

Focus point analytics e Big Data

Dott. Laura Spinsanti

Di cosa parliamo

- Una visione d'insieme sui Big Data
- Chi produce i Big Data? Quanti sono i Big data?
- Basta avere tanti dati per parlare di Big Data?
- Chi usa i Big data in campo istituzionale?
- Big Data Analysis: nuove possibilità... e nuove sfide
- Nuove figure professionali: data scientist & data system integrator
- Big data step by step

Una visione d'insieme sui Big Data

2013

Big data è come il sesso tra adolescenti: tutti ne parlano, nessuno sa davvero come fare, tutti pensano che tutti gli altri lo stiano facendo, quindi tutti affermano che lo stanno facendo...

2016

Cosa è cambiato?



Professor Dan Ariely

Chi produce i Big Data?

Tutti noi produciamo dati attraverso:

dati diretti,

dati indiretti,

dati ufficiali,

dati volontari,

dati di mobilità,

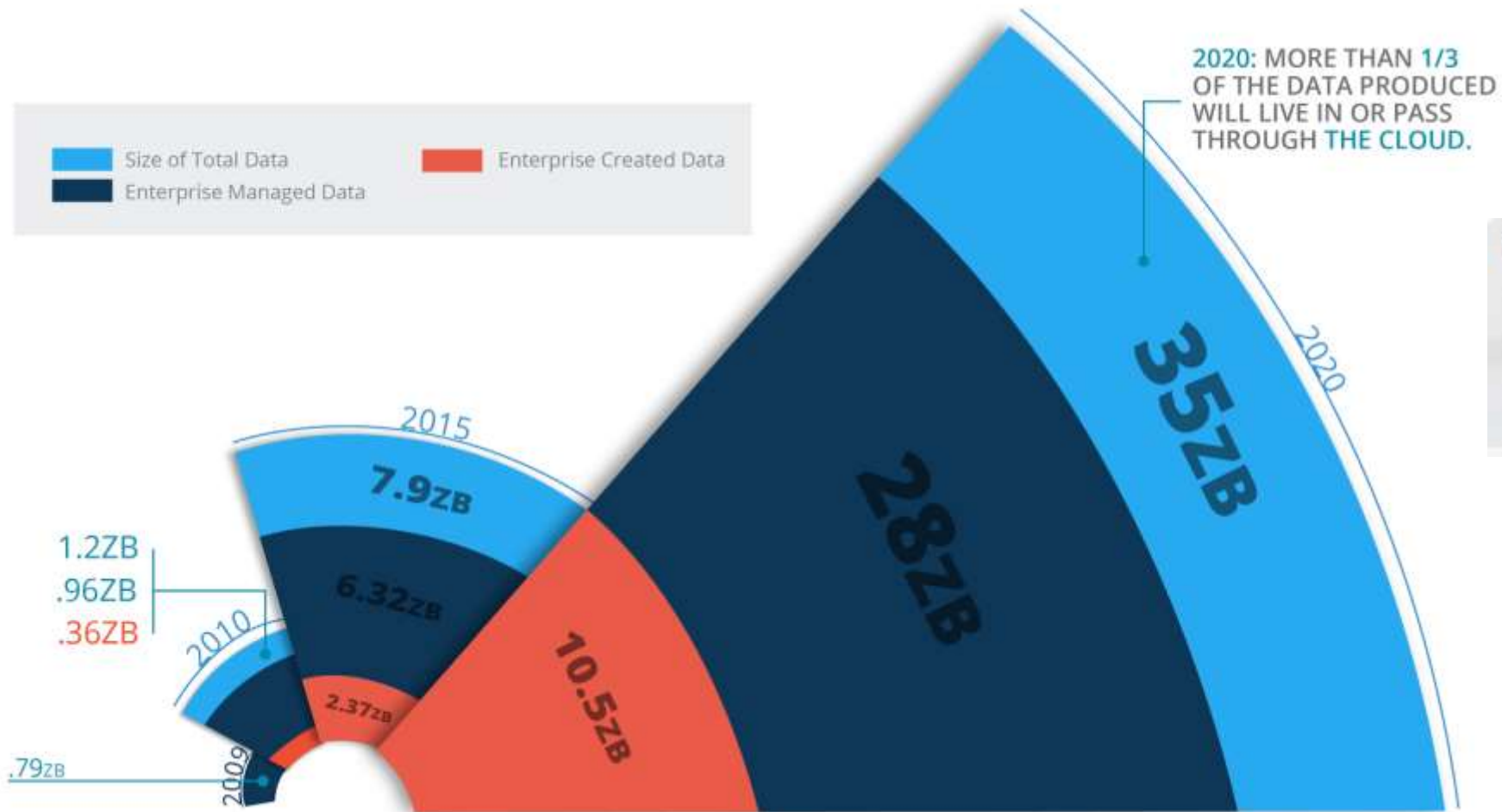
dati di telecomunicazione,

dati delle reti sociali fisiche e virtuali



<https://whatsthebigdata.com>

Quanti sono i Big Data?



WHAT IS A ZETTABYTE?

1,000,000,000,000	gigabytes
1,000,000,000,000	terabytes
1,000,000,000,000	petabytes
1,000,000,000,000	exabytes
1,000,000,000,000	zettabyte

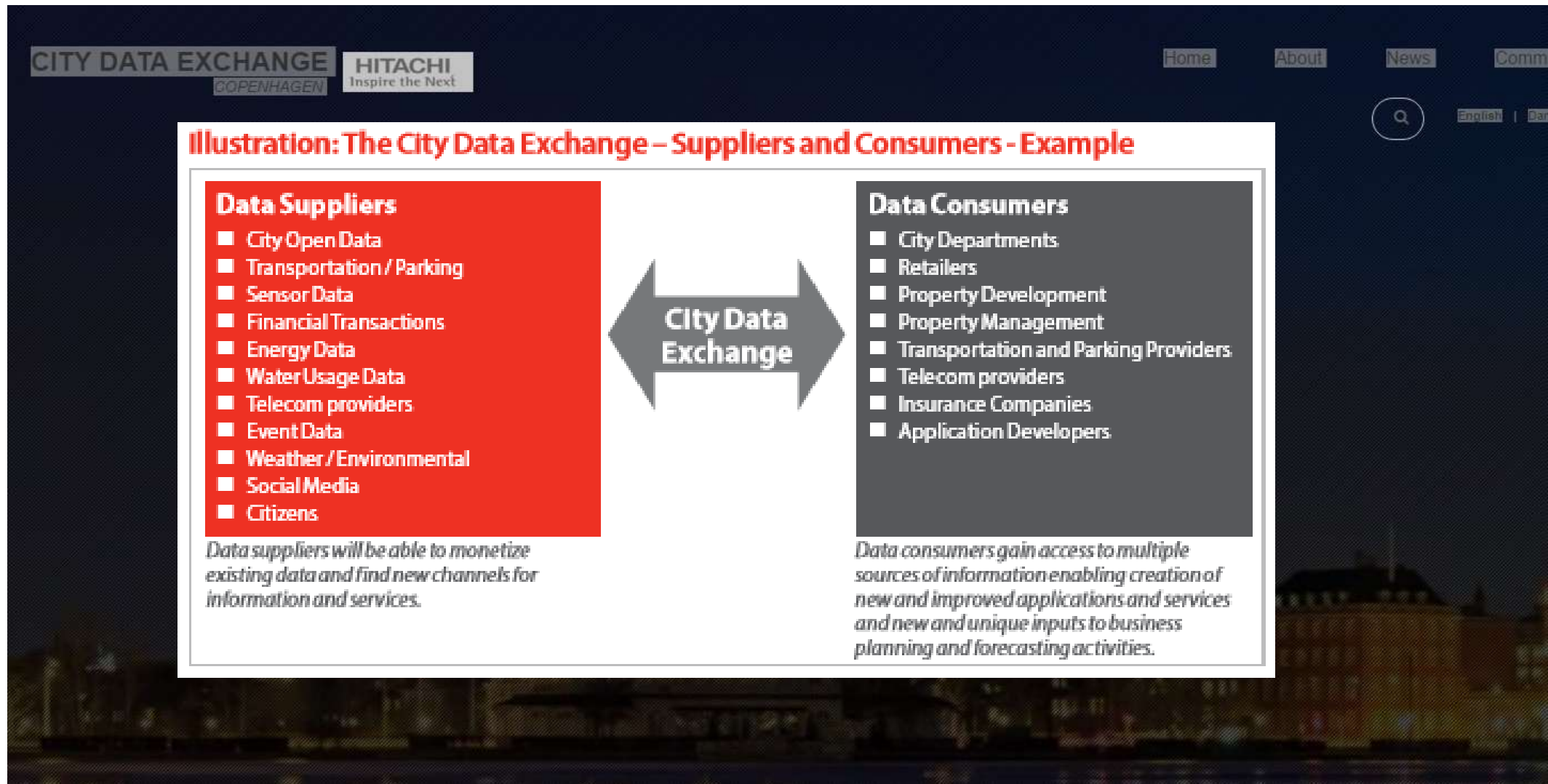
Basta avere tanti dati per parlare di Big Data?

Le 7 V dei Big data: Volume, Velocità, Varietà, ...

- **volume**: rappresenta la dimensione effettiva del dataset;
- **velocità**: si riferisce alla velocità di generazione dei dati;
- **varietà**: si riferisce alle varie tipologie di dati (strutturati e non), provenienti da fonti diverse;
- **variabilità**: si riferisce al *contesto* in cui è necessario analizzare le informazioni;
- **veridicità**: si riferisce all'*affidabilità* delle informazioni stesse e delle decisioni da essa derivanti;
- **visualizzazione**
- **valore**

Chi usa i Big data in campo istituzionale?

City Data Exchange: il caso di Copenhagen - Hitachi



Carlo Ratti @ SENSEable City Lab, MIT

SENSEABLE CITY LAB

MIT

Urban imagination and social innovation through design & science

The real-time city is real! As layers of networks and digital information blanket urban space, new approaches to the study of the built environment are emerging. The way we describe and understand cities is being radically transformed—as are the tools we use to design them. The mission of the *Senseable City Laboratory*—a research initiative at the Massachusetts Institute of Technology—is to anticipate these changes and study them from a critical point of view.

Not bound by the methodologies of a single field, the Lab is characterized by an omni-disciplinary approach: it speaks the language of designers, planners, engineers, physicists, biologists and social scientists. Senseable is as fluent with industry partners as it is with metropolitan governments, individual citizens and disadvantaged communities. Through design and science, the Lab develops and deploys tools to learn about cities—so that cities can learn about us.

Projects



Roboat
2016



Urban Exposures
2016



Monitour
2016



Forum on Future Cities: ...
2016



Light Traffic
2016



Urban Lens
2016



Underworlds
2015



A Tale of Many Cities
2014



The Road Ahead
2014

Contacts

Senseable City Laboratory

MIT 9-209

Publications

→ 'Exposure Track' - The Impact of Mobile Device Based Mobility Patterns on Quantifying Population Exposure to Air

News

→ Book Review: The city of tomorrow

2016-09-22 / Download

Positions

We always seek outstanding candidates from a variety of backgrounds besides our specific position openings.

Researchers

Carlo Ratti *Director, Professor of the Practice*,
Assaf Biderman *Associate Director*, Erin Baumgartner *Assistant Director*, Rex Britter

United Nations Global Pulse



UNITED NATIONS GLOBAL PULSE

Harnessing big data for development and humanitarian action

SEARCH



ABOUT

PROJECTS

LABS

NEWS

CHALLENGES

PRIVACY

PARTNERSHIPS

CONTACT

HOME

PUBLIC HEALTH PROJECTS



Monitoring In Real Time The Implementation Of HIV Mother-To-Child Prevention Programme



Understanding Immunisation Awareness And Sentiment Through Analysis Of Social Media And News Content (2015)



Analysing Social Media Conversations To Understand Public Perceptions Of Sanitation (2014)



Data Visualisation And Interactive Mapping To Support Response To Disease Outbreak



Analyzing Attitudes Towards Contraception & Teenage Pregnancy Using Social Data (2014)

BROWSE BY LAB

Jakarta

Kampala

New York

BROWSE BY PROGRAMME

Climate & Resilience

Data Privacy & Protection

Economic Well-being

Food & Agriculture

Gender

Humanitarian Action

Public Health

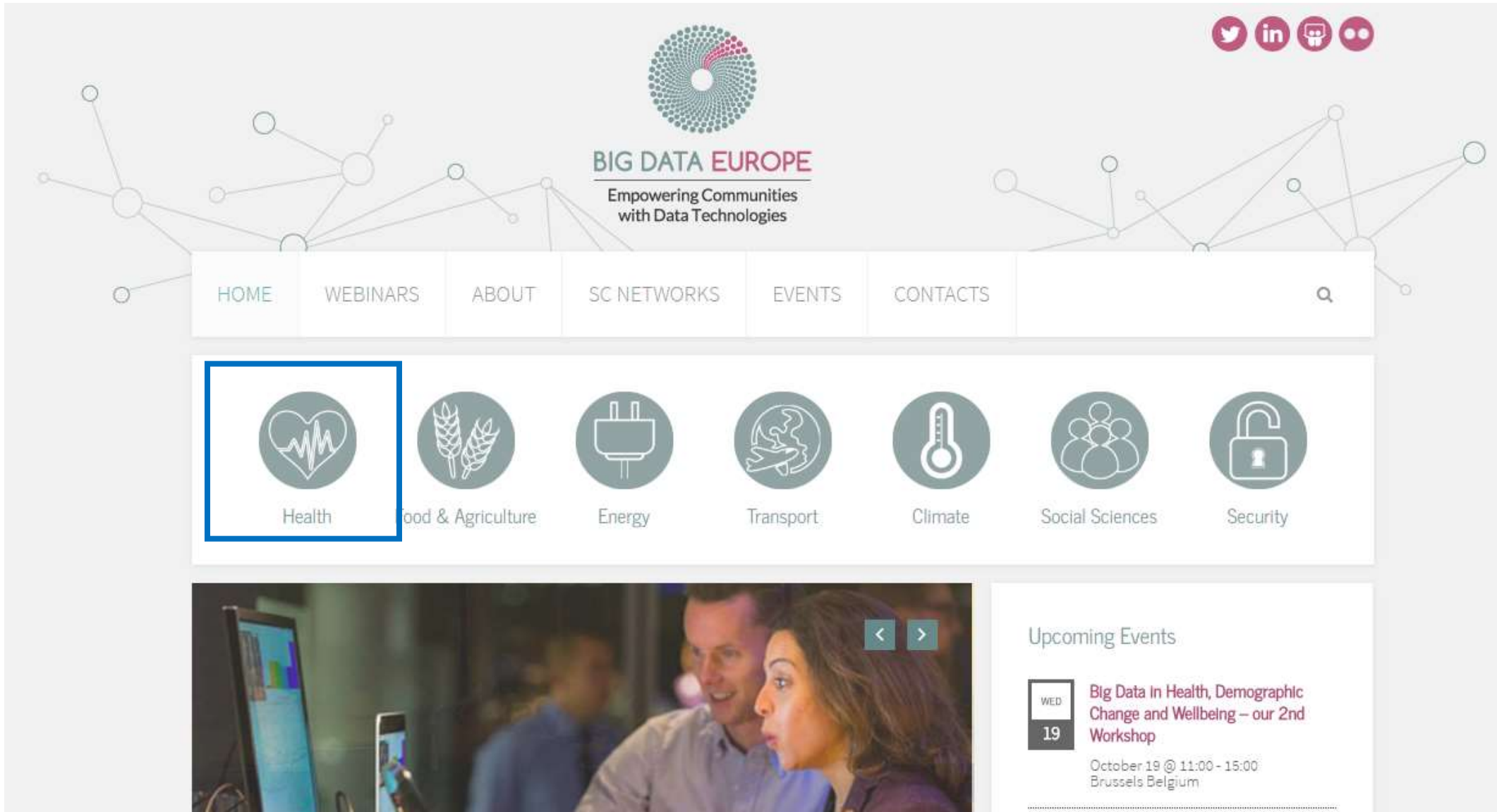
Real-time Evaluation

The Sustainable Development Goals (SDGs)

SUBSCRIBE TO OUR NEWSLETTER

GO

EU nell'ambito Horizon 2020 (un esempio): Big Data Europe



The screenshot shows the homepage of the Big Data Europe website. At the top, there is a navigation bar with links: HOME, WEBINARS, ABOUT, SC NETWORKS, EVENTS, and CONTACTS. Below the navigation bar is a row of seven circular icons representing different sectors: Health (heart with ECG), Food & Agriculture (wheat), Energy (plug), Transport (globe with car), Climate (thermometer), Social Sciences (group of people), and Security (padlock). The 'Health' icon is highlighted with a blue border. Below the icons is a section titled 'Upcoming Events' featuring a workshop on 'Big Data in Health, Demographic Change and Wellbeing' scheduled for October 19 in Brussels.

BIG DATA EUROPE
Empowering Communities
with Data Technologies

HOME WEBINARS ABOUT SC NETWORKS EVENTS CONTACTS

Health Food & Agriculture Energy Transport Climate Social Sciences Security

Upcoming Events

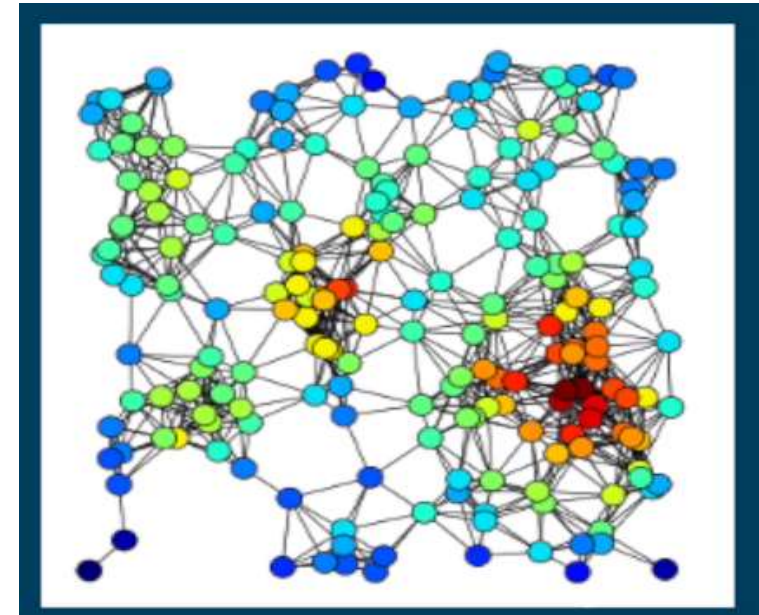
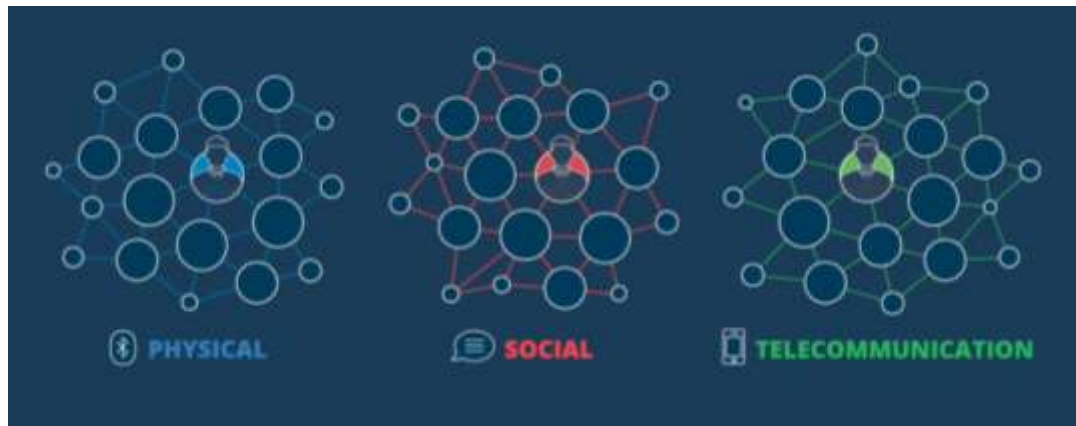
WED 19
Big Data in Health, Demographic Change and Wellbeing – our 2nd Workshop
October 19 @ 11:00 - 15:00
Brussels Belgium

Big Data Analysis

- una nuova sfida concettuale...
 - Esaminare grandi moli di dati
 - Riconoscere l'appropriatezza delle informazioni
 - Identificare pattern nascosti e correlazioni sconosciute
 - Modelli no-SQL
- ...e computazionale
 - nuovi modelli di rappresentazione in grado di gestire tale mole di dati con elaborazioni in parallelo dei database, per esempio **MapReduce**.

Nuove possibilità

Contagio e social networks analisi: Big Data
per predire la diffusione epidemica e
programmare vaccinazioni mirate

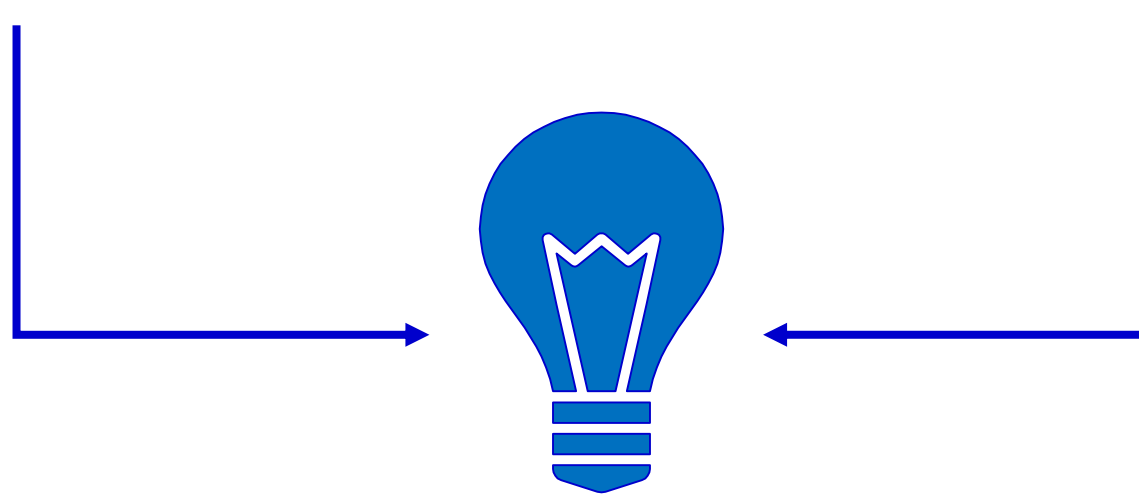


Vaccination and Complex Social Dynamics, 2016

Indurre al cambiamento

Indurre al cambiamento di comportamenti a rischio: un caso di studio per valutare gli incentivi all'attività fisica

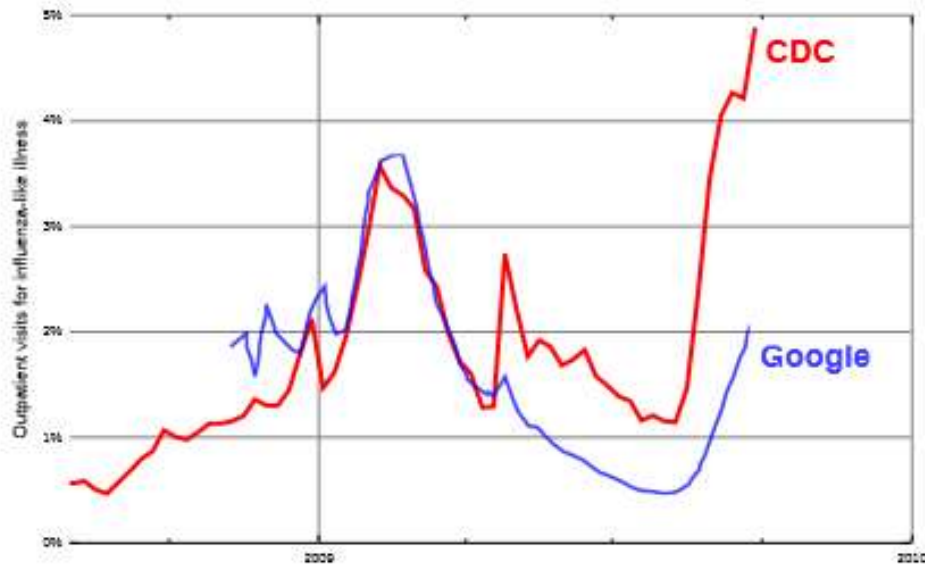
Aderenza ai trattamenti farmacologici: predictive analytics per capire quali pazienti sono a rischio di non prendere medicinali



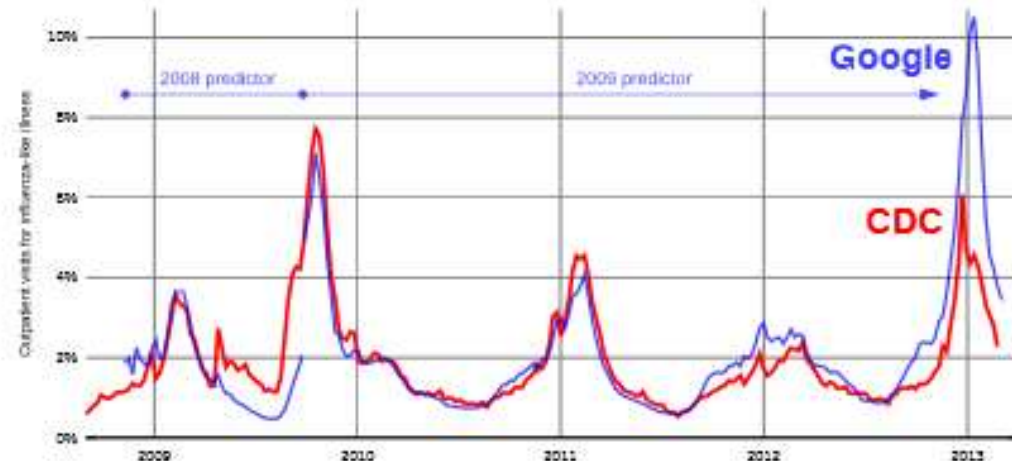
Non è tutto oro..

Trappole nell'analisi dei Big Data: il caso di Google Flu Trends per misurare l'incidenza dell'influenza

Performance in the first year



Second divergence in 2012–2013 for U.S.



Nuove sfide: privacy un problema aperto

- Anonymity: soppressione o generalizzazione di dati quali nome, indirizzo, numero di telefono...
- Unicità: il rischio di re-identificare un soggetto conoscendo n pezzi di informazione su un utente
- *“Unique in the shopping mall: On the reidentifiability of credit card metadata”* (Montjoye et al., 2015)
 - Studiando tre mesi di dati di utilizzo di carte di credito su 1.1 milione di persone (in una città europea) mostra che 4 punti spaziotemporali sono sufficienti per reidentificare in maniera univocal il 90% degli individui

“Anonymity is no longer a solution to the privacy problem in the big data context.”
Yves-Alexandre de Montjoye

Big data & health 2020

Strategic objectives

Health 2020 recognizes that governments can achieve real improvements in health if they work across government to fulfil two linked strategic objectives.



The two strategic objectives are:

- improving health for all and reducing health inequalities
- improving leadership and participatory governance for health.

- Una grossa percentuale dei dati medici NON è strutturata ed è clinicamente rilevante;
- I dati sono disseminati in molteplici posti;
 - cartelle cliniche (elettroniche),
 - laboratori,
 - sistemi di immagini,
 - corrispondenza tra medici,
 - reclami
- Usare abilmente i big data per:
 - Costruire sistemi sostenibili
 - Collaborare per migliorare l'assistenza e i risultati
 - Incrementare e migliorare l'accesso al sistema sanitario nel suo complesso

THEMATRIX

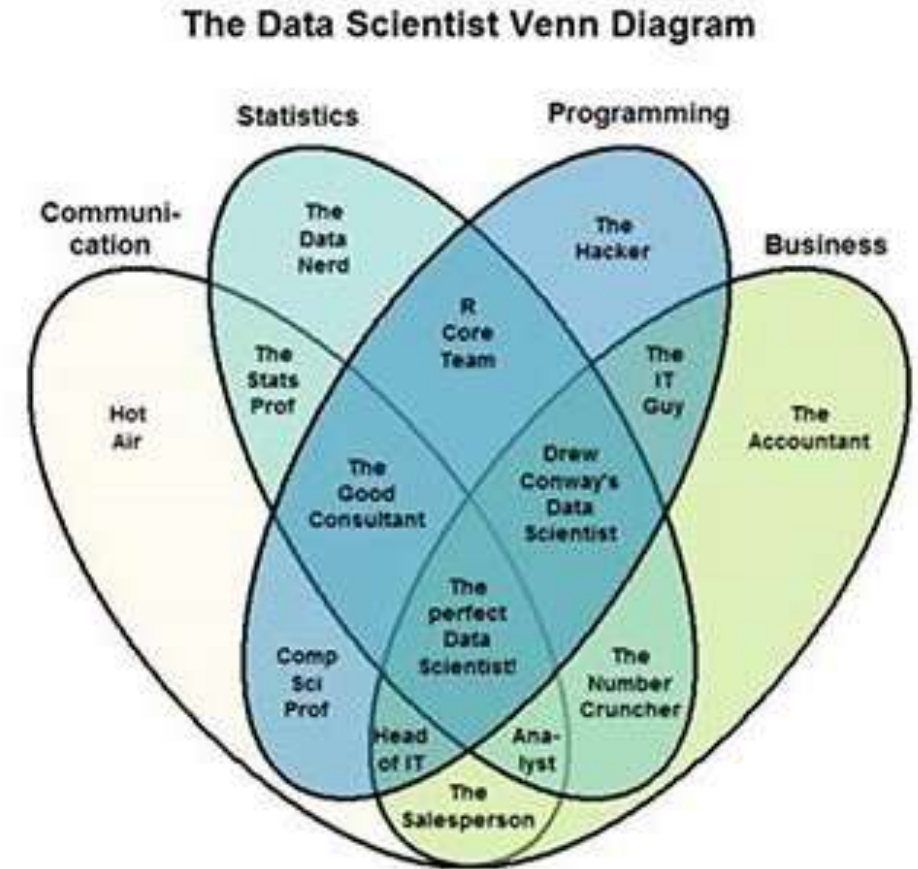
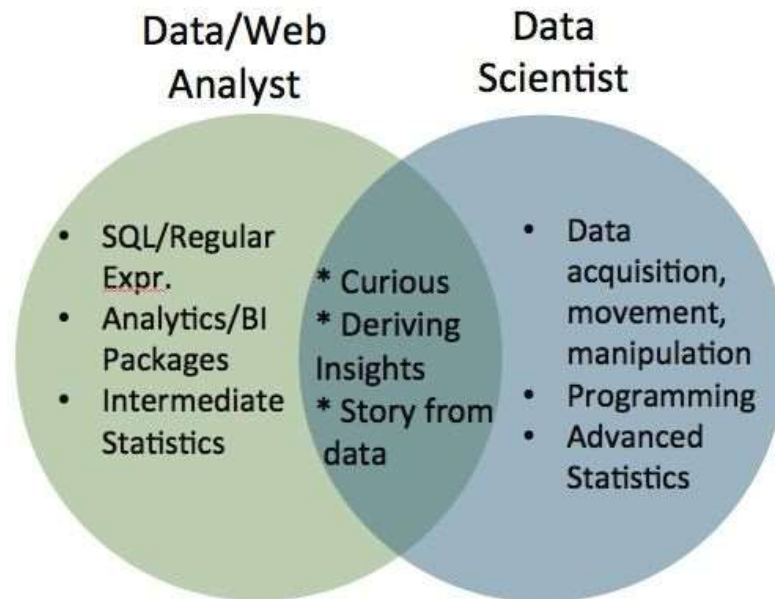


(un progetto Regione Toscana, CNR, AGENAS):

Una piattaforma di big data analytics con lo scopo di produrre strumenti di analisi unificati sui dati amministrativi legati alla salute in possesso delle unità regionali.

Nuove figure professionali

- C'era una volta l'analista di dati...
il ruolo del **Data Scientist**
- C'era una volta il project manager...
il ruolo del **Data System Integrator**



Big data step by step

1	Esaminare il problema che vorrei affrontare con i big data
2	Definire quali big data sono utili
3	Acquisire nuove competenze tecnologiche di gestione
4	Sviluppare nuove abilità statistiche, previsionali, intuitive, tecnologiche
5	Gestire la conoscenza di Business Intelligence
6	Decidere come mantenere i dati
7	Modificare la cultura dell'organizzazione
8	Preparare un piano di integrazione

Recap

- Siamo ancora nell'**adolescenza** dei Big Data
- Straordinarie **possibilità** di utilizzi futuri e nuove **responsabilità**
- **Condividere** esperienze (e fallimenti) è fondamentale
- Gli strumenti da soli non bastano, sono importanti le nuove **figure professionali**

Grazie per l'attenzione



Laura Spinsanti

Data Scientist per VITA S.p.A.

Integrazione e analisi di dati da fonti eterogenee.

l.spinsanti@vita.it

Riferimenti

- City Data Exchange: il caso di Copenhagen
 - Sito <https://www.citydataexchange.com/#/home>
 - Video <https://www.hitachiinsightgroup.com/en-us/smart-cities.html>
- Carlo Ratti @ SENSEable City Lab, MIT (pubblica su VITA per l'inserito RI-DESIGN)
 - Lezione alla Camera dei deputati 24/03/2016 <https://www.youtube.com/watch?v=6kniEkO3tYQ>
- United Nations Global Pulse
 - <http://www.unglobalpulse.org/about-new>
- EU nell'ambito Horizon 2020 (un esempio):
 - Big Data Europe: <http://www.big-data-europe.eu/>
- Big data in Italia
 - <http://www.sciencedirect.com/science/article/pii/S2095809916309444> molto interessante su cosa si fa sul territorio Italiano
- TheMatrix presentazione
 - <http://thematrix.isti.cnr.it/>
 - <http://aie.volonet.it/wp-content/uploads/2015/11/Gini-AIE-Milano-20151.pdf>
- Google Flu
 - <https://www.wired.com/2015/10/can-learn-epic-failure-google-flu-trends/>
- La privacy
 - <http://owlitalia.com/la-privacy-ai-tempi-dei-big-data>