Processi Diffusivi e Reti Sociali

Giulio Rossetti, Letizia Milli, Salvatore Citraro
KDD Lab., UniPi & ISTI-CNR

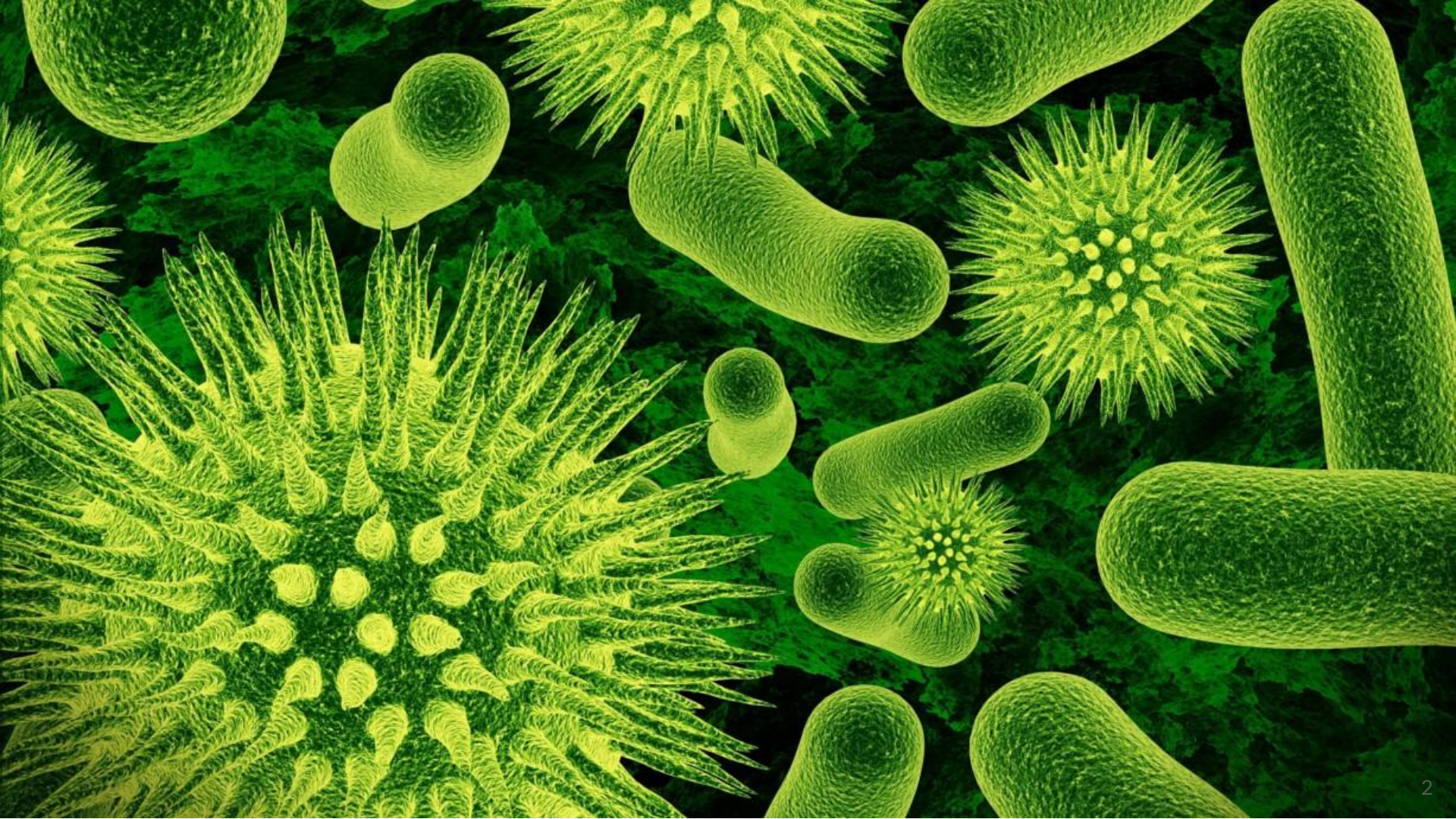
@GiulioRossetti Giulio.Rossetti@isti.cnr.it

SoBigData

ISTITUTO DI SCIENZA E TECNOLOGIE DELL'INFORMAZIONE "A. FAEDO"

IN SUPREMÆ DIGNITATIS 1343 UNIVERSITÀ DI PISA

KDD Lab




SAMSUNG


Strategy Innovation
Vision
Creativity
Support
Solutio

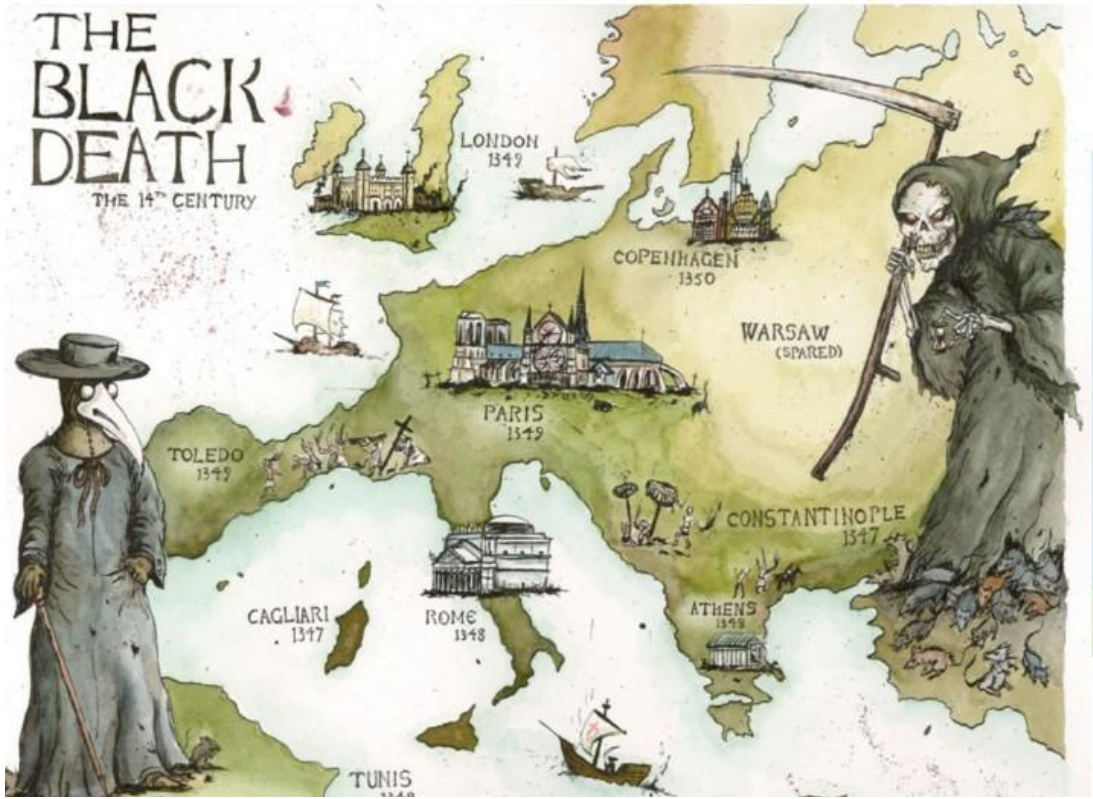

INNOVATION
BUSINESS
PLAN
IDEA

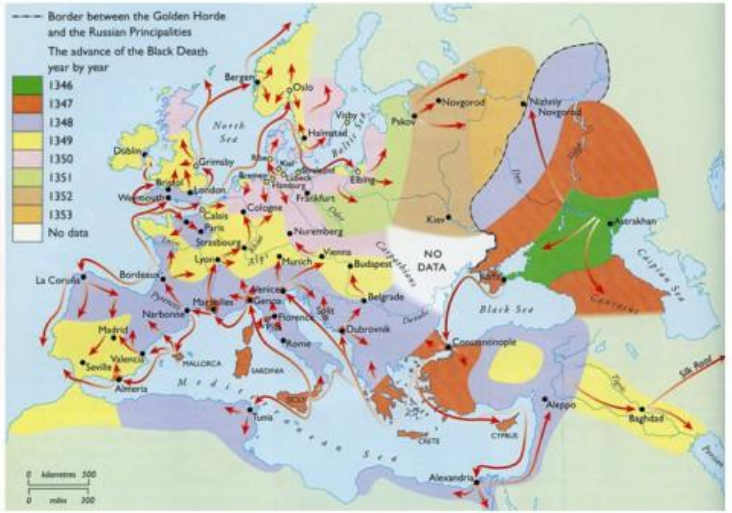
Perché è importante studiare fenomeni diffusivi?

Elevata Mobilità

Elevata Densità

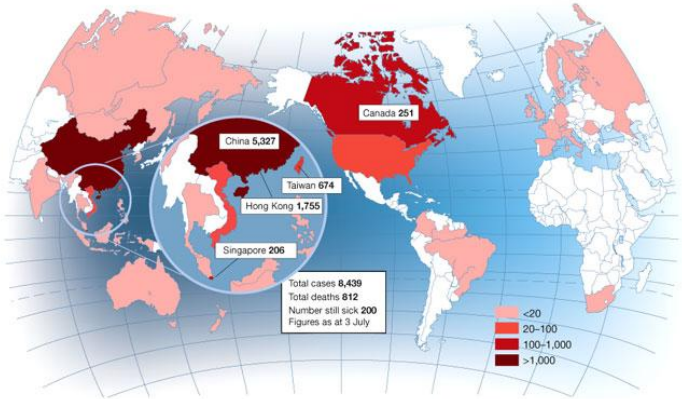
La “Grande” Peste (14° secolo)

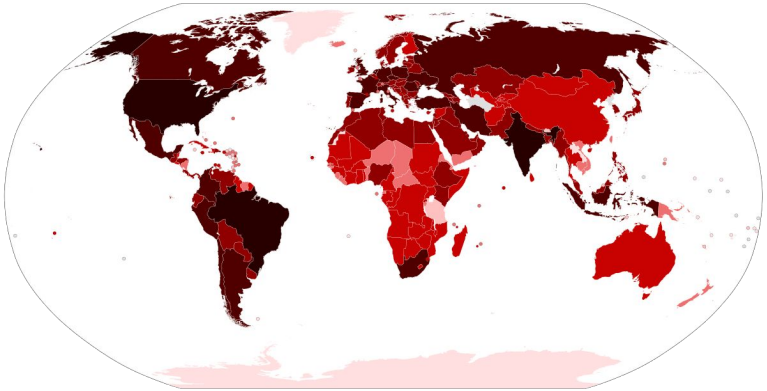




4 anni dalla Francia alla Svezia

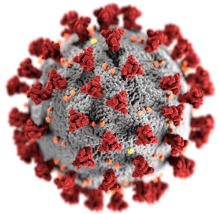
SARS & Covid19



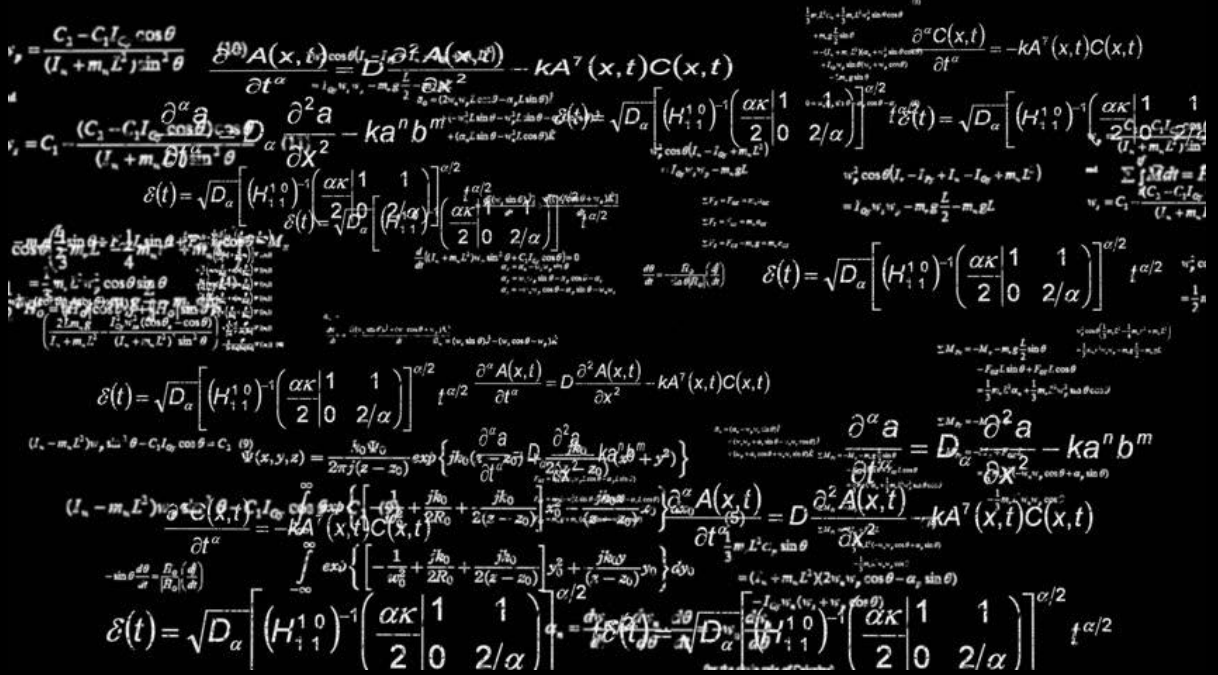


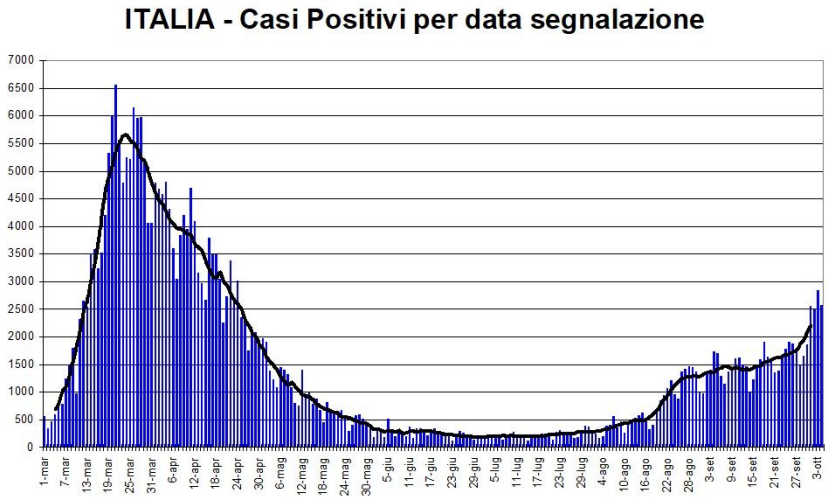
6 mesi..

..2 mesi

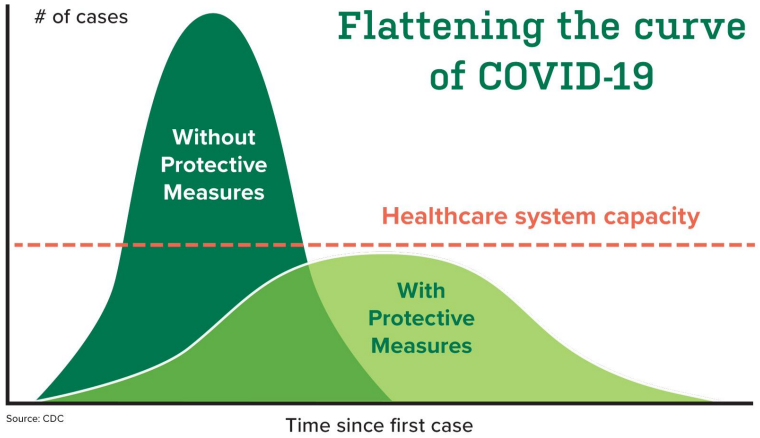


Come descrivere un fenomeno diffusivo?

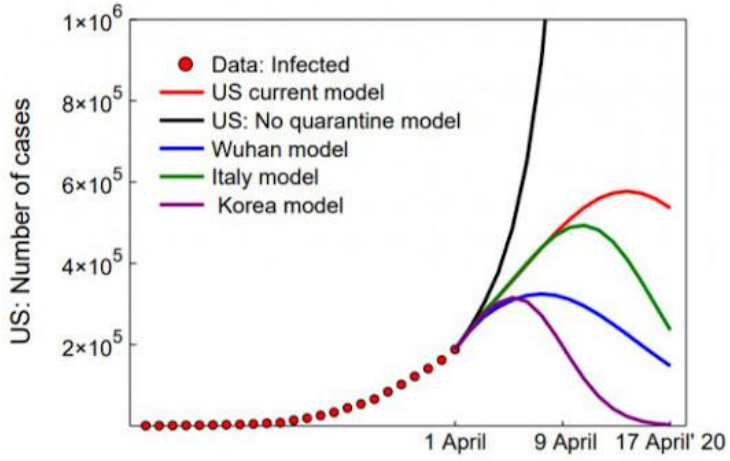




Dati



Obiettivo



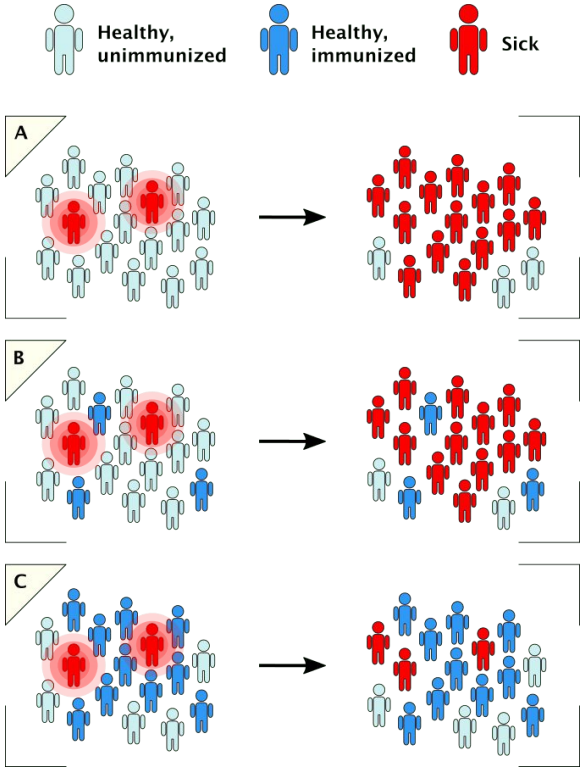
Modelli

$$\begin{cases} \dfrac{dS}{dt} = -aSI, \quad S(0) = S_0 > 0, \\ \dfrac{dI}{dt} = aSI - bI, \quad I(0) = I_0 > 0, \\ \dfrac{dR}{dt} = bI, \quad R(0) = R_0 \geq 0. \end{cases}$$

Una prima approssimazione: Mean Field

Data una popolazione si assume:

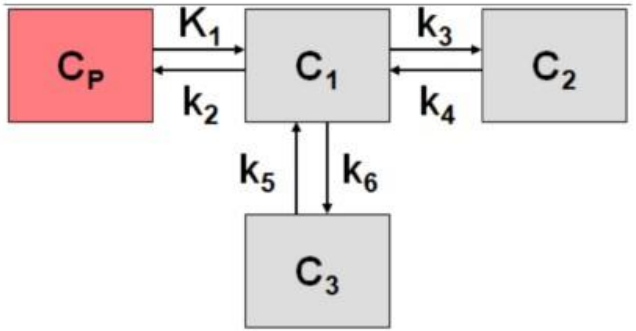
- Perfect mixing: tutti gli individui che ne fanno parte hanno la stessa probabilità di entrare in contatto



Modelli a “Compartimenti”

Ipotesi:

Compartimentalizzazione:
Ad ogni individuo è associato uno stato
(e.g., Suscettibile, Infetto, Rimosso)

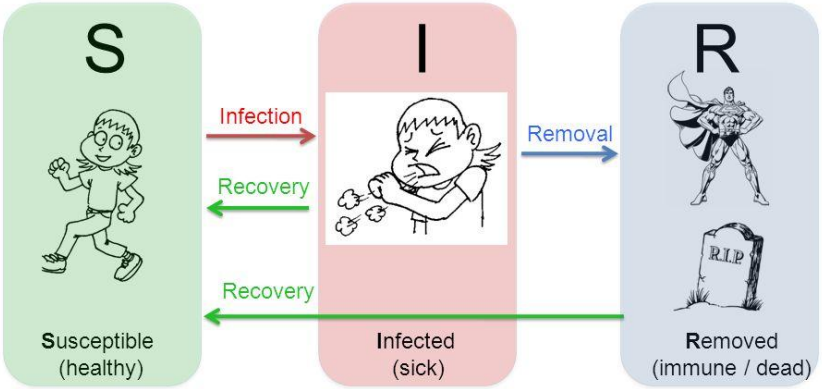


W. O. Kermack and Ag McKendrick. A Contribution to the Mathematical Theory of Epidemics. 1927

Esempi:

Modelli SI/SIS/SIR

$$\begin{cases} \dfrac{dS}{dt} = -aSI, \quad S(0) = S_0 > 0, \\ \dfrac{dI}{dt} = aSI - bI, \quad I(0) = I_0 > 0, \\ \dfrac{dR}{dt} = bI, \quad R(0) = R_0 \geq 0. \end{cases}$$



2 (3) compartimenti:

Suscettibile/Infetto/(Rimosso)

Definizione di regole (probabilistiche) per passare dall'uno all'altro

Assumere un contesto “mean field” è realistico?

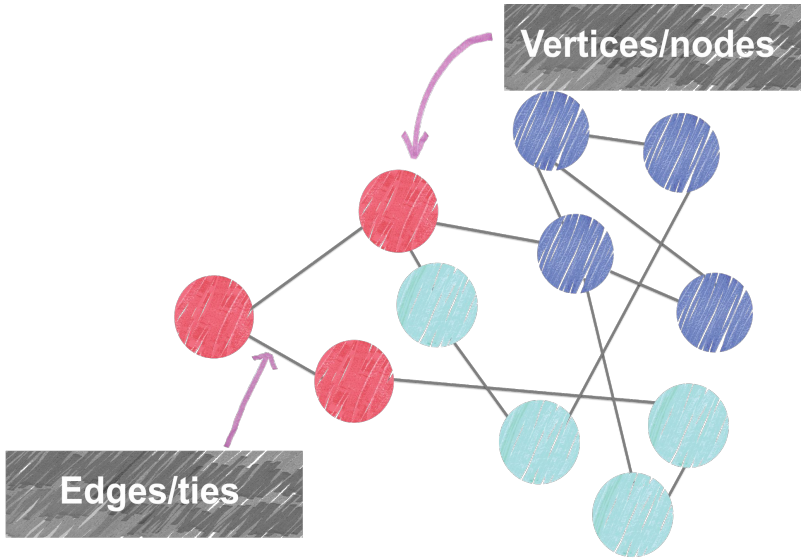
"Behind each complex system
there is a network,
that defines the interactions
between the components."

Keith Shepherd's "Sunday Best . http://baseballart.com/2010/07/shades-of-greatness-a-story-that-needed-to-be-told/

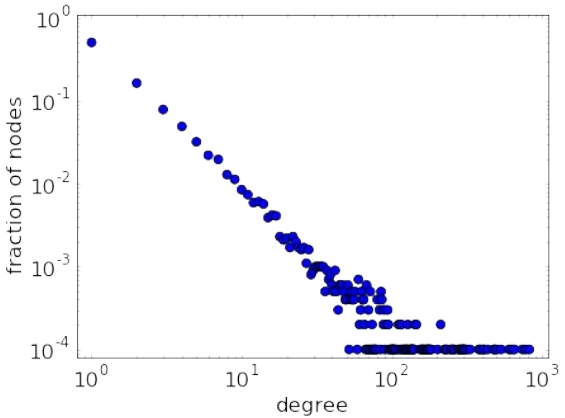
Social Networks
sono composte di...

- Individui (nodi)
- relazioni (sociali) tra di essi (archi)

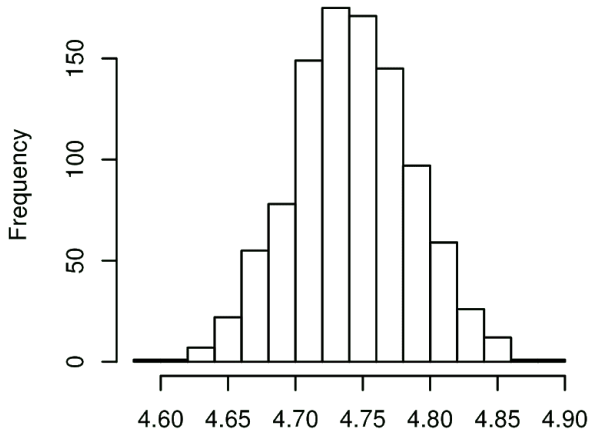
Il modo in cui gli individui sono connessi tra di loro è importante!



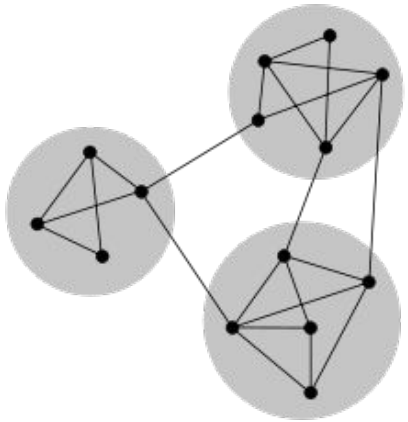
Social Networks
e le loro caratteristiche...



Eterogeneità nel numero delle connessioni



Piccolo diametro
"6 gradi di separazione"

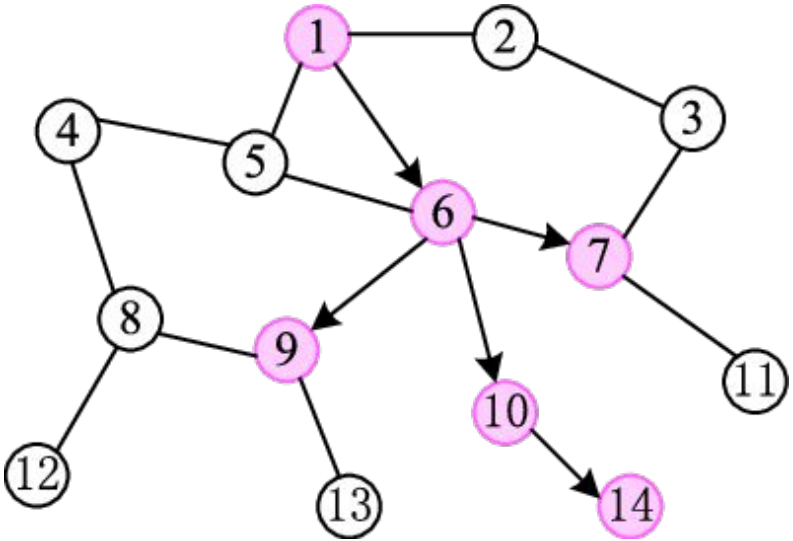


Esistenza di Comunità

Social Networks
… ed il loro impatto sulla diffusione

La velocità e le modalità di diffusione dipendono da:

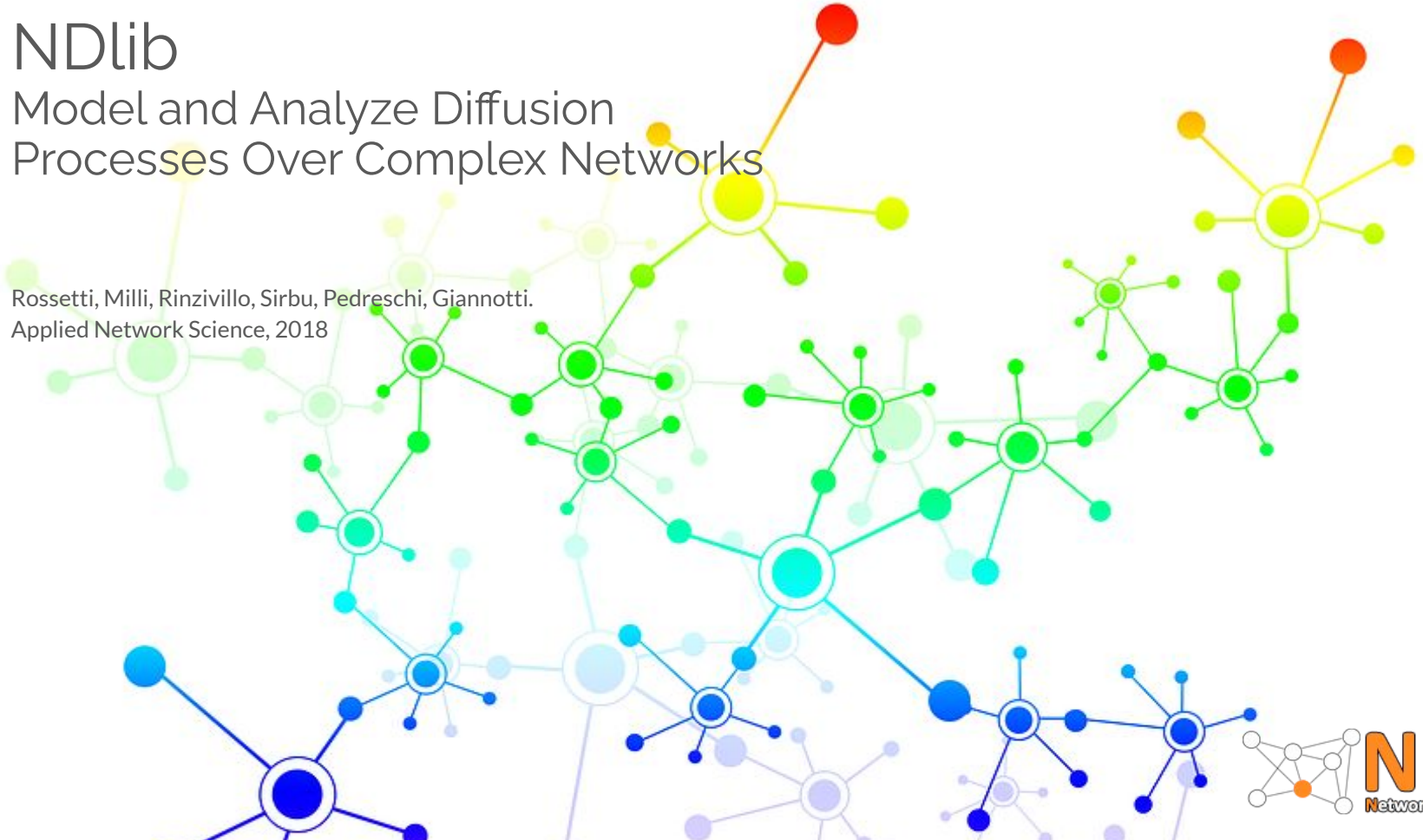
- modello usato per simularla
- Insieme iniziale degli individui infetti
- Struttura della rete sociale dei soggetti infetti
  (e.g., esistenza di comunità)
- …



NDlib

Model and Analyze Diffusion Processes Over Complex Networks

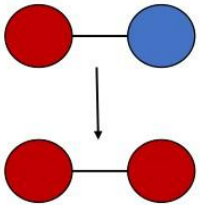
Rossetti, Milli, Rinzivillo, Sirbu, Pedreschi, Giannotti.
Applied Network Science, 2018



NDlib
Network Diffusion Library

Processi diffusivi e reti sociali

Un “processo diffusivo” ha luogo quando i portatori della malattia/virus/idea sono connessi a nodi “suscettibili”.



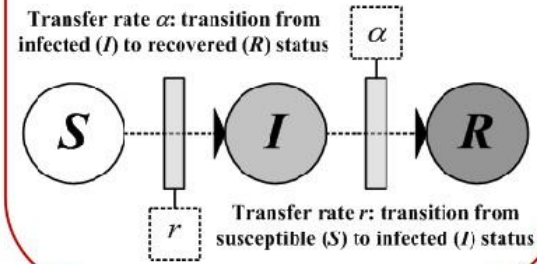
Cosa serve per descrivere un processo diffusivo?

- compartimenti
- regole di transizione tra compartimenti

Esempio
SIR model

Tre compartimenti
(**S**)uscettibile
(**I**)nfetto,
(**R**)imosso

Due regole di transizione:
S->I;
I->R

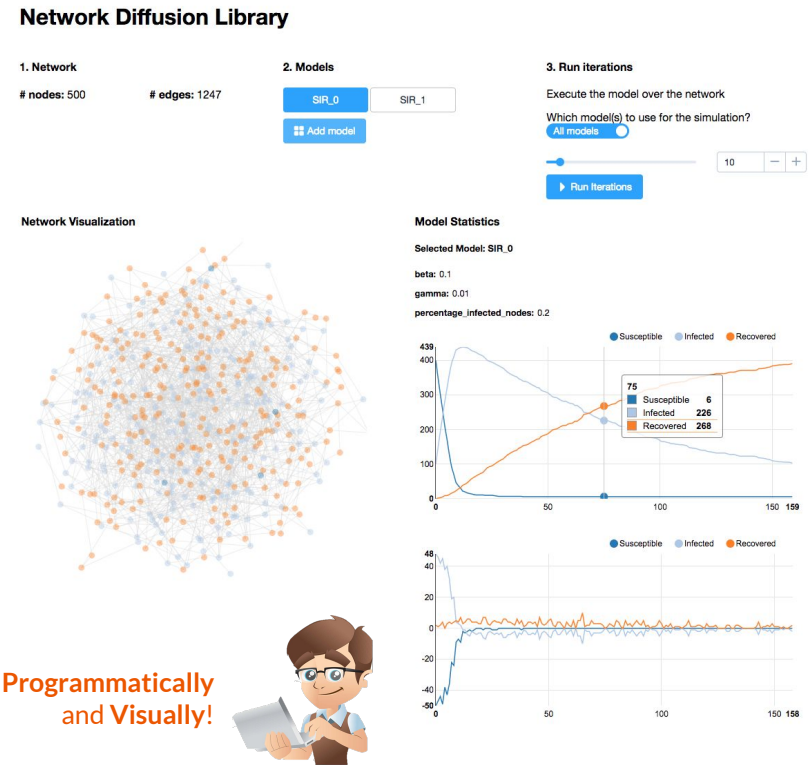


A Network Diffusion Framework!

NDlib
Network Diffusion Library

Simulate
Epidemics and
Opinion Dynamics
processes

Unfolding on top of
complex network structures



Programmatically
and **Visually**!

Available Models

Epidemics

(16 Models)

- SI / SIS / SIR
- SEIS / SEIR /SWIR
- Threshold / Generalized Threshold/ Profile / Profile-Threshold / Threshold-Blocked
- Independent Cascades /Independent Cascades with community
- Ising

Opinion Dynamics

(9 Models)

- Majority Rule
- Voter / Q-Voter
- Sznajd
- Cognitive Opinion Dynamics
- Algorithmic Bias
- Hegselmann-Krause / Weighted Hegselmann-Krause/ Attraction-Repulsione Weighted Hegselmann-Krause

SIR Code Example

Poche azioni semplici:

- ❏ Carichiamo la rete sociale (grafo)
- ❏ Selezioniamo il modello diffusivo
- ❏ Eseguiamo la simulazione

```
import networkx as nx
import ndlib.models.ModelConfig as mc
import ndlib.models.epidemics.SIRModel as sir

# Network topology
g = nx.erdos_renyi_graph(1000, 0.1)

# Model selection
model = sir.SIRModel(g)

# Model Configuration
cfg = mc.Configuration()
cfg.add_model_parameter('beta', 0.01)
cfg.add_model_parameter('gamma', 0.005)
cfg.add_model_parameter("percentage_infected", 0.05)
model.set_initial_status(cfg)

# Simulation execution
iterations = model.iteration_bunch(200)
```

Visual Analysis

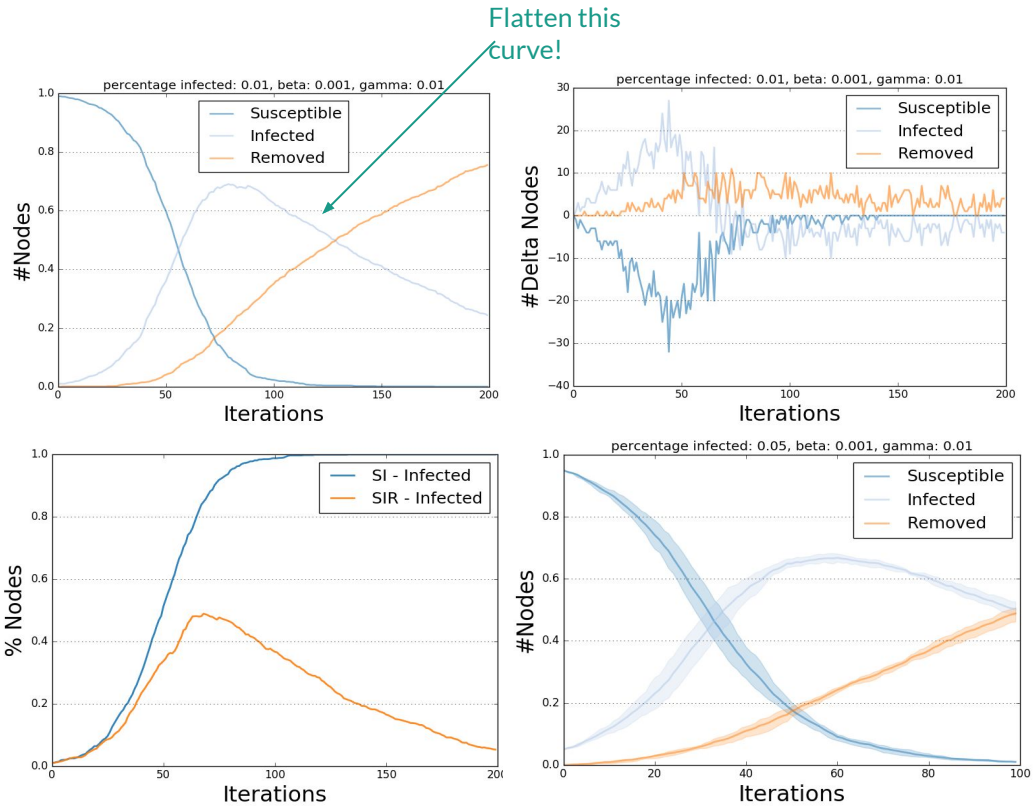
Come interpretare i risultati della simulazione?

Base Viz

- ❑ Diffusion Trends
- ❑ Incidence/Prevalence
- ❑ Opinion Dynamics

Advanced Viz

- ❑ Compare Models
- ❑ Multiple Run



Advanced Features

Definizione di modelli ad-hoc (composite)

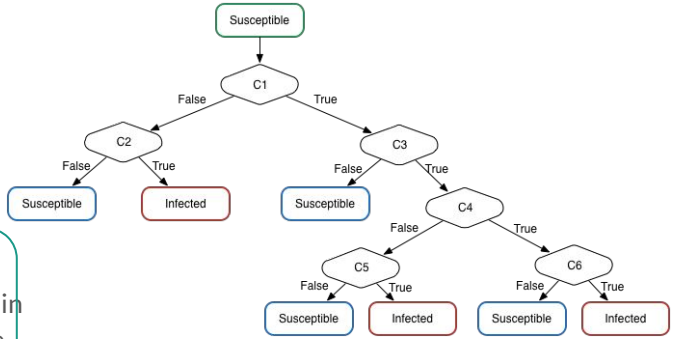
- Definire modelli di diffusione componendo in modo incrementale compartimenti e regole di transizione

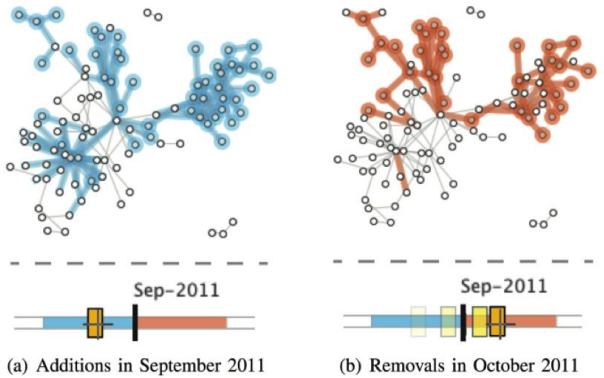
Support for Dynamic Network models

- Integration with DyNetX (ad-hoc library by CNR-UNIPI)

NDQL: Network Diffusion Query Language

- High-level query language for defining diffusion processes





(a) Additions in September 2011 (b) Removals in October 2011

```
CREATE_NETWORK g1
TYPE erdos_renyi_graph
PARAM n 300
PARAM p 0.1

MODEL SI

STATUS Susceptible
STATUS Infected

# Compartment definitions

COMPARTMENT c1
TYPE NodeStochastic
PARAM rate 0.1
TRIGGER Infected

# Rule definitions

RULE
FROM Susceptible
TO Infected
USING c1

# Model configuration

INITIALIZE
SET Infected 0.1

EXECUTE SI ON g1 FOR 100
```

When

Right now, NDlib v5.1
(codename: Enterovirus) is out!

Where

- Pypi:
  https://pypi.python.org/pypi/ndlib
- GitHub NDlib:
  https://github.com/GiulioRossetti/ndlib
- Documentation:
  http://ndlib.readthedocs.io/
- SoBigData:
  http://www.sobigdata.eu
- Tutorial:
  https://github.com/KDDComplexNetworkAnalysis/CNA_Tutorials

User Base	Research impact
~50k	35
Installations (2021-Q1 only)	Publications citing NDlib since 2018 (First release 12/2017)

Perché NDlib?

Definire analisi “What if”

- Quali sono gli individui da vaccinare per ridurre sensibilmente la velocità di un processo diffusivo?
- Chi devo far “parlare” per ridurre la polarizzazione su temi dibattuti?
- Quale è la risposta più probabile di una popolazione ad una nuova policy?
- ...

Definire e testare modelli ad-hoc per specifici fenomeni

- Misinformation
- Fake-news diffusion
- Flu-like illnesses
- Reactions to new laws
- ...

E ora, sporchiamoci le mani!