

# MACHINE LEARNING E STRUMENTI ANTICRIMINE NELL'ERA DEI BIG DATA

Gaetano Bruno Ronsivalle

Università degli Studi di Verona  
Tecnologie informatiche e multimediali



**BANCHE E SICUREZZA 2015**

 #banchesicurezza

Roma - Palazzo Altieri  
4<sup>th</sup>/5<sup>th</sup> June

# Argomenti

**Big Data: una galassia online**

**Machine Learning: la nuova alleanza**

**Big Data, Machine Learning  
& Anticrimine**



# Argomenti

**Big Data: una galassia online**

**Machine Learning: la nuova alleanza**

**Big Data, Machine Learning  
& Anticrimine**



# Cosa avviene ogni minuto sul Web?

**195.000** App download  
**72 ore** Video → YouTube  
**348.000** tweet  
**+83.000 \$** → Amazon  
**4 milioni** ricerche → Google  
**7 milioni** messaggi → FB



# Big Data: una galassia online

Quantità di dati  
generati ogni giorno online:

**2,5 miliardi di gigabytes!!!**



# Big Data & Complessità = $V^4$

**V**olume = quantità dati

**V**ariety = differenti formati

**V**elocity = rapidità flussi

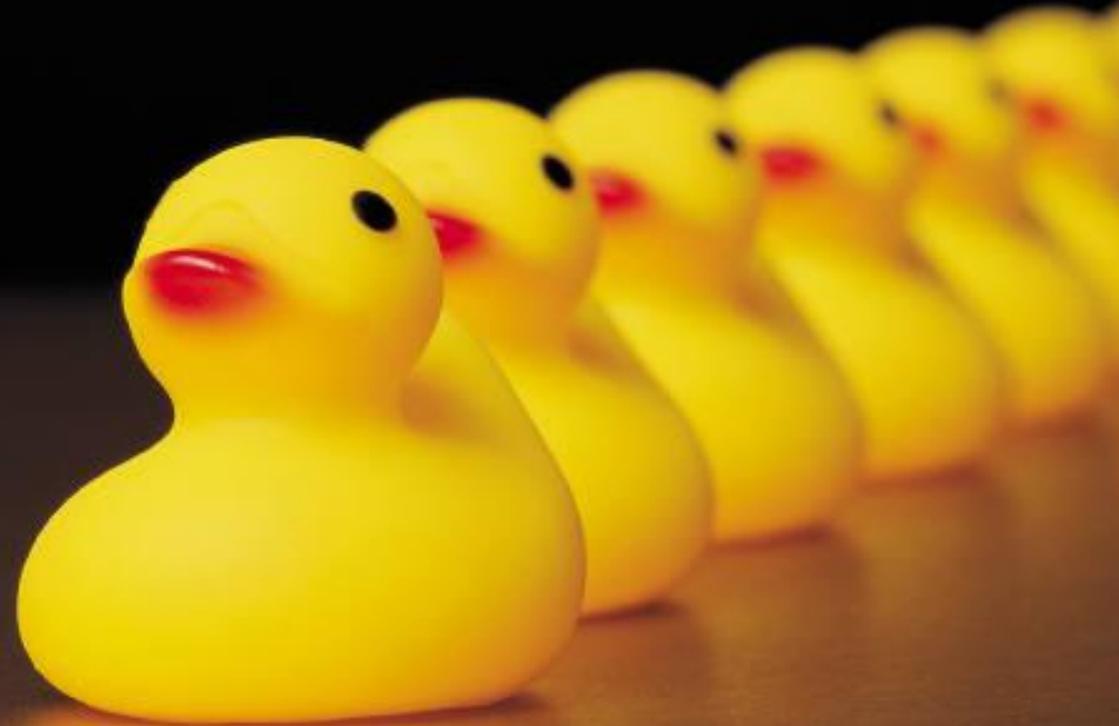
**V**eracity = qualità dati



# Volume: quantità di dati

**2020**

Crescita esponenziale:  
dati 2005 x 300 volte!!!



# Variety: diversi formati e tipologie

Dati strutturati  
Dati semi-strutturati  
Dati non strutturati  
Testi  
Immagini  
Audio  
Video



**Velocity = velocità di generazione dati**

**2016**

2,5 connessioni  
per abitante del pianeta Terra

Gestione dei flussi di dati

Real Time

Aggiornamenti



Veracity = qualità dei dati

Da dove vengono i dati online che utilizziamo per prendere decisioni?

Il problema delle **fonti**



# Dalla molteplicità di DB relazionali ...

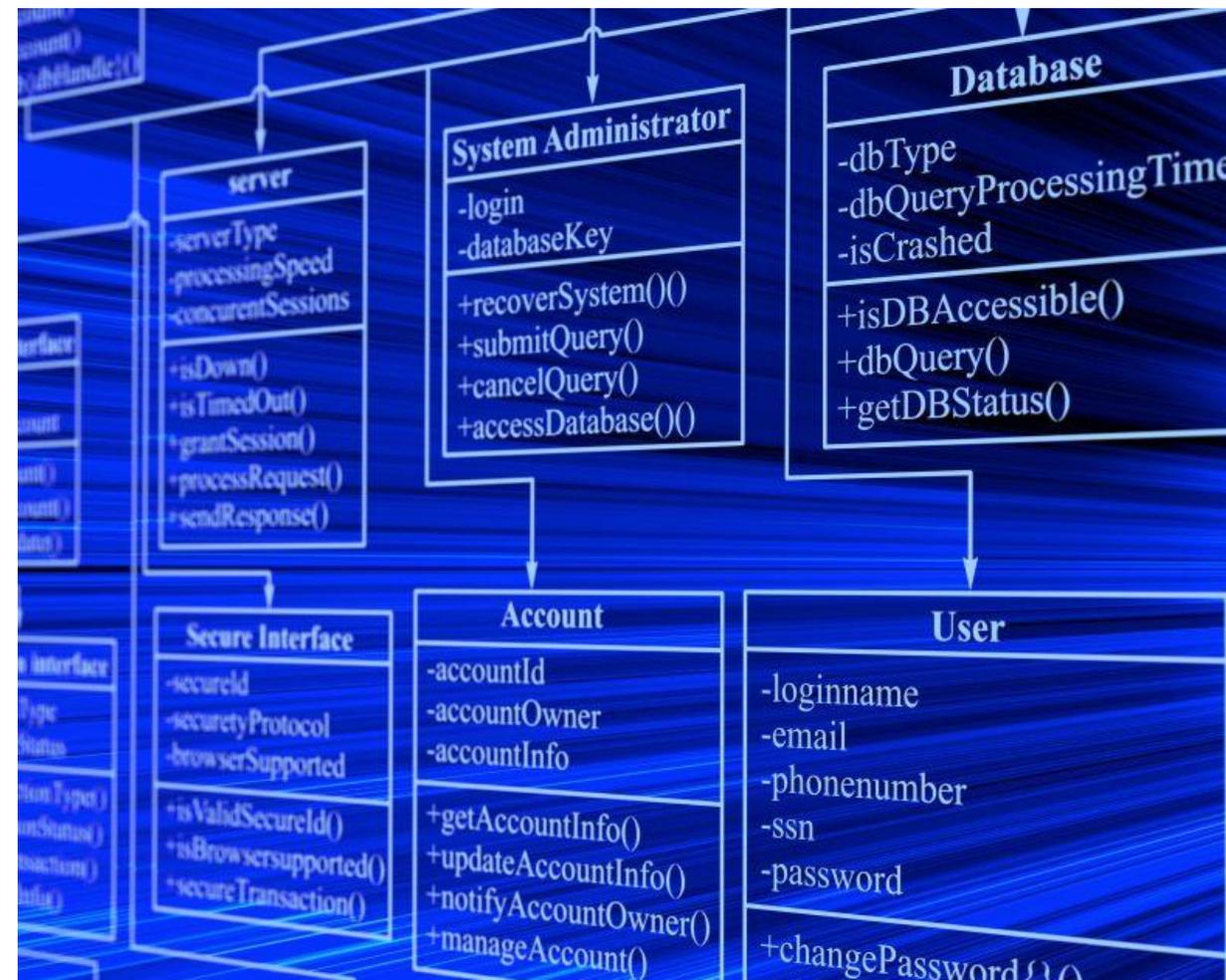
Database relazionale:

Campi

Record

Chiavi primarie

Mappe di relazioni

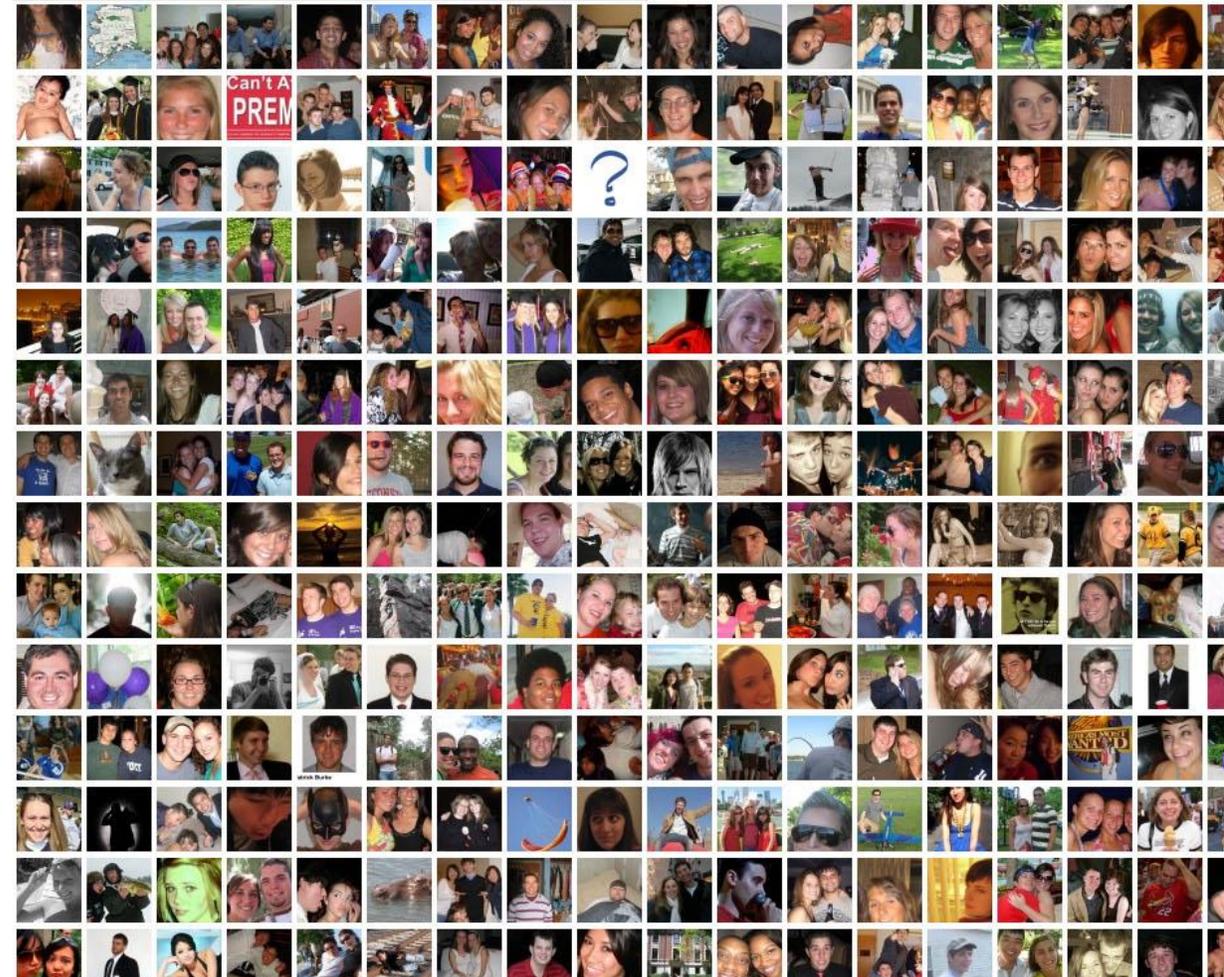




# I dati «nascosti» nelle applicazioni social

Alla ricerca di un **filo rosso**:

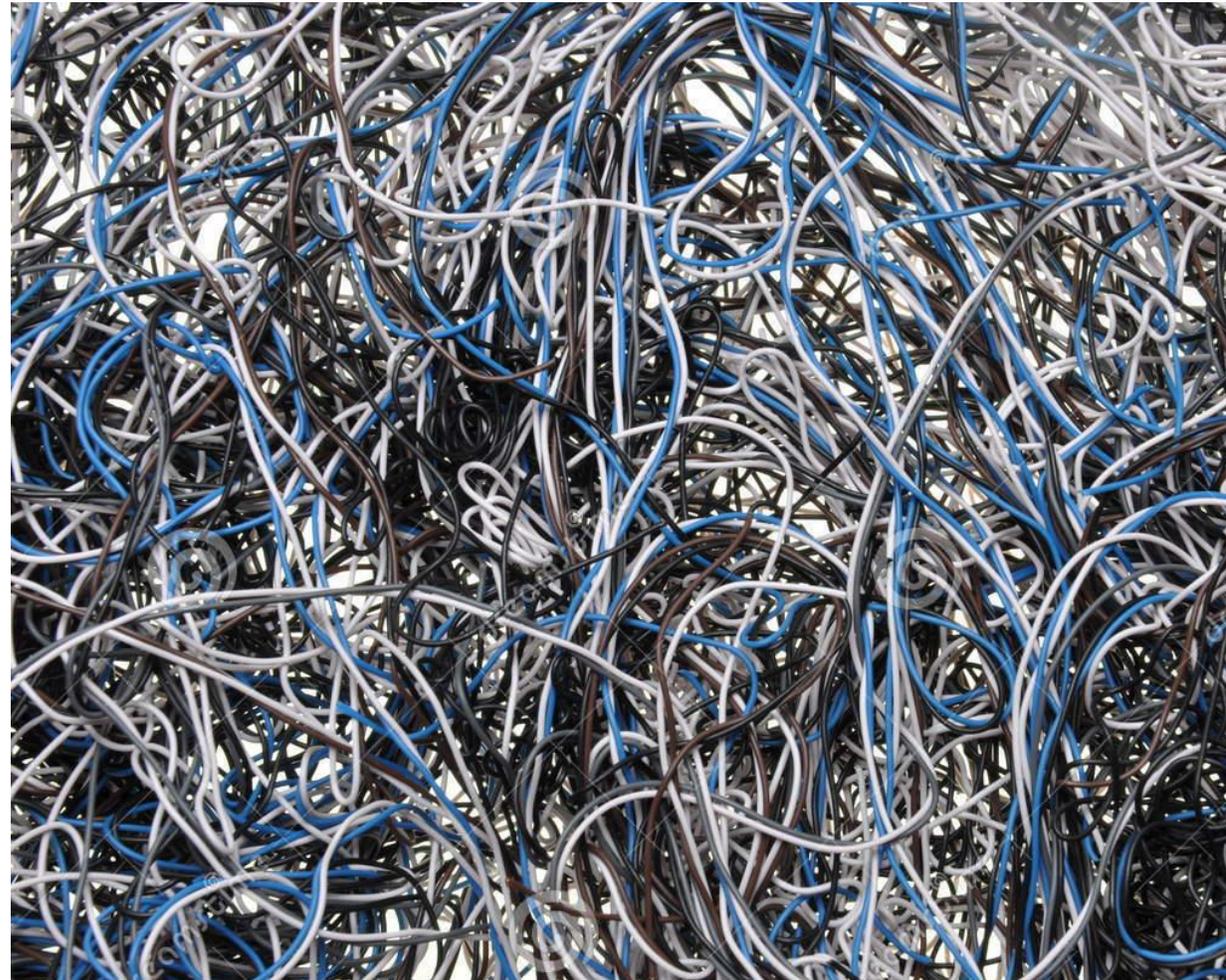
Facebook, Twitter, LinkedIn, ...



Dai «Big Data» ...

Come **aggredire** i Big Data?

Modalità non convenzionali  
e innovative  
di gestione,  
elaborazione  
e analisi dei dati



... alle «informazioni»

Come passare  
dai dati alle **informazioni**  
dotate di **senso**?



# Argomenti

**Big Data: una galassia online**

**Machine Learning: la nuova alleanza**

**Big Data, Machine Learning  
& Anticrimine**



# Big Data & Machine Learning

La nuova alleanza:

**il pensiero meccanico** come  
punto di accesso  
ai Big Data

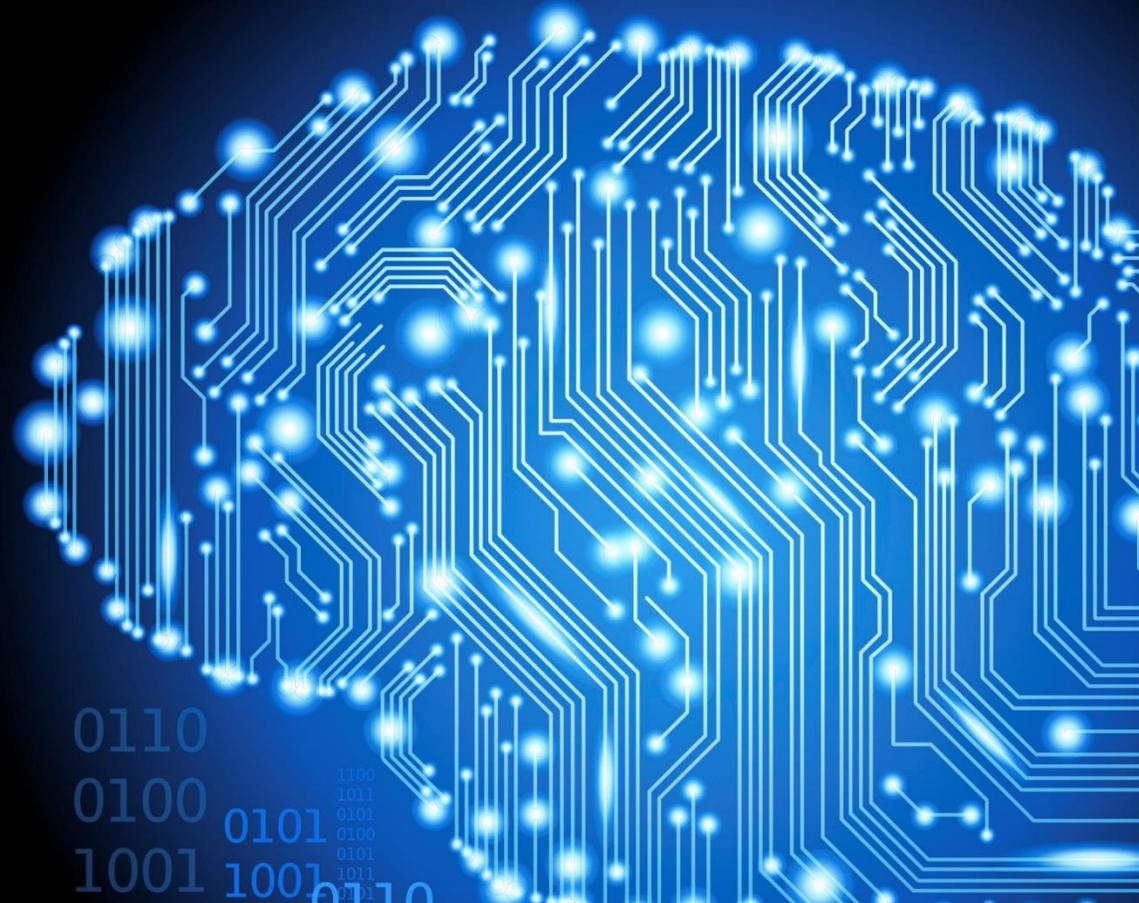


# Machine «Learning»

Macchine che **apprendono**

Reti Neurali Artificiali

Mappe di Kohonen

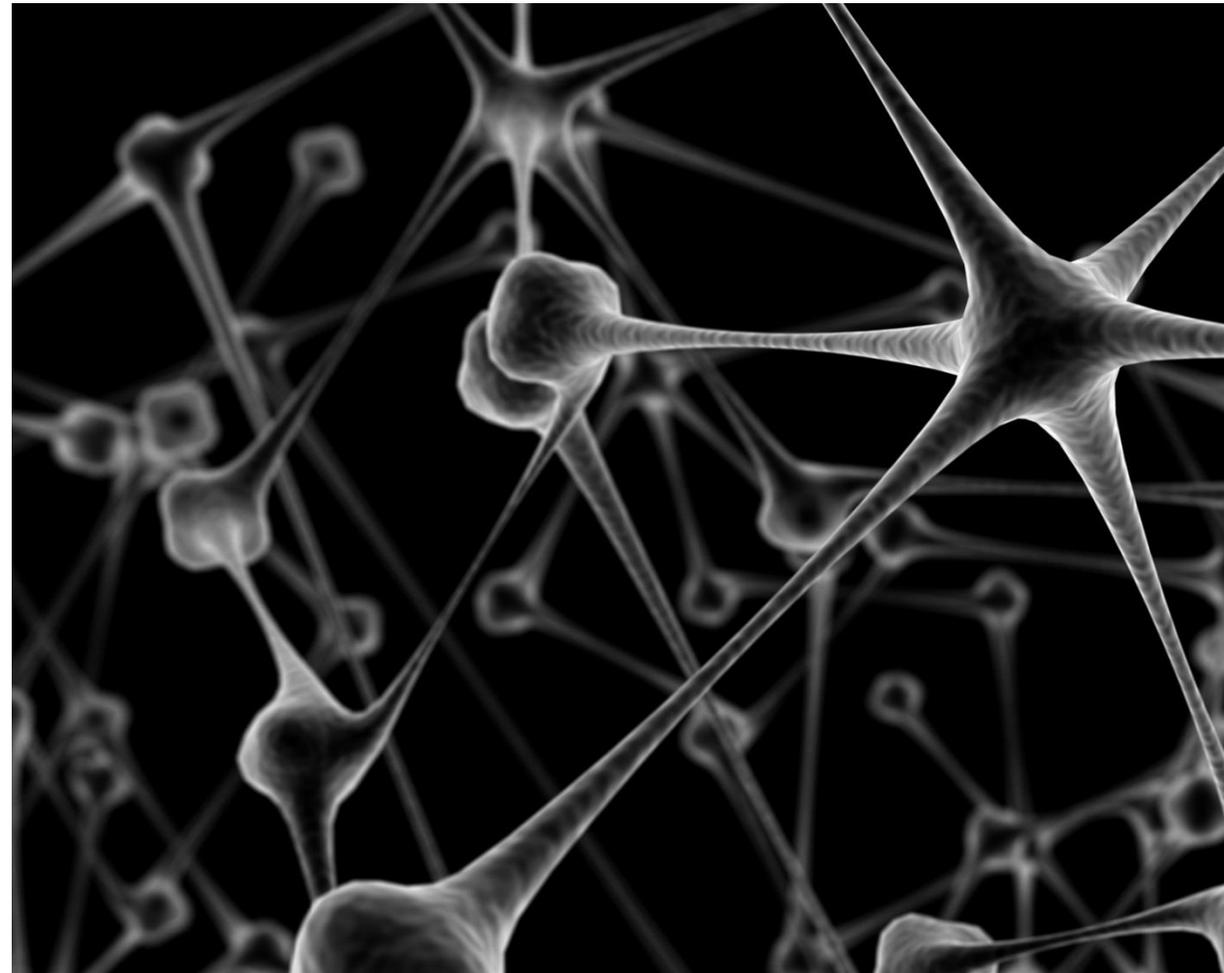


# Reti Neurali Artificiali

Strumenti hardware  
e/o software

**Apprendimento  
supervisionato**

Simulazione/previsione



# Mappe di Kohonen

Topologie neurali  
multidimensionali

**Cluster Analysis  
& Data filtering**

Classificazione e  
generalizzazione



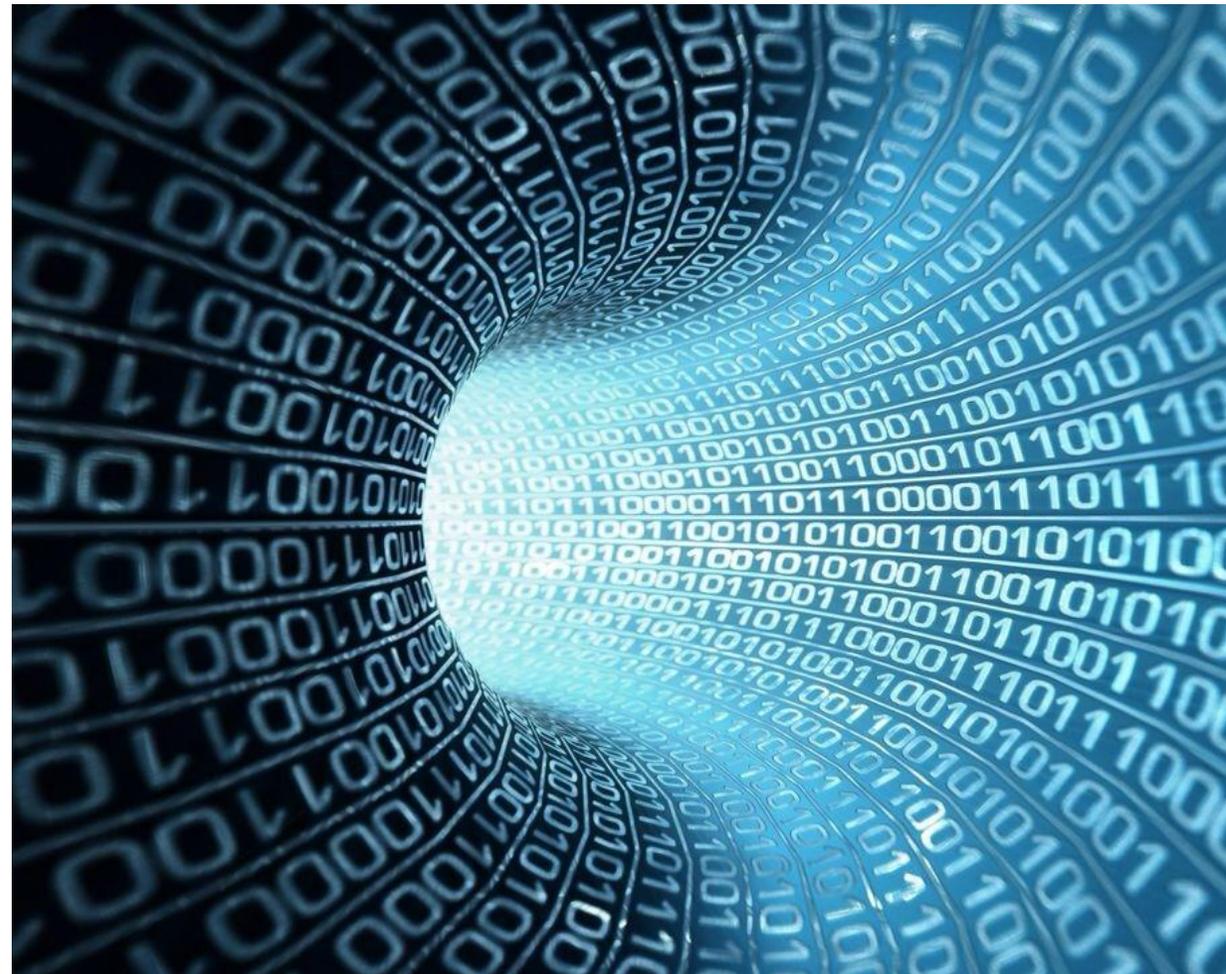
# Machine Learning vs. V<sup>4</sup> (Big Data)

**V**olume = algoritmi neurali

**V**ariety = visione reticolare

**V**elocity = parallelismo

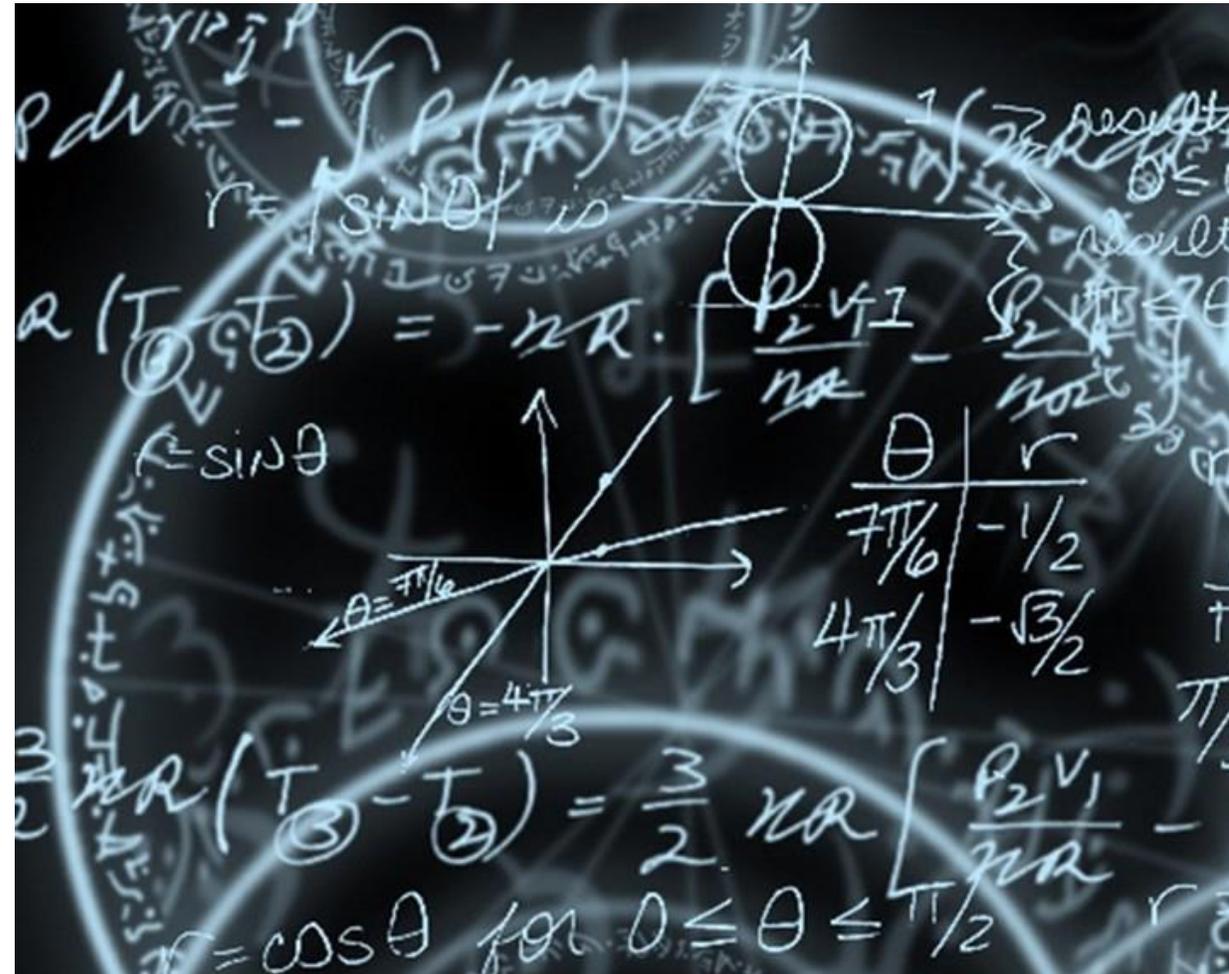
**V**eracity = data filtering



# Elementare Watson!

## Super Computer IBM

Interazione uomo-macchina  
mediante  
linguaggio naturale



# Argomenti

Big Data: una galassia online

Machine Learning: la nuova alleanza

Big Data, Machine Learning  
& Anticrimine



# Analisi anticrimine

**Classificazione**  
di dati eterogenei

**Simulazione**  
di funzioni non lineari

**Previsione**  
di fenomeni complessi



# Big Data, Machine Learning e anticrimine

*Rapine in banca*  
*Attacchi agli impianti ATM*  
*Terrorismo*  
*“Lupi solitari”*  
*Criminalità organizzata*  
*E-crime*



# Rapine in banca

## Analisi del rischio

DB Ossif = relazionale

Architettura neurale MLP  
simulazione/previsione



# Attacchi agli impianti ATM

## Analisi del rischio

DB Ossif = relazionale

Architettura neurale ibrida  
simulazione/previsione

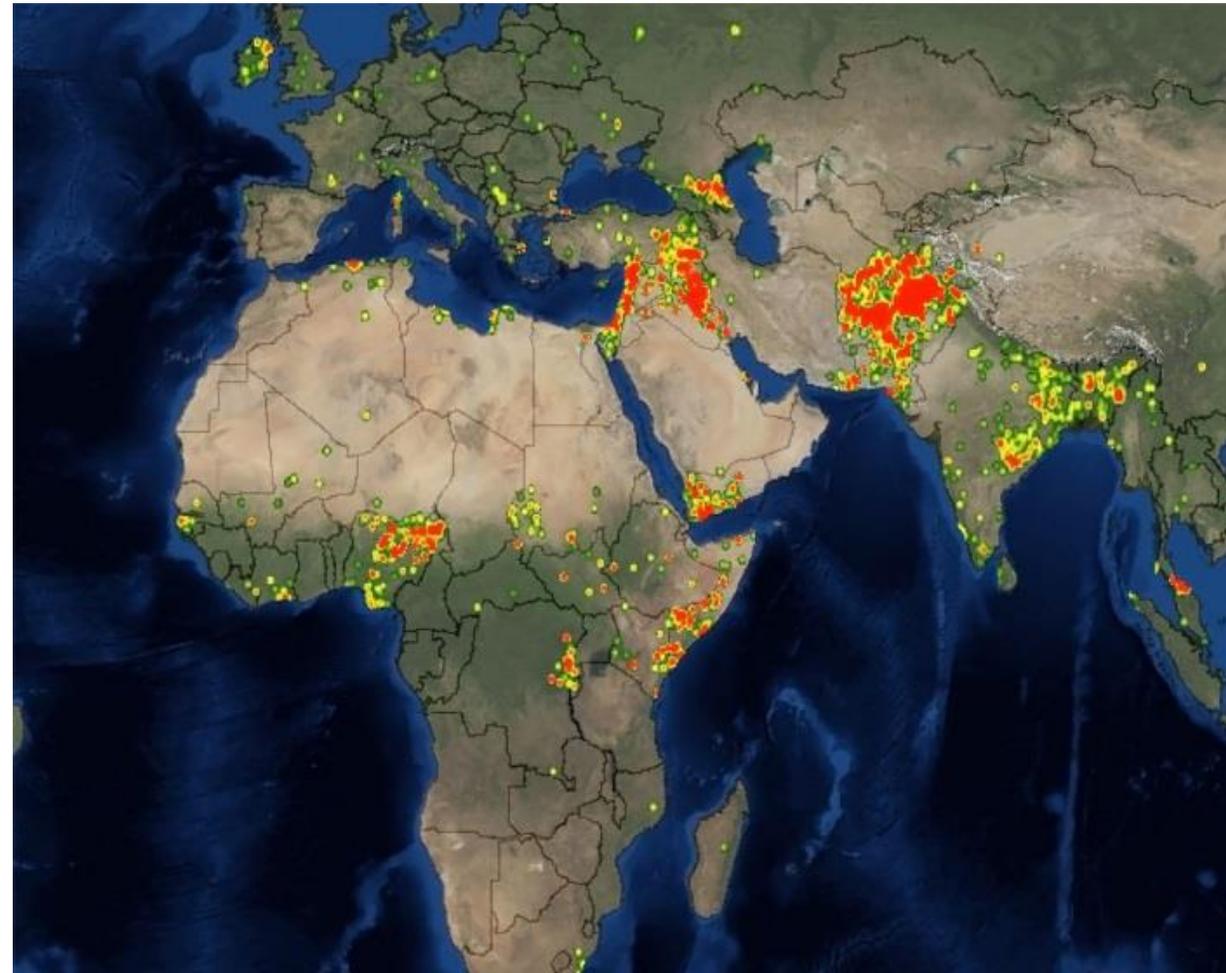


# Terrorismo

## Geo-Network Analysis

DB non relazionali:  
text mining da Web/Social

Machine Learning →  
previsione georeferenziata



**“Lupi solitari”**

## **Geo-Analysis**

DB non relazionali:  
text mining da social app

Machine Learning →  
previsione geo-referenziata



# Criminalità organizzata

## Geo-Network Analysis

DB non relazionali:  
text mining da intercettazioni

Machine Learning →  
previsione evoluzione rete



## Analisi del rischio

DB relazionali e non:  
data/text mining

Machine Learning →  
simulazione/previsione



# Il grande diluvio di dati

## Big Data

=

*“datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze”*

(McKinsey Global Institute 2011)



# Data Science

La nuova figura professionale  
dello «**scienziato dei dati**»  
per governare il grande diluvio

Gestione, elaborazione, archiviazione,  
analisi e trasformazione dei Big Data



# Argomenti

**Big Data: una galassia online**

**Machine Learning: la nuova alleanza**

**Big Data, Machine Learning  
& Anticrimine**



[gaetanobruno.ronsivalle@univr.it](mailto:gaetanobruno.ronsivalle@univr.it)

# MACHINE LEARNING E STRUMENTI ANTICRIMINE NELL'ERA DEI BIG DATA



**BANCHE E SICUREZZA 2015**

 #banchesicurezza

Roma - Palazzo Altieri  
4<sup>th</sup>/5<sup>th</sup> June