBERISMO DIGITALE



S come Silicolonizzazione, parola che indica come il modello della Silicon Valley rappresenti la nuova forma del capitalismo globale neoliberista per conquistare, in modo felpato, nuovi spazi.

CdSera 13 settembre 2018



 **BUONENOTIZIE**

Natale, ex pastore con l'Alzheimer: le pecore sono finte ma lui le cura lo stesso

di Michele Farina

Da quando cura il gregge in polistirolo, l'uomo che vive in una struttura a Bezzecca (Trento) è più felice e più calmo. Questa e altre storie all'Alzheimer Fest



NYT -

*Rejecting Puerto Rican Death Toll,
Trump Accuses Democrats of Inflating It*



President Trump during his visit in Puerto Rico last year. Doug Mills/The New York Times

Washington Post

Trump rebuked after questioning number of deaths attributed to Hurricane Maria

President Trump falsely said that a high death count had been generated by Democrats to “make me look as bad as possible.” A George Washington University study estimated there were 2,975 “excess deaths” in the six months after Maria made landfall in Puerto Rico last September.

By John Wagner and Joel Achenbach • 2 hours ago

- **Analysis: Why Trump’s tweets about Puerto Rico are obviously untrue**
- **Fact Checker: Four Pinocchios for Trump’s complaint about Maria deaths**



Piattaforme digitali

L'Europa punta i fari sui dati di Amazon

La commissaria Vestager annuncia un'indagine conoscitiva sull'uso delle informazioni raccolte dai venditori

GABRIELLA COLARUSSO, ROMA

Dopo le indagini su Apple e Google che si sono chiuse con due maximulte, una per elusione fiscale, l'altra per abuso di posizione dominante, la Commissione europea accende i fari anche sul modello di business di Amazon, il colosso americano dell'e-commerce.

L'Antitrust guidato da Margrethe Vestager vuole capire come l'azienda raccoglie i dati sulle vendite dei commercianti che usano la sua piattaforma e se queste informazioni diano ad Amazon un vantaggio sulla concorrenza. La società di Jeff Bezos, infatti, opera nel doppio ruolo di piattaforma per il commercio elettronico a disposizione di altri venditori e di venditore diretto.

«Se Amazon prende i dati dai piccoli commercianti, cosa che può essere assolutamente legittima perché migliora il suo servizio, può usare questi dati per i propri



Jeff Bezos

calcoli su qual è la nuova grande novità, su quello che vuole la gente, su che tipo di offerta vuole ricevere, su cosa li spinge a comprare cosa», ha spiegato ieri Vestager. Questo viola le regole europee sulla concorrenza, è un abuso di posizione dominante? Per ora si tratta solo di una indagine preliminare e

non formale, ha precisato il Commissario. Entro due mesi l'Antitrust riceverà le risposte ai questionari che sono stati inviati ai commercianti e deciderà se aprire formalmente il caso. L'anno scorso, per la prima volta nella storia di Amazon, più del 50% dei prodotti venduti sulla piattaforma sono arrivati da "terze parti", da venditori indipendenti. I servizi a terzi hanno portato nelle casse della multinazionale 31,88 miliardi di dollari di ricavi, rispetto ai 22,99 miliardi dell'anno precedente: la seconda voce di fatturato dopo le vendite al dettaglio. L'azienda non ha commentato la notizia dell'indagine conoscitiva avviata dall'Antitrust, ma altre volte in passato si è difesa dalle accuse sostenendo che le vendite online sono solo una parte del suo fatturato e che comunque opera in un mercato con molti competitor. Una posizione che tuttavia non convince i regolatori europei e nemmeno l'antitrust americano.

ISERVIZI A TERZI

31,88 mld

Le vendite per conto terzi hanno fruttato ad Amazon 31,88 miliardi di ricavi

A fine agosto, il presidente degli Stati Uniti, Donald Trump, ha accusato Amazon di «distruggere» i piccoli negozi, e le grandi aziende digitali di essere «in una posizione molto anticoncorrenziale». La Federal Trade Commission e il Dipartimento della giustizia americano hanno entrambi aperto indagini per verificare se queste società violino le regole sulla concorrenza. In Europa poi potrebbe aprirsi un nuovo fronte quando la Commissione e il Parlamento discuteranno della nuova tassa al 3% sulle imprese digitali, e potrebbe avvenire già entro l'anno.

La pressione politica tuttavia non sembra preoccupare Bezos: «Tutte le grandi istituzioni saranno e devono essere controllate. Siamo così innovativi che qualsiasi regolazione dovesse arrivare non smetteremo di fornire servizi ai nostri clienti», ha dichiarato nelle scorse settimane.

© RIPRODUZIONE RISERVATA

Carofiglio Repubblica settembre 2018

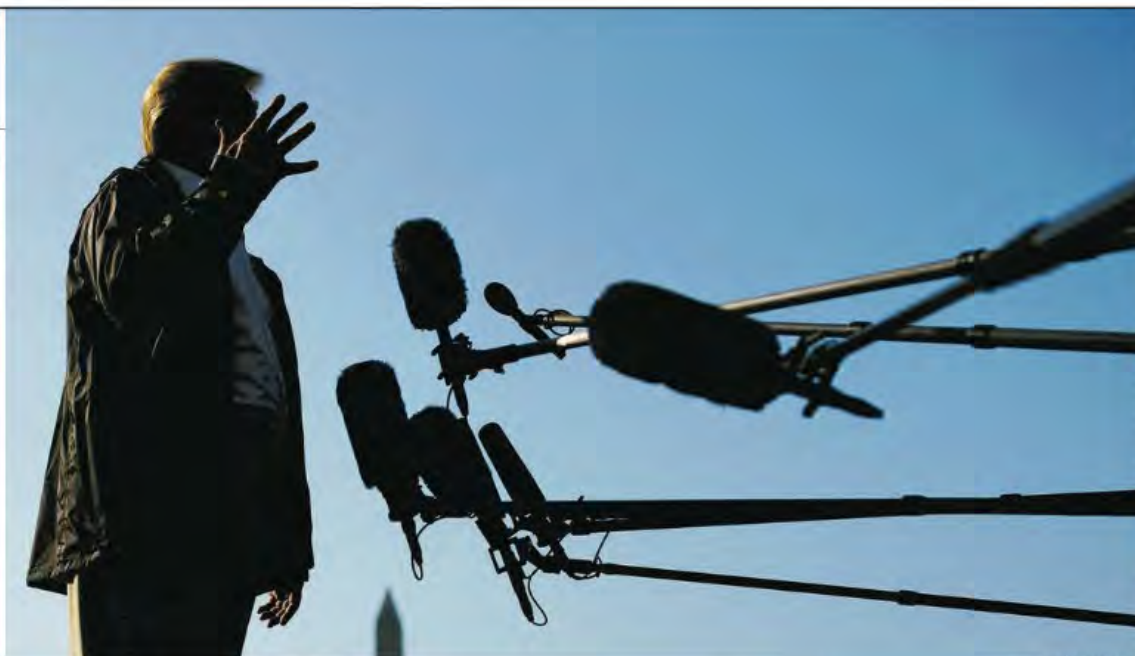
- Il filosofo John Searle, teorico del rapporto fra linguaggio e istituzioni, sostiene che le società vengano costruite e si reggano, essenzialmente su una premessa linguistica: sul fatto, cioè, che formulare un'affermazione comporti un impegno di verità e di correttezza nei confronti dei destinatari.
- Nel caso della politica: formulare un'affermazione comporta un obbligo di verità e correttezza nei confronti dei cittadini. Siano essi sostenitori, siano essi avversari. Non osservare questo impegno mette in pericolo il primario contratto sociale di una comunità, cioè la fiducia in un linguaggio condiviso.
- Le società nelle quali prevalgono le asserzioni vuote di significato sono in cattiva salute: in esse, alla perdita di senso dei discorsi, consegue una pericolosissima caduta di legittimazione delle istituzioni. È per questo che occuparsi del linguaggio pubblico e della sua qualità non è dunque un lusso da intellettuali o un esercizio da accademici. È un dovere cruciale dell'etica civile

Resti

Online News Association Fake news, propaganda, manipolazioni, speculazioni finanziarie: così a Austin in Texas gli stati generali dell'informazione fanno il punto sulla crisi della stampa. E lanciano la controffensiva: "Non facciamo il loro gioco"

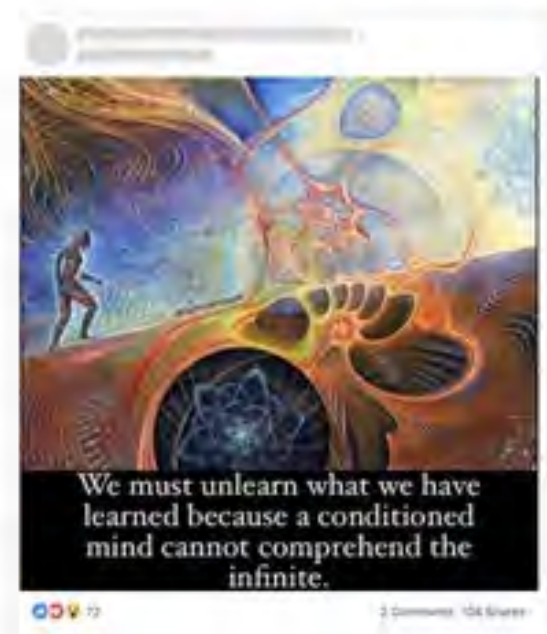
American Troll il giornalismo sotto attacco

RAFFAELLA MENICHINI, AUSTIN (TEXAS)



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Which post about self care is from a fake page?



Can You Spot the Deceptive Facebook Post?

Facebook has been plagued with disinformation posts placed by foreign operatives. Can you tell the difference between a real post and one designed to fool you?



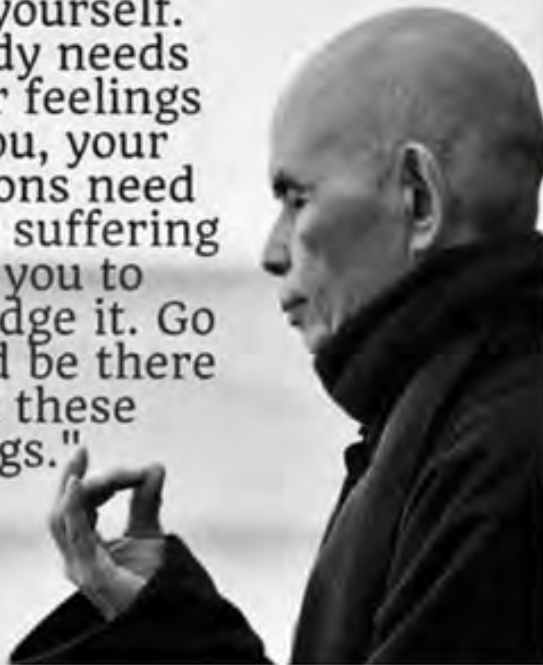
Revolutionary Self-Care

April 23 · 🌐



you can't heal what you won't acknowledge.

"Go back and take care of yourself. Your body needs you, your feelings need you, your perceptions need you. Your suffering needs you to acknowledge it. Go home and be there for all these things."



👍❤️ 225

95 Shares

REAL PAGE ✖



Mindful Being

July 1 at 8:00 PM · 🌐



We must unlearn what we have learned because a conditioned mind cannot comprehend the infinite.

👍❤️👎 72

2 Comments 104 Shares

INFLUENCE CAMPAIGN

Breaking News by Alan Rusbridger review – the remaking of journalism and why it matters now

The former Guardian editor details a revolution in journalism.
Can it still perform its vital, truth-telling role?



▲ Eloquent in his argument for well-resourced journalism ... Alan Rusbridger. Photograph: Greg James

Truth is a small word liable to sanctimonious overuse and philosophical dispute, but in its humblest sense of accurate and verifiable information we like to think we know it when we see it. In Alan Rusbridger's view, journalism should be among its leading providers: societies depend on good journalism to distinguish fact from fiction, to form a realistic view of their problems and futures. And here, he

Repubblica 22 9 2018

Nomine

Istat, la disfida gialloverde per conquistare la presidenza

Il M5S non vuole lasciare a Salvini il controllo dell'Istituto che diffonde i dati chiave

Hints from cognitive psychology

It's not just what people think that matters, but how they think.

Refuting misinformation involves dealing with complex cognitive processes

A simple myth is more cognitively attractive than an over-complicated correction

For those who are strongly fixed in their views, encountering counter arguments, can cause them to strengthen their views

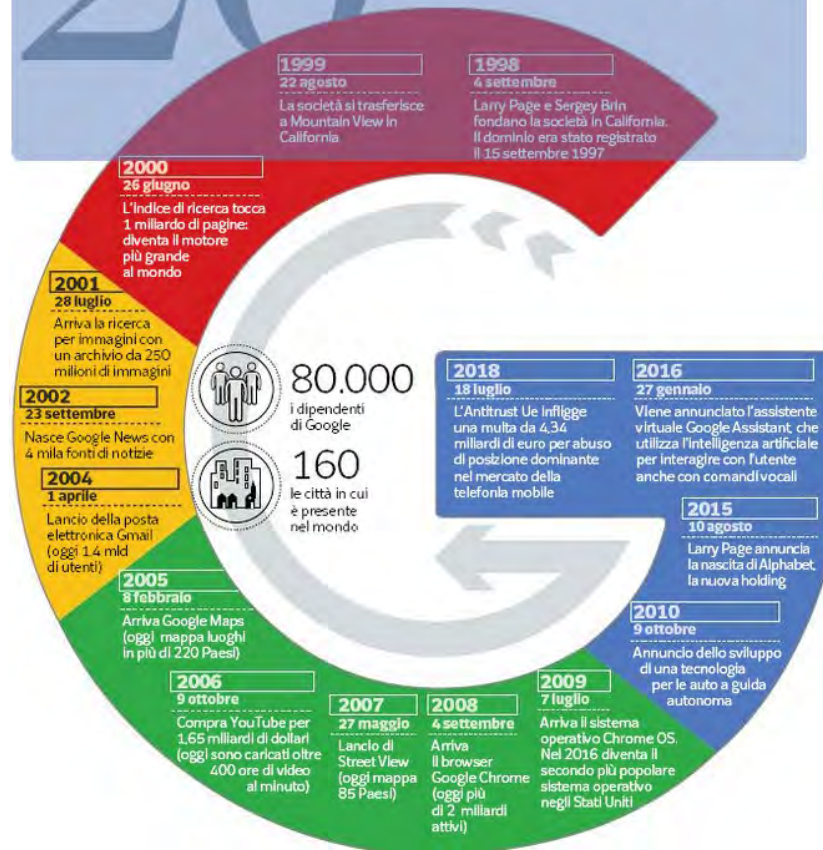
Tutto è iniziato in un garage con due ragazzi dagli intenti filantropici



Co-fondatori Larry Page, a sinistra, e Sergey Brin nella sede di Google

Dal lavoro agli affetti, così ha cambiato (non sempre in meglio) le nostre vite

20 anni di Google



- Globale verso locale – I dati digitali disponibili nelle reti sociali, nell’ internet delle cose, nel Web, subiscono un fenomeno di progressiva frammentazione. A fronte di questa frammentazione il rischio sta nella creazione di un vastissimo universo di “echo chambers”, recinti entro cui conosciamo e decidiamo, senza più confrontarsi con conoscenze globali e condivise. Certamente, siti come Wikipedia, DBPedia, OpenMap, forniscono conoscenza globale, ma se teniamo conto che il 95% degli accessi al Web tramite motori di ricerca si ferma alla prima pagina, si comprende come costruzione e la condivisione di conoscenza globale e la relazione tra conoscenza locale e globale sia un tema determinante per il formarsi di una conoscenza consapevole non affetta da bias.
- Analogico verso digitale – Quanto più la nostra esperienza sensibile è basata sui dati digitali, piuttosto che sulla esperienza concreta, tanto più si estende e l’ambito del virtuale nella sua relazione con il mondo reale. Il confine tra i due mondi è in continua evoluzione e solo un equilibrio cognitivo consapevole della relazione tra essi può ridurre il rischio di far prevalere il virtuale (una nuova edizione de “l’ha detto il computer”).
- Il “cosa” e il “perché” – L’utilizzo sempre più diffuso di algoritmi e tecniche di apprendimento e la produzione di modelli si basa sulla necessità di utilizzare grandi moli di dati. Spesso accade che la tecnica utilizzata (e i dati su cui si basa la elaborazione) siano noti solo al proprietario della tecnica e dei dati, e ignoti all’utente, per cui il proprietario attraverso il modello prodotto definisce il “cosa”, e l’utente non conosce il “perché”.

- **Condivisione (dei dati)** – per condividere i dati tra diversi soggetti che li utilizzano servono forme di codifica e modellazione che ne permettano la diffusione e la comprensione condivisa.
- **Significato (dei dati)** – spesso i dati ci arrivano grezzi, per cui è necessario applicare metodi e modelli per attribuire loro un significato (nell'esempio precedente della temperatura l'attribuzione di significato è relativamente semplice, in generale è un processo complesso che richiede la conoscenza di modelli descrittivi di tipo concettuale, cioè universali e generali, indipendenti dal tipo di dato e dal fenomeno osservato).
- **Qualità (dei dati)** – i dati digitali sono affetti da errori, sono spesso imprecisi, incompleti, inconsistenti, non aggiornati. E' necessario definire le dimensioni della qualità e metodi per la loro misurazione e miglioramento.
- **Valore (dei dati)** – il valore che i dati ci trasmettono può essere di tipo decisionale, economico, sociale, etico, ed è in parte oggettivo, in parte soggettivo. Sono necessari modelli per trattare le diverse tipologie del valore.
- **Astrazioni (sui dati)** – gli esseri umani hanno una razionalità e capacità cognitive limitate. A fronte della crescente disponibilità di dati digitali, estrarre e organizzare conoscenza richiede la capacità di descrivere ed rappresentare la conoscenza a diversi livelli di astrazione.
- **Integrazione (dei dati)** – i dati digitali forniscono modelli della realtà sempre più frammentati ed eterogenei, sono necessarie tecniche e metodologie per riconciliare e integrare i dati e la conoscenza, accrescendo in questo modo il loro valore.
- **Tecniche di apprendimento (sui dati)** – sempre di più è necessario definire e utilizzare tecniche che permettano di apprendere a partire di dati di training classificazioni, modelli interpretativi, predittivi e prescrittivi.
- **Linguaggi e ambienti l'analisi (dei dati)** – I dati sono rappresentazioni del mondo reale. Per estrarre conoscenza da essi, è necessario analizzarli ed elaborarli. Ciò richiede l'uso di linguaggi e ambienti automatizzati che siano facili da apprendere ed utilizzare

Nope. “Mindful Being” was part of an influence operation, and “Revolutionary Self-Care” is a real Facebook page.

The Mindful Being page existed for only two months before Facebook removed it. The page mainly focused on innocuous topics like health and well-being, though some of its posts challenged the media and the pharmaceutical industry, suggesting it was laying the groundwork to tread into more controversial content.

One of these posts was from a genuine Facebook page that supports feminism, and the other was part of an influence campaign. **Can you guess which post is from a fake page?**



Corriere
16 settembre



L'appuntamento Torna A seminar la buona pianta e supera il concetto di «sostenibilità», allargandosi al «bene comune». Uno studioso di innovazione spiega perché oggi alle imprese non basta più un algoritmo: serve una visione. Coraggiosa

INTERESSI
CONDIVISI

«NUMERI, SÌ, MA ANCHE **VALORI**»
È LA RIVOLUZIONE DEL **CAPITALE**

NYT 4 settembre 2018

Hard Lessons (Thanks, Amazon) Breathe Life Into Retail Stores

- It may be too early to declare the death of retail, as Americans have started shopping more in stores, reflecting a broad reordering of the \$3.5 trillion industry.
- Fewer retailers are capturing more of the gains, and many successful stores are now a cross between a fast-food drive-through and a hotel concierge.

4h ago



Growth Has Lifted Counties That Voted for Trump. Mostly, It's the Wealthy Ones.

President Trump's economy has left the most distressed swaths of the country waiting for their share of the good times.

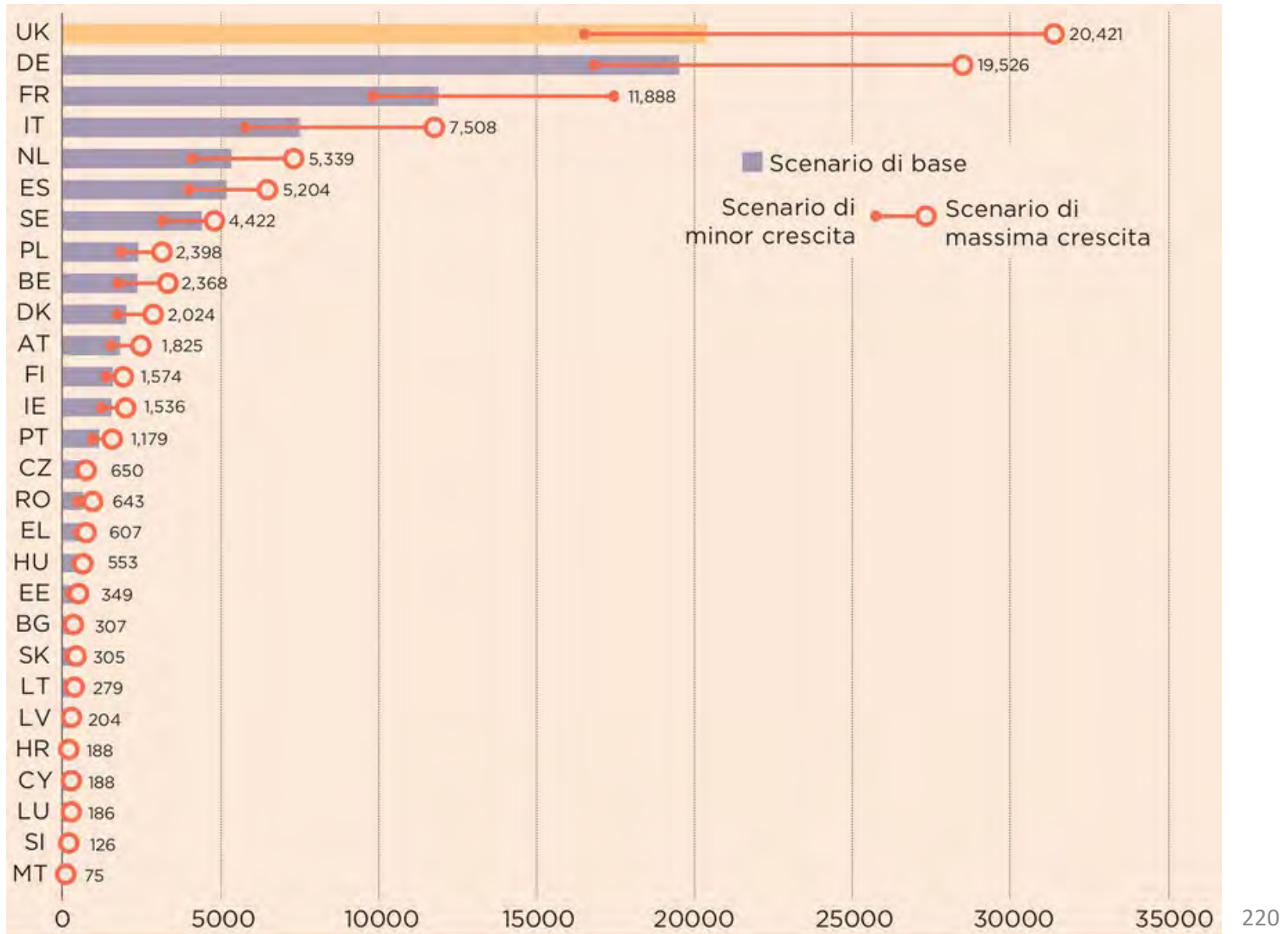
Amazon is moving aggressively into digital advertising, emerging as a competitor to Facebook and Google.

8h ago

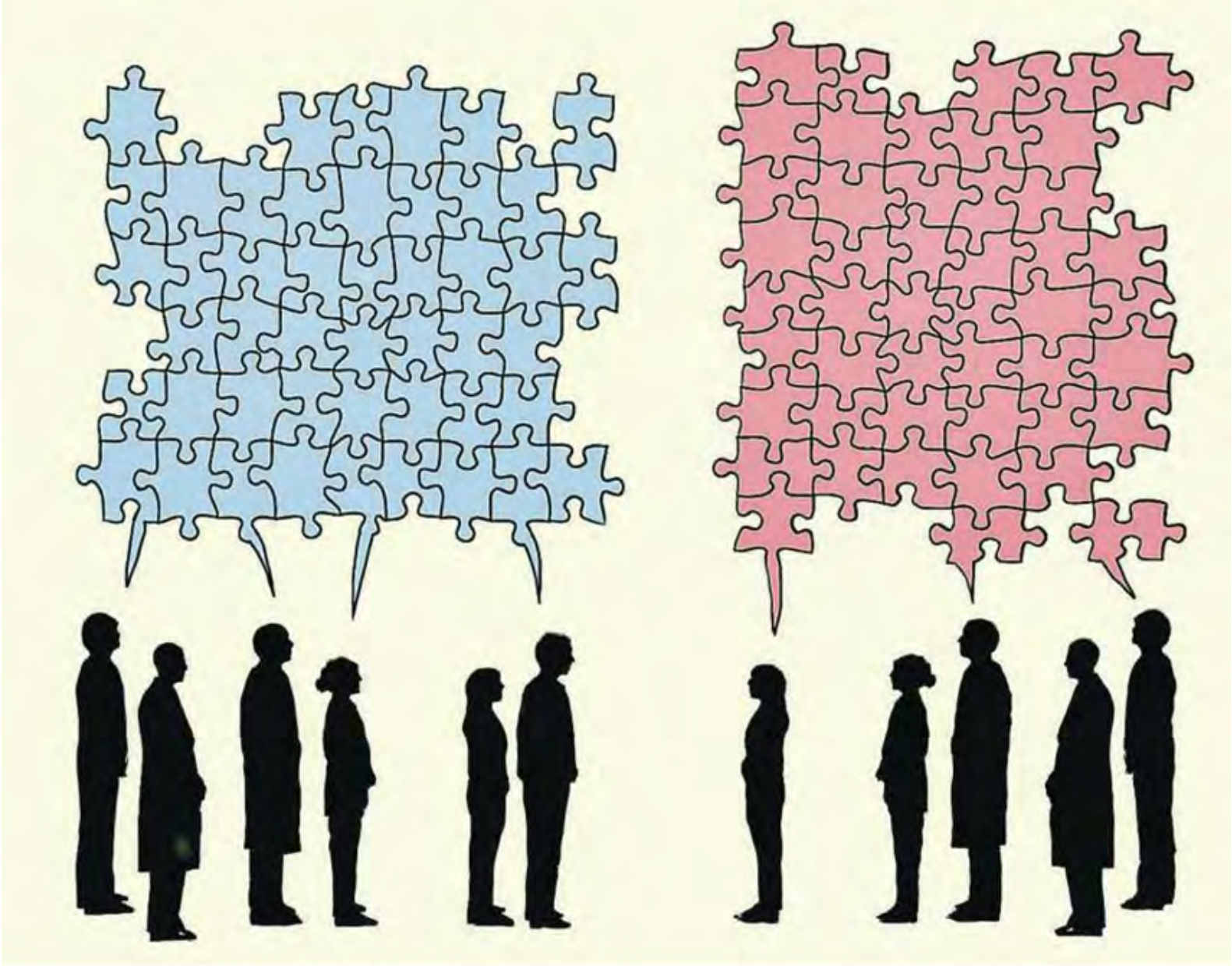
"Our stores are at the center of our strategy, and they are at the center of our success right now," said Target's chief executive, Brian Cornell.

Jeenah Moon for The New York Times

Ricavi della industria dei dati, previsioni al 2020







Il pericolo della falsa realtà Rep 26/9/2018



Moisés Naim venezuelano, è membro del Carnegie Endowment, think tank con sede a Washington. Il suo ultimo libro in Italia è "La fine del potere" (Mondadori, 2013) @moisesnaim www.moisesnaim.com

“
Le immagini manipolate possono intorbidire le relazioni politiche, economiche, diplomatiche ed esacerbare i conflitti internazionali
”

Alla fine dell'anno scorso cominciarono a girare sulla Rete dei video pornografici che avevano come protagoniste principali alcune delle attrici e cantanti più famose del momento. Naturalmente i video diventarono virali e furono visti da milioni di persone in tutto il mondo. Nel giro di pochi giorni uscì fuori che Scarlett Johansson, Taylor Swift, Katy Perry e altre artiste famose non erano le vere protagoniste dei filmati, ma le vittime di una nuova tecnologia che, utilizzando sistemi di intelligenza artificiale e altri strumenti digitali avanzati, permette di inserire l'immagine facciale di qualsiasi persona in un video.

Quello fu solo l'inizio. Ben presto Angela Merkel, Donald Trump e Mauricio Macri finirono anche loro vittime del cosiddetto *deepfake*, la falsificazione profonda. Barack Obama fu utilizzato, senza il suo consenso, per esemplificare i possibili impieghi nefasti di questa tecnologia: si vedeva l'ex presidente Usa dire in un discorso quello che il falsificatore voleva che dicesse, e che l'ex presidente non aveva mai detto. Ma il risultato era un video molto reale.

La manipolazione di immagini non è certo una novità. I governi autoritari hanno una lunga storia di leader caduti in disgrazia che venivano fatti «sparire» dalle foto ufficiali. E già dal 1990 *Photoshop* consente agli utenti di alterare le fotografie digitali.

Il *deepfake* però è differente. E molto più pericoloso. Differente perché, da quando sono stati messi in circolazione i falsi video delle attrici a oggi, la tecnologia è migliorata moltissimo. L'immagine corporea e l'espressione facciale sono iperrealistiche e l'imitazione della voce e della gestualità di una persona è talmente accurata che risulta impossibile capire che si tratta di una falsificazione, a meno di poter contare su sofisticati programmi di verifica digitale. E il pericolo del *deepfake* è che questa tecnologia è alla portata di chiunque.

Un ex fidanzato indispettito e psicopatico può produrre e diffondere anonimamente sui *social network* un video che imiti perfettamente la voce, i gesti e la faccia di una donna che lo ha lasciato, e dove la si vede fare o dire le bestialità più vergognose. Le immagini di poliziotti che pestano brutalmente una donna anziana che partecipa a una protesta contro il governo possono provocare violenti scontri tra manifestanti e agenti. Il rispettato leader di un gruppo razziale o religioso può istigare i suoi seguaci ad attaccare membri di un'altra razza o religione. Alcuni studenti possono produrre un video compromettente di un professore che ripudia. Estorsori digitali possono minacciare un'impresa di divulgare un video che danneggerà la sua reputazione, se non riceveranno la somma richiesta.

I possibili impieghi del *deepfake* in politica, in economia o nelle relazioni internazionali sono tanto vari quanto inquietanti.

La divulgazione di un video che mostra un candidato alla presidenza di un Paese che dice o fa cose riprovevoli poco prima del voto diventerà uno stratagemma elettorale di uso comune. Magari il candidato rivale non approverà questo trucco indecente, ma i suoi seguaci più radicali potrebbero realizzare il video e distribuirlo senza chiedere il permesso a nessuno.

I video falsificati possono anche intorbidire drammaticamente le relazioni tra Paesi ed esacerbare i conflitti internazionali.

E non è una mera ipotesi, è già avvenuto: lo scorso anno l'emiro del Qatar, Tamim bin Hamad al-Thani, comparve in un video in cui elogiava e appoggiava Hamas, gli Hezbollah, i Fratelli musulmani e l'Iran, provocando la reazione furibonda dell'Arabia Saudita, degli Emirati Arabi Uniti, del Bahrein e dell'Egitto, che già avevano forti attriti con il Qatar. Questi Paesi denunciarono il discorso dell'emiro come un appoggio al terrorismo e ruppero le relazioni diplomatiche, chiudendo le frontiere e imponendo un embargo aereo, marittimo e terrestre. La realtà, però, è che l'emiro del Qatar non aveva mai pronunciato quel discorso: il video che aveva scatenato l'escalation del conflitto era un falso. Molto reale è invece il boicottaggio, che rimane in vigore.

La minaccia che rappresenta il *deepfake* per l'armonia sociale, la democrazia e la sicurezza internazionale è evidente. Gli antidoti lo sono molto meno, anche se alcune proposte ci sono.

Tutte le organizzazioni che producono o distribuiscono fotografie o video devono impegnarsi a usare blocchi tecnologici che garantiscano l'inalterabilità dei loro materiali visivi. Anche le persone devono avere accesso a tecnologie che le proteggano dal rischio di diventare vittime di *deepfake*. Bisogna adeguare le leggi per fare in modo che coloro che diffamano o arrecano danno ad altri attraverso l'uso di queste tecnologie siano chiamati a risponderne di fronte alla giustizia. L'uso dell'anonimato sulla Rete dev'essere reso più difficile.

Tutto questo è necessario, ma insufficiente. Bisognerà fare molto di più.

Siamo entrati in un'era in cui la capacità di distinguere la verità dalla menzogna, i fatti dalle falsità, si è andata progressivamente erodendo. E con essa la fiducia nelle istituzioni e nella democrazia. Il *deepfake* è un'altra arma nuova e potente nell'arsenale dei mercanti di menzogne. Bisogna combatterli.

(Traduzione di Fabio Galimberti)