Webinar

SESSTIM UMR912

Sciences Economiques et Sociales de la Santé et Traitement de l'Information médicale

sesstim.univ-amu.fr

&

OpenHealth INSTITUTE

openhealth-institute.org

Jacques DEMONGEOT

Professeur, IUF, Institut Universitaire de France, Membre honoraire.
UGA, Université Grenoble Alpes Professeur Emérite.

Big data,
Functional modelling, Time Data.

juin 2017

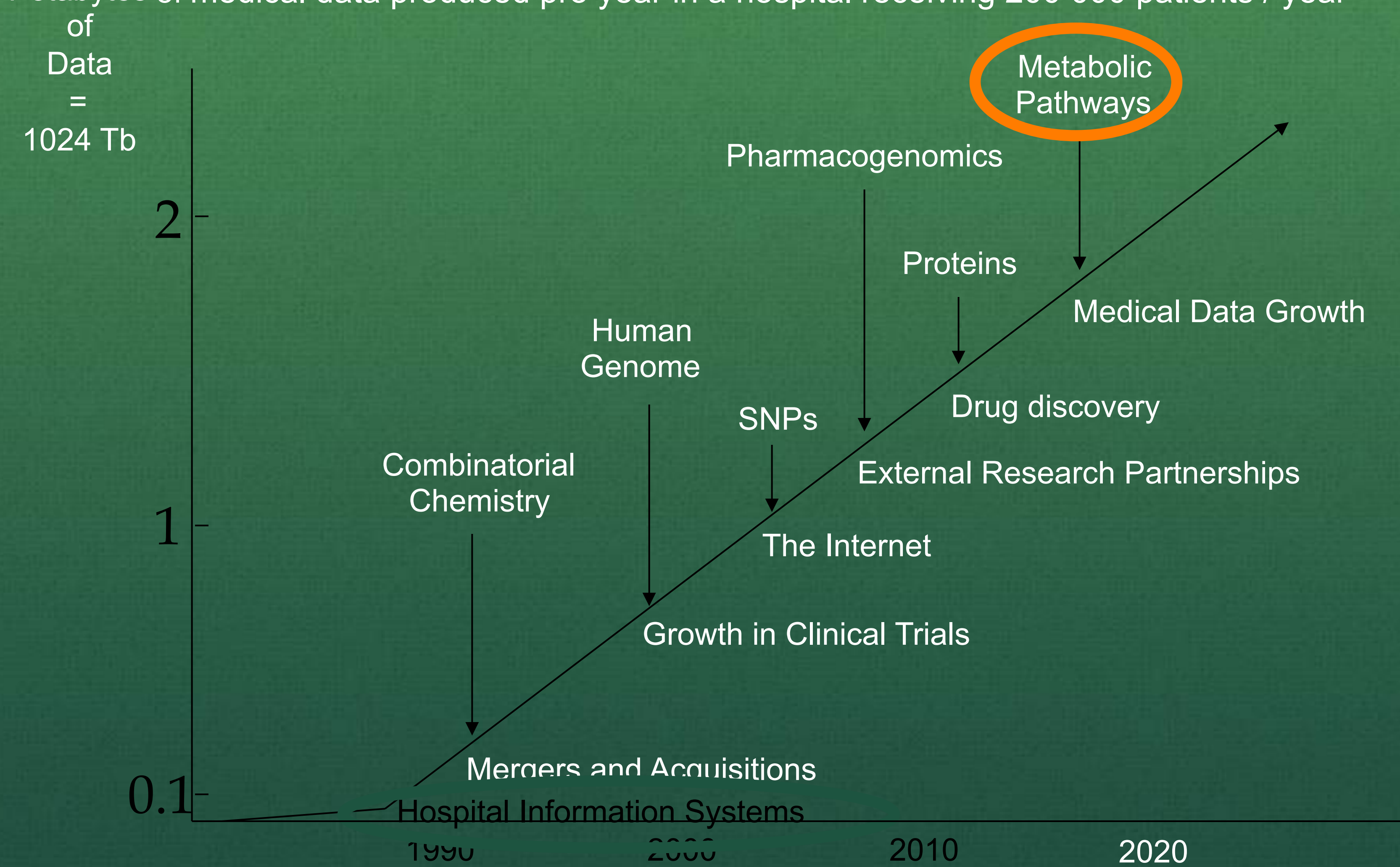
Cliquez ici pour voir l'intégralité des ressources associées à ce document

Big data, Functional modelling, Time data

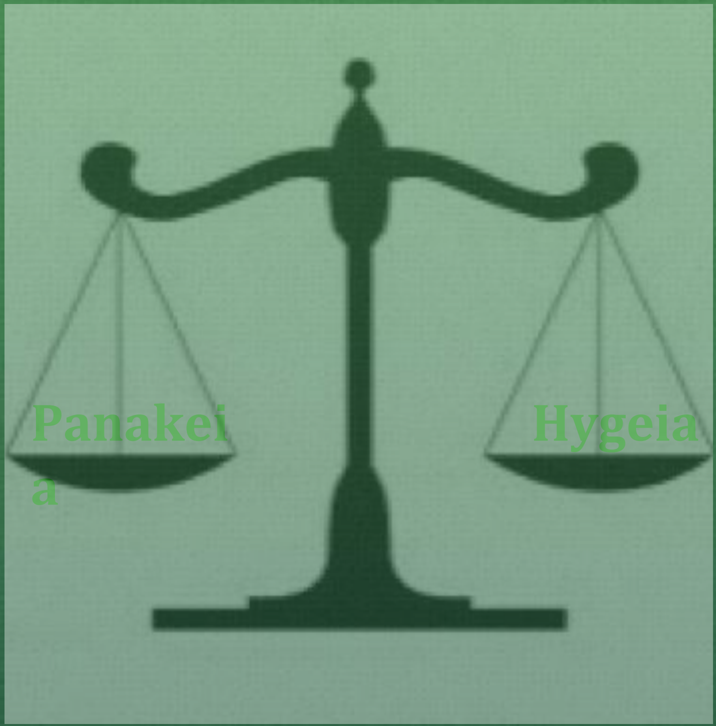
J. Demongeot IUF (honorary) & UGA (emeritus)

New Health Technologies Produce Data Tsunami

of medical data produced pro year in a hospital receiving 200 000 patients / year



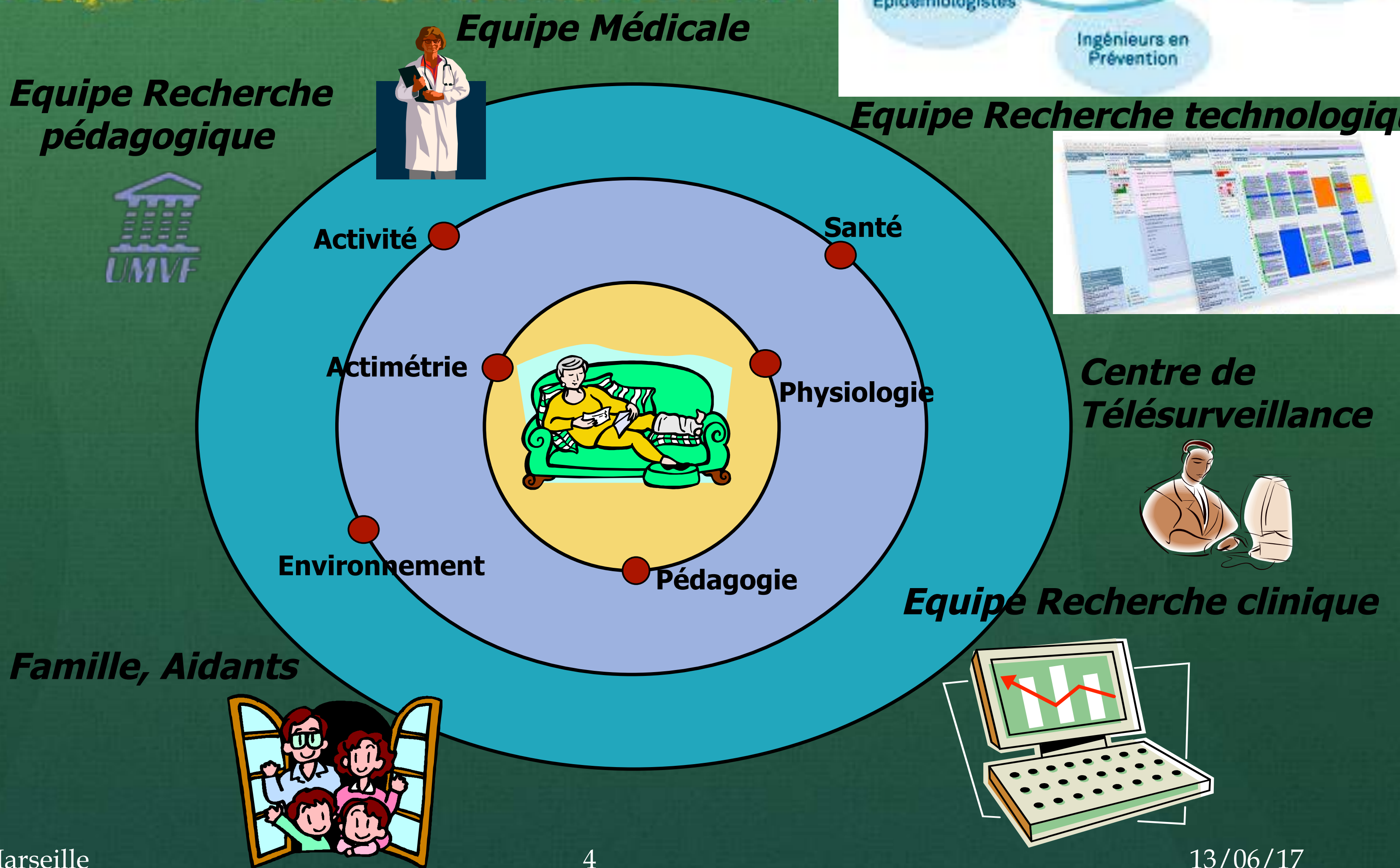
Source: Jeff Augen, IBM

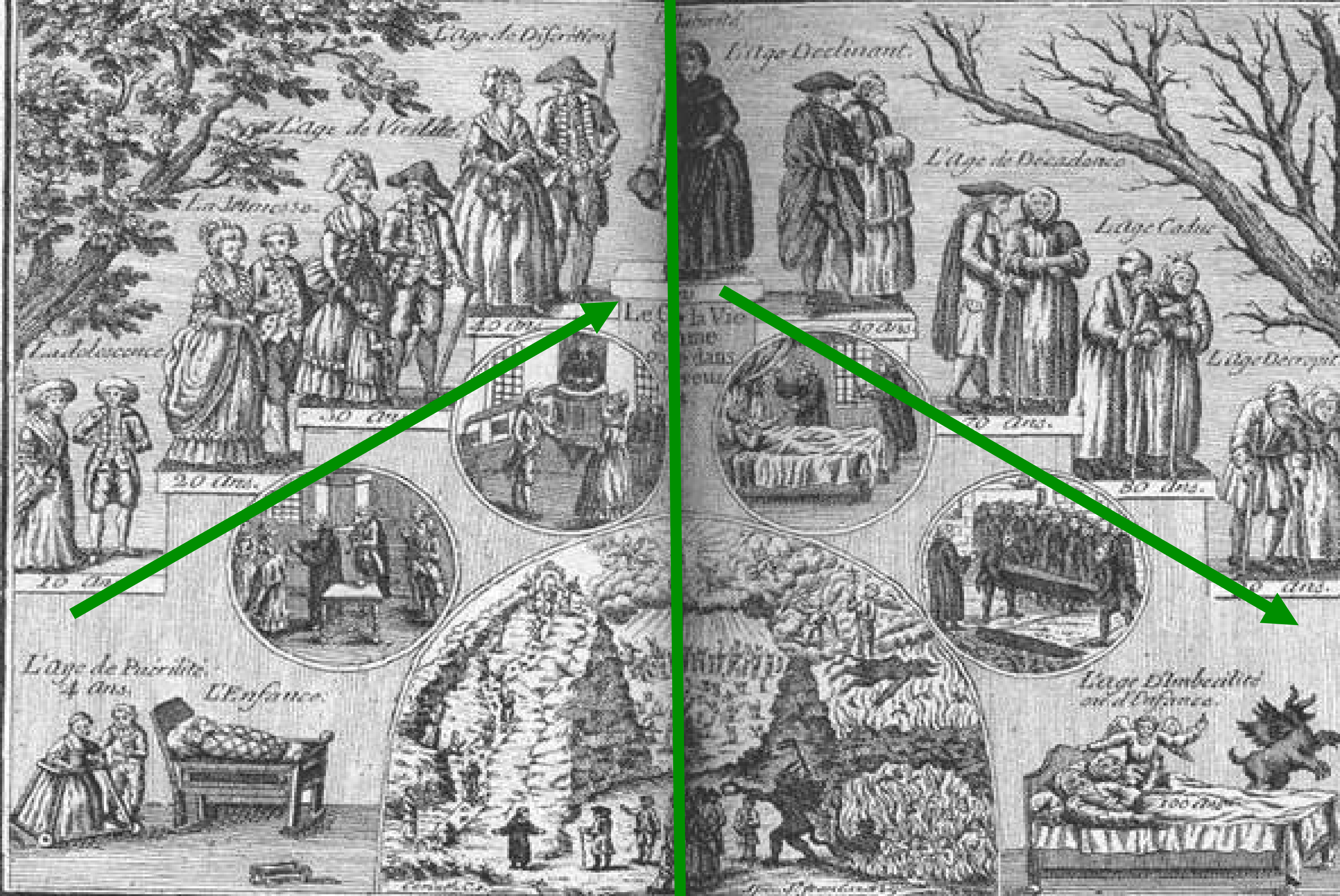


Le bâton symbolise l'activité
du médecin qui promène
sa science secourable

Le serpent symbolise la personne
qui fait le choix de participer
ou non au soin, prenant ainsi en
main son propre bien-être

e-santé asservie au patient





L'Homme commence par l'Enfance,
Il y souffre beaucoup de maux,
mais ses peines et ses travaux
Le suivent dans l'adolescence.

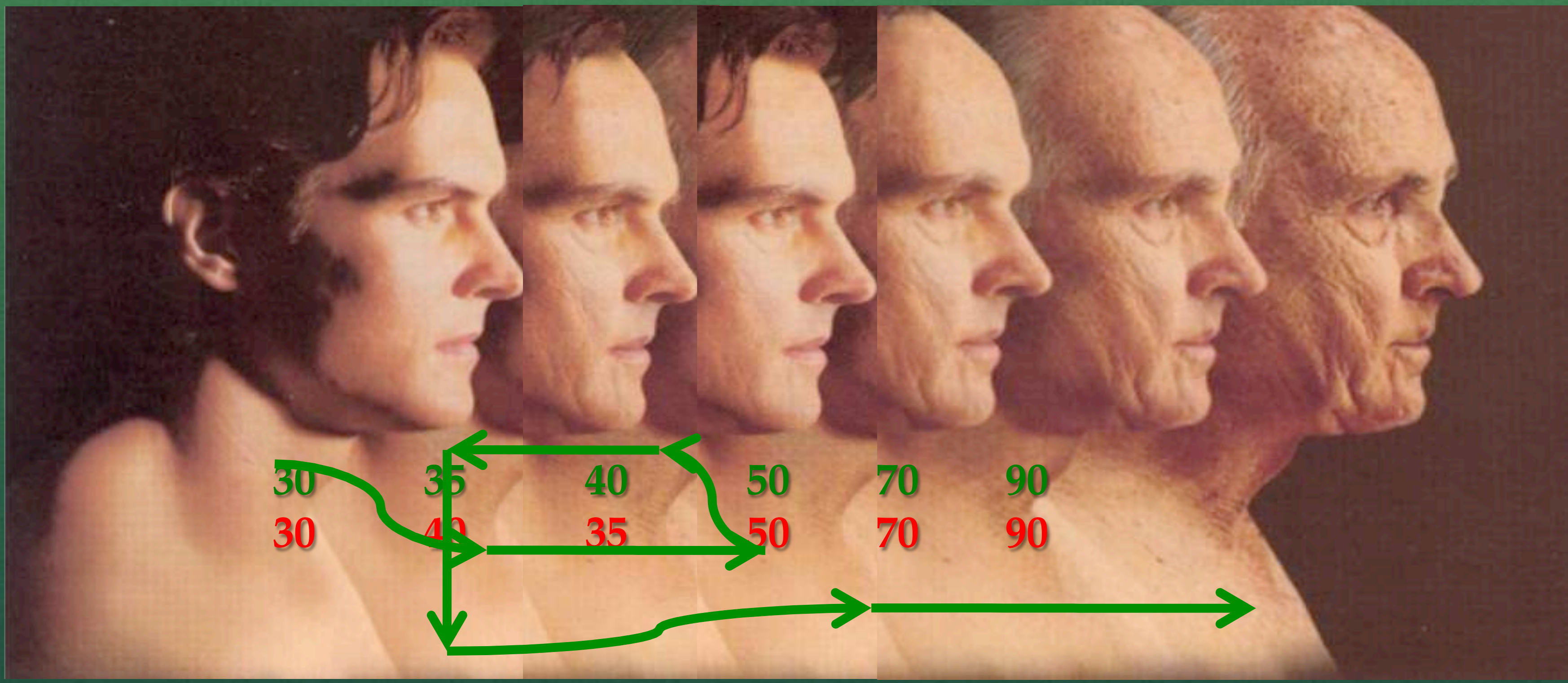
Si la jeunesse a l'avantage,
D'être la saison des plaisirs,
L'âge viril a des desirs
Et de grands soins pour son partage.

L'âge de Maturité
Qui commence à cesser de vivre,
Et dans l'âge qui va le suivre,
Il ... a dit de la Santé.

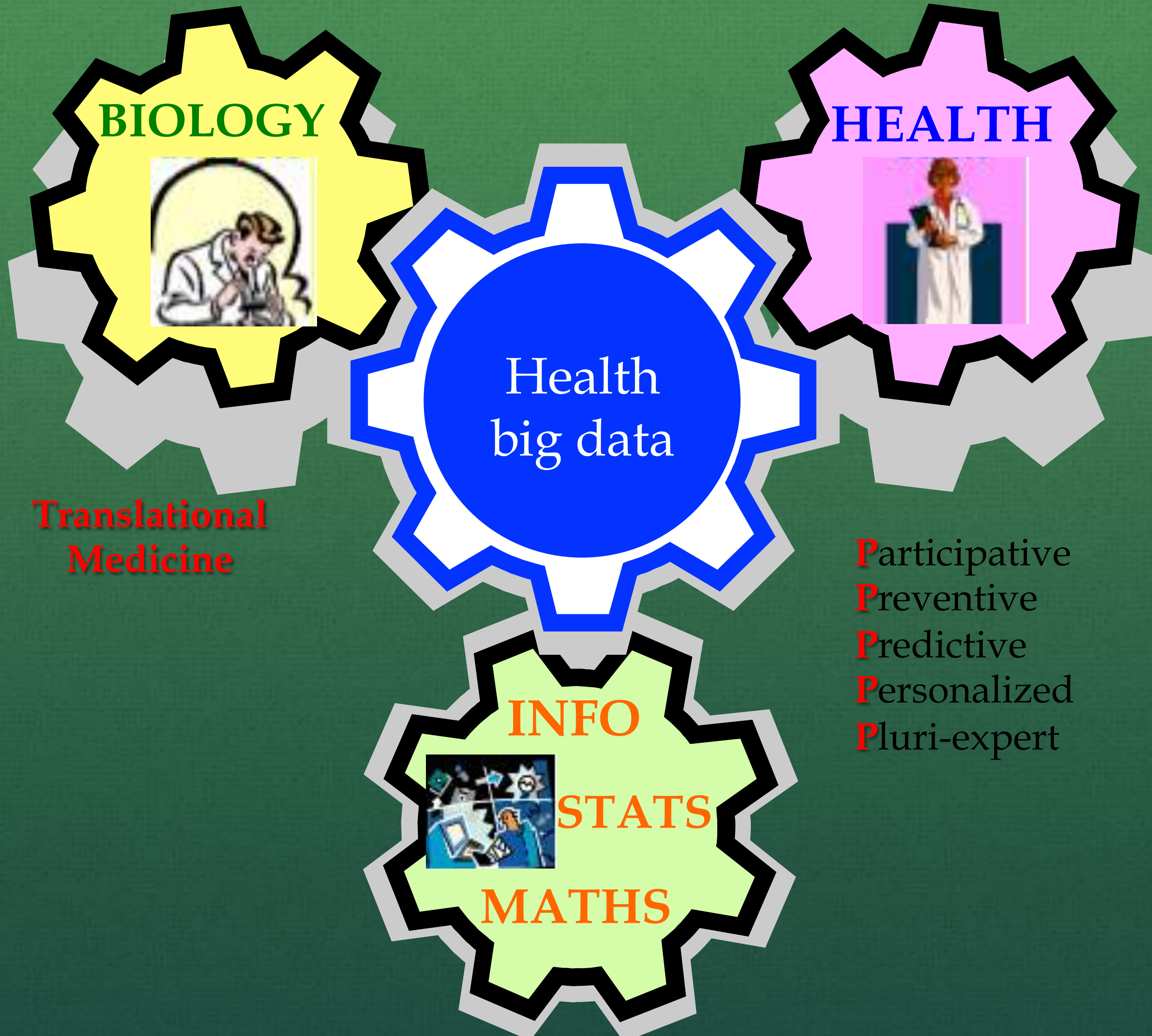
L'âge Caduc et imbécile
N'a prompt plus de jours heureux
Alors sont éteints tous nos feux
Et le tombeau est notre Asile.

Biological Time

Chronological Time



JD, Acta Bio, 2009


BIOLOGY
HEALTH
Health
big data
Translational
Medicine
Participative
Preventive
Predictive
Personalized
Pluri-expert
INFO
STATS
MATHS

Methods of investigation, compression & interpretation

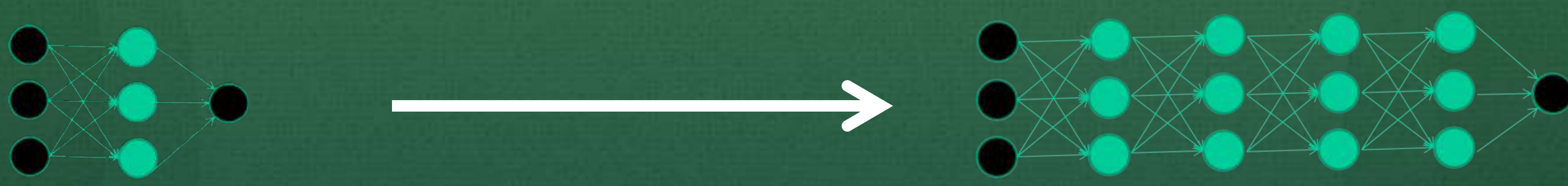
Methods of investigation

- Classical data analysis
- Deep learning
- No SQL interrogation

Le “deep learning est une méthode non supervisée, qui utilise un classifieur réseau de neurones à plusieurs couche.

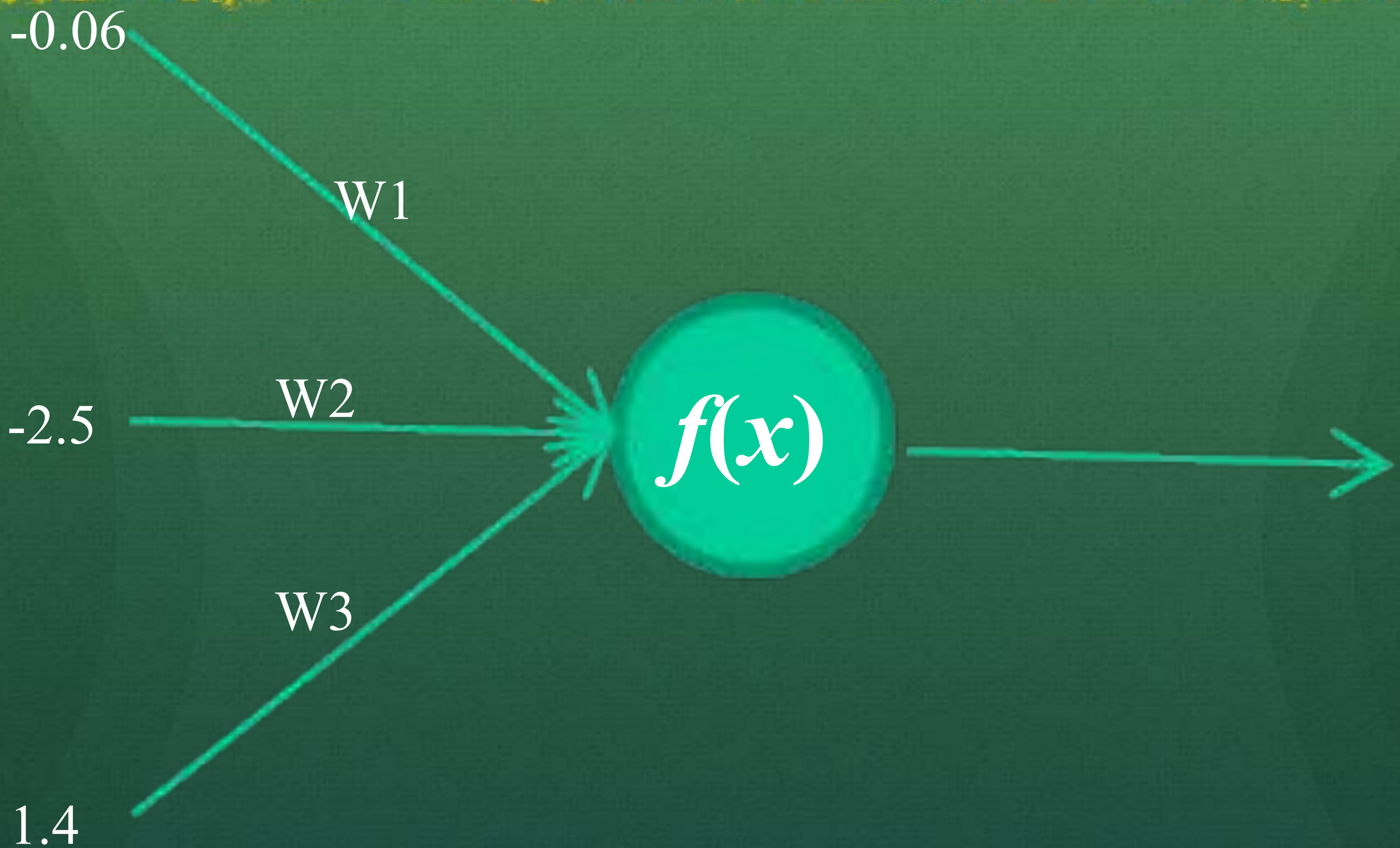
Par rapport aux réseaux classiques (Hopfield, Kohonen, etc),

adjonction de plusieurs couches

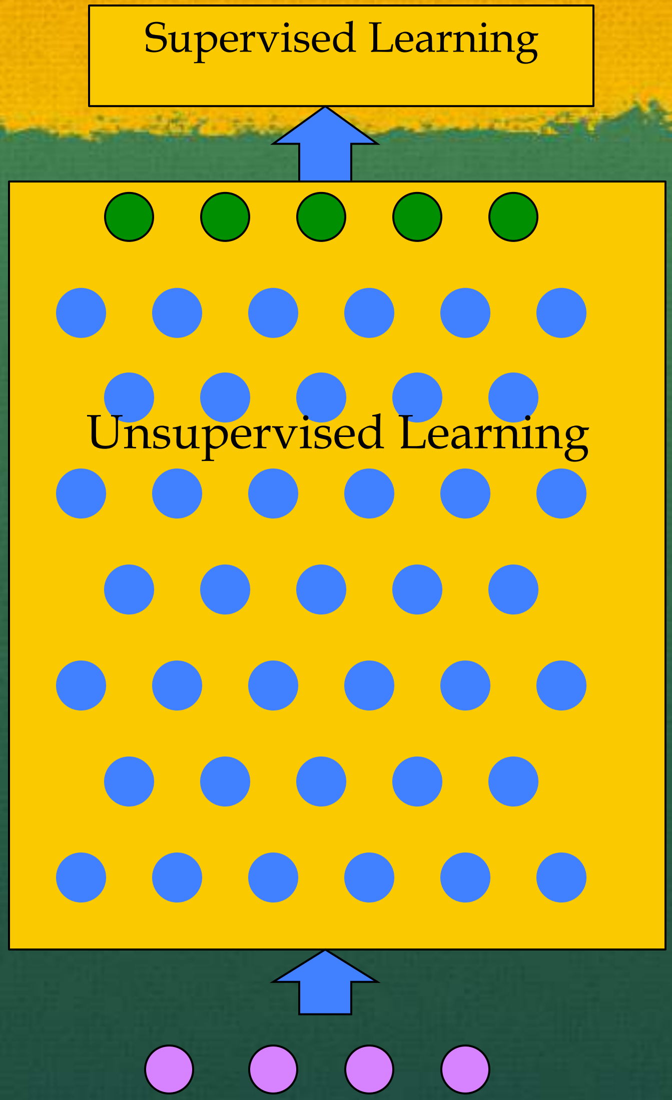


McCulloch-Pitts 1940

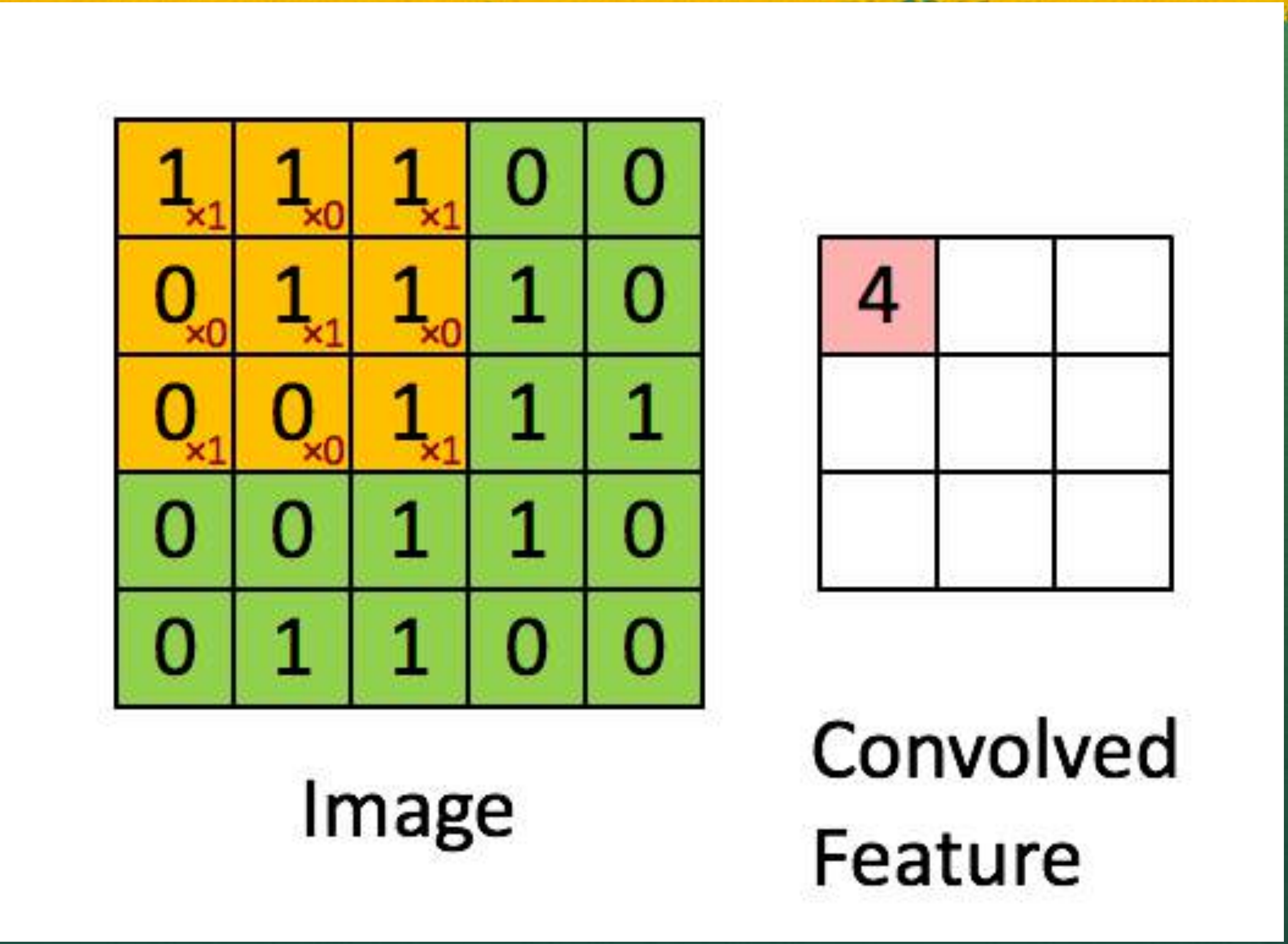
$$f(x) = \frac{1}{1 + \mathrm{e}^{-x}}$$



Deep Learning



Convolution (L. Schwartz, Y. Lecun) Pour régulariser, on convole…



Convolutional Neural Networks

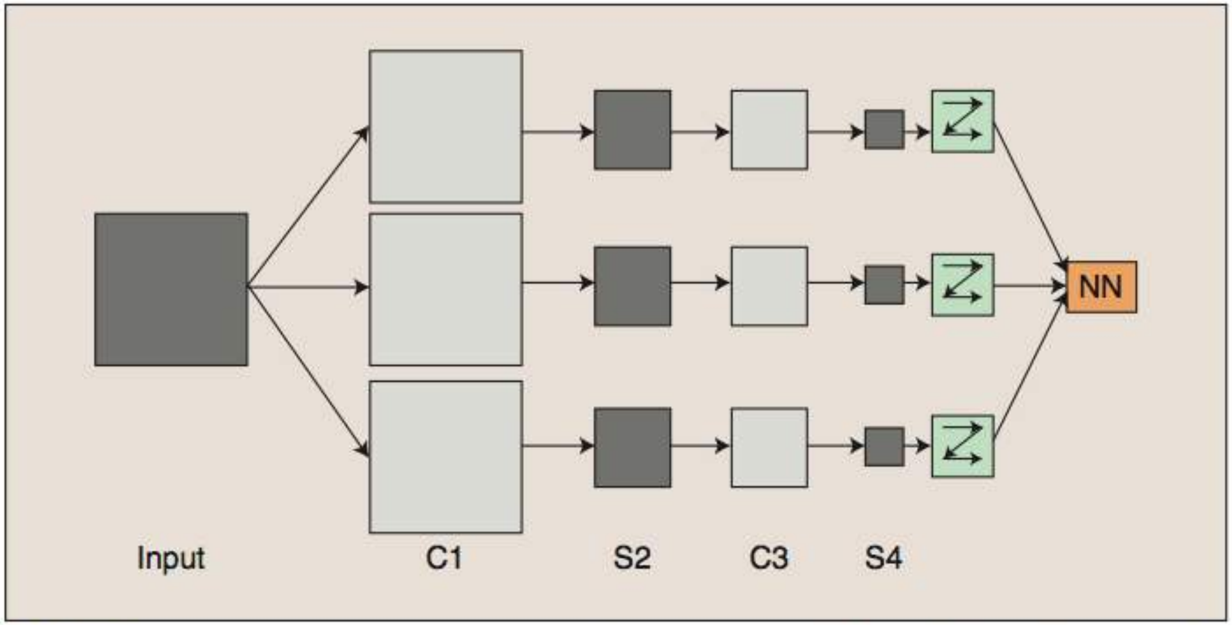
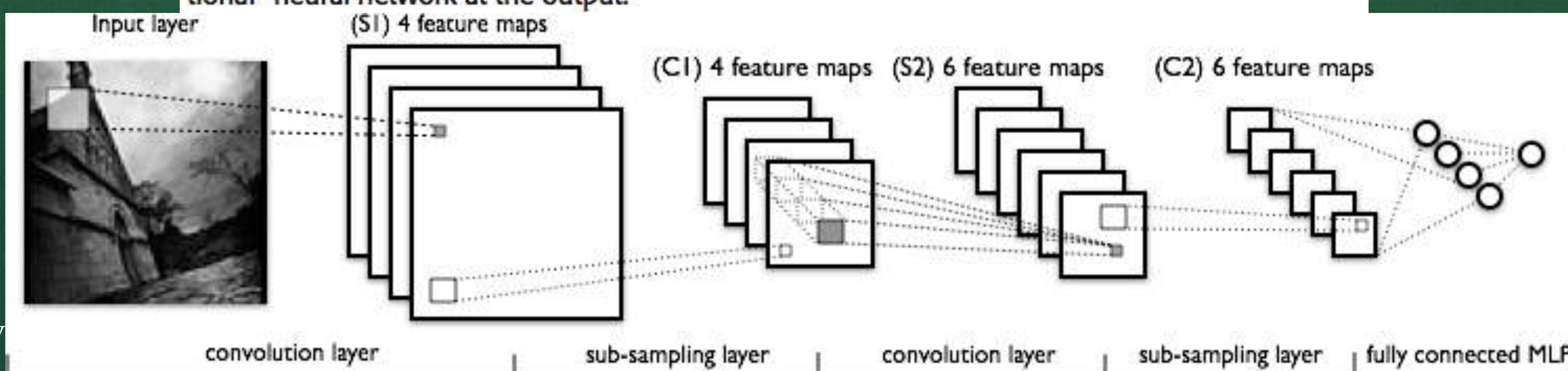


FIGURE 2 Conceptual example of convolutional neural network. The input image is convolved with three trainable filters and biases as in Figure 1 to produce three feature maps at the C1 level. Each group of four pixels in the feature maps are added, weighted, combined with a bias, and passed through a sigmoid function to produce the three feature maps at S2. These are again filtered to produce the C3 level. The hierarchy then produces S4 in a manner analogous to S2. Finally these pixel values are rasterized and presented as a single vector input to the "conventional" neural network at the output.



Cascade of convolutions

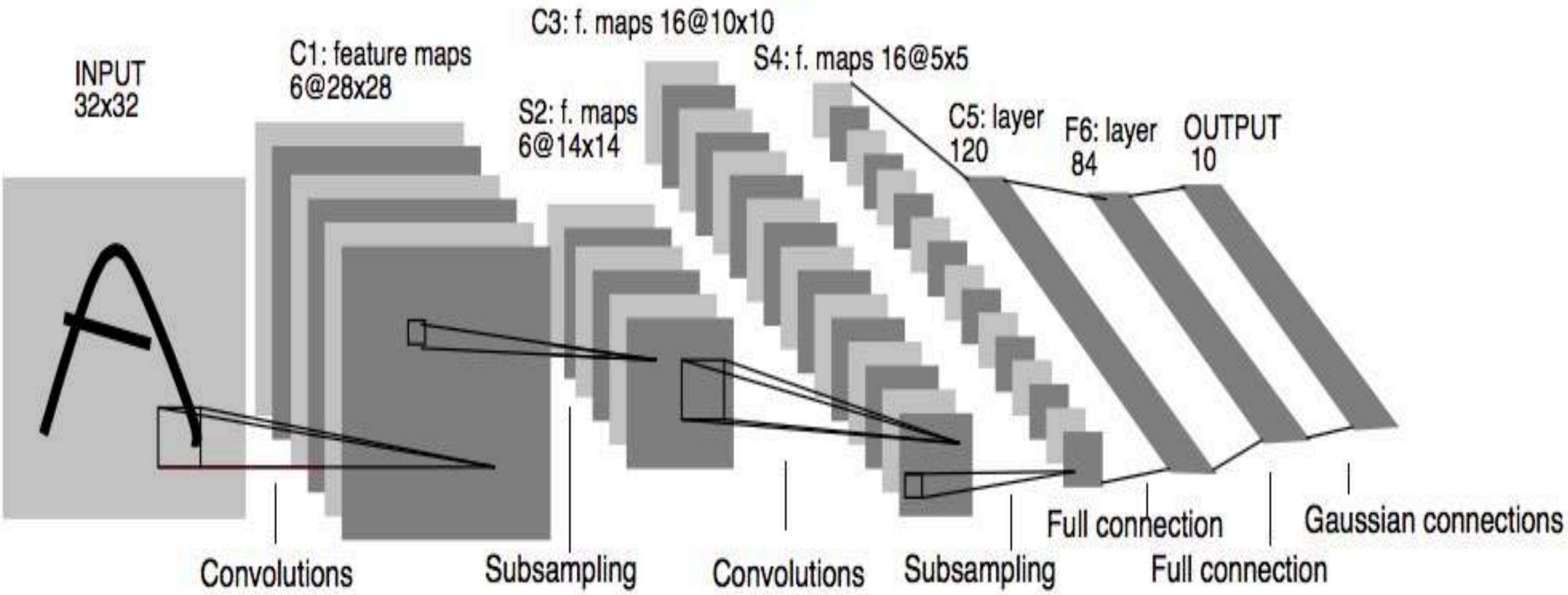
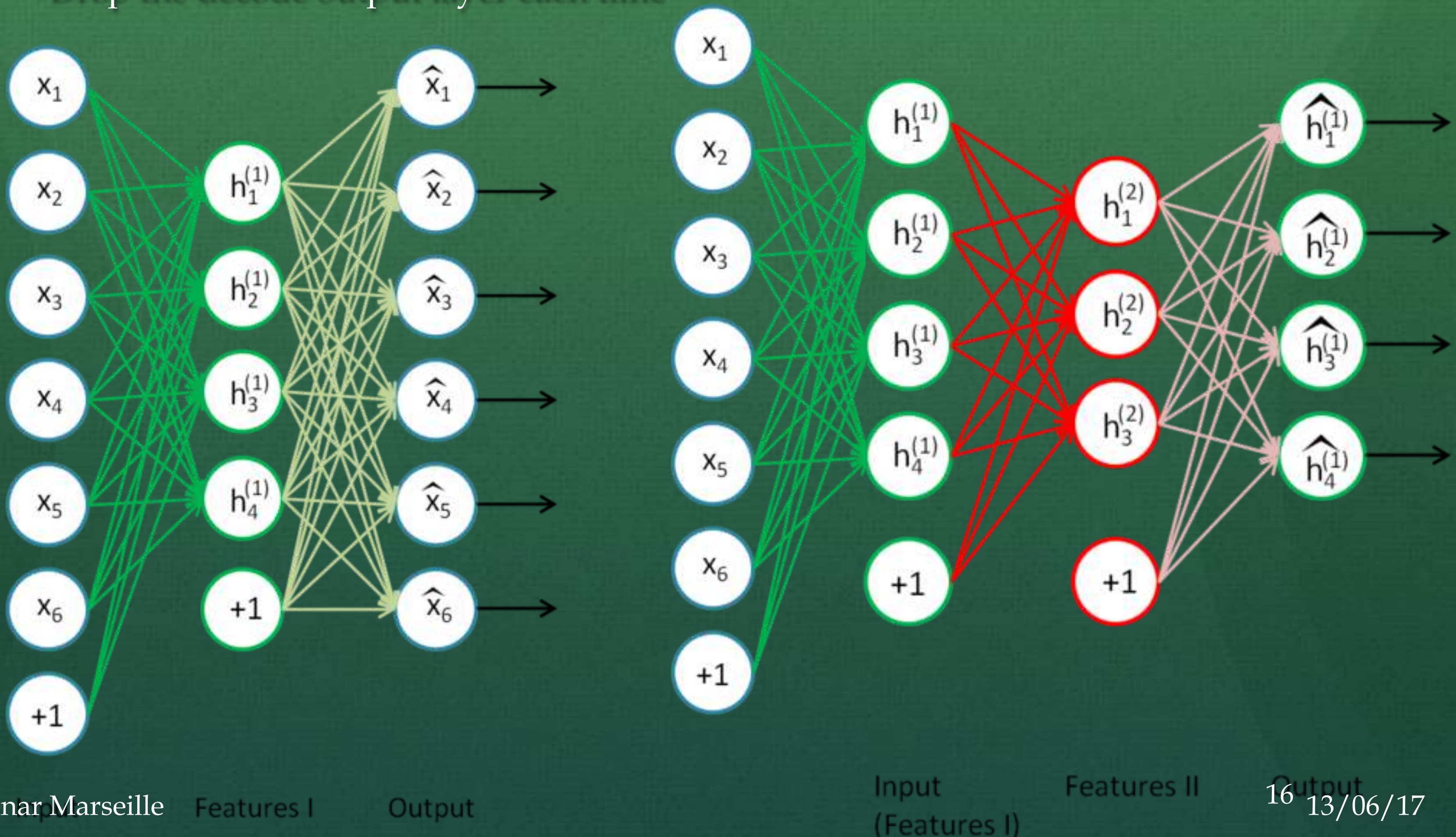


Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

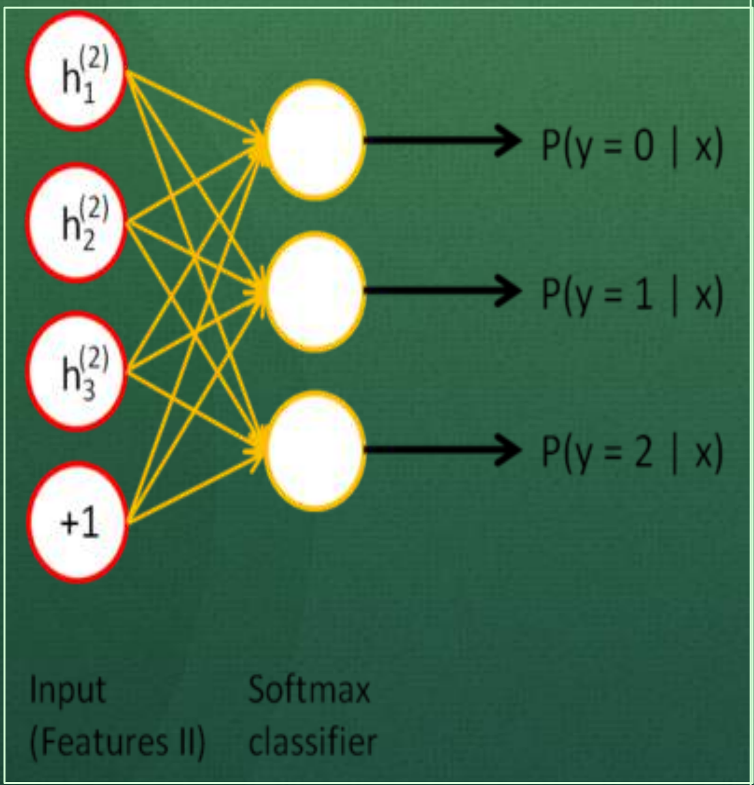
5x5 2x2 5x5 2x2 Fully Connected

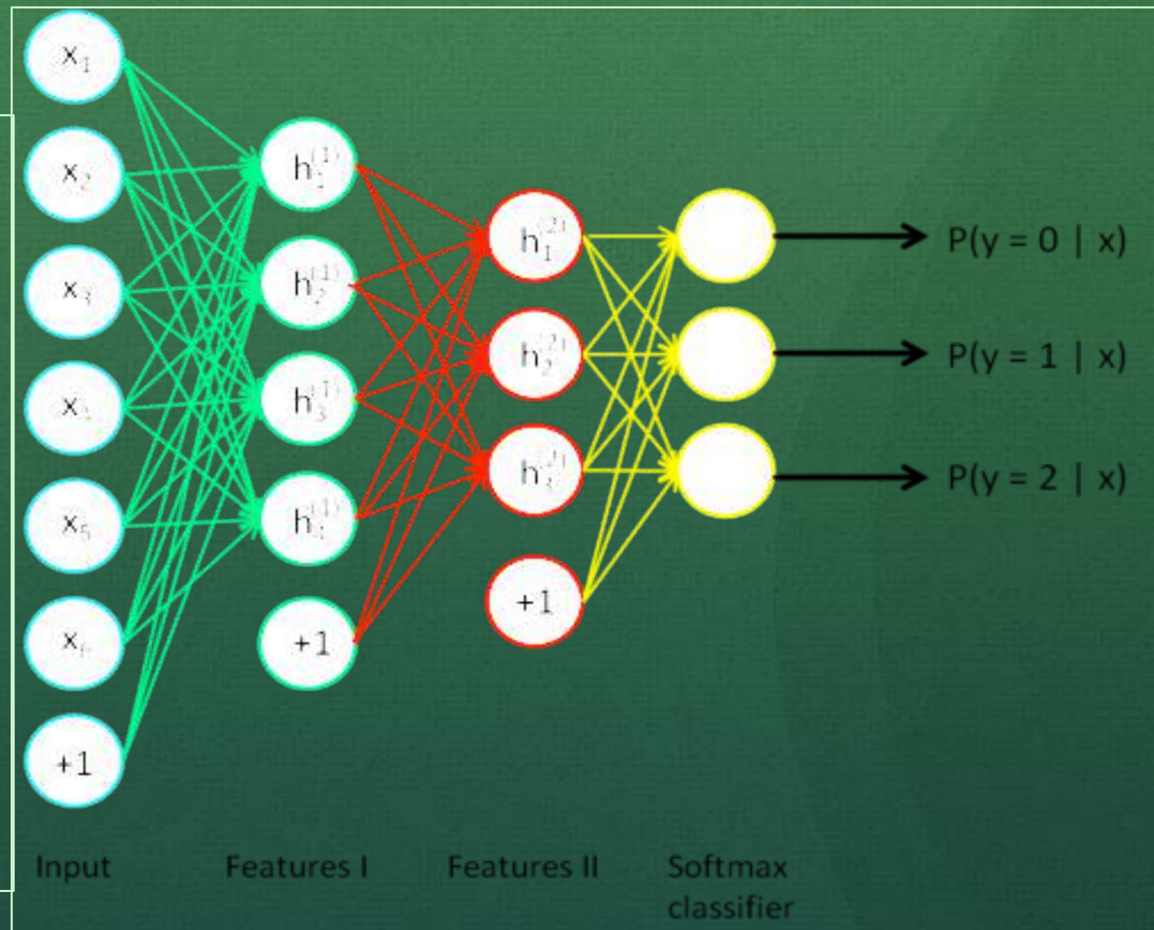
Stacked Auto-Encoders

- Bengio (2007) – After Deep Belief Networks (2006)
- Stack many (sparse) auto-encoders in succession and train them using greedy layer-wise training
- Drop the decode output layer each time



Auto-encodeurs empilés





Methods of compression

- Statistical compression
- Model driven compression
- Qualitative ontologic compression

Statistical compression

for reducing health big data

into a few data representations

based on data analysis restituting

the essential of the medico-social

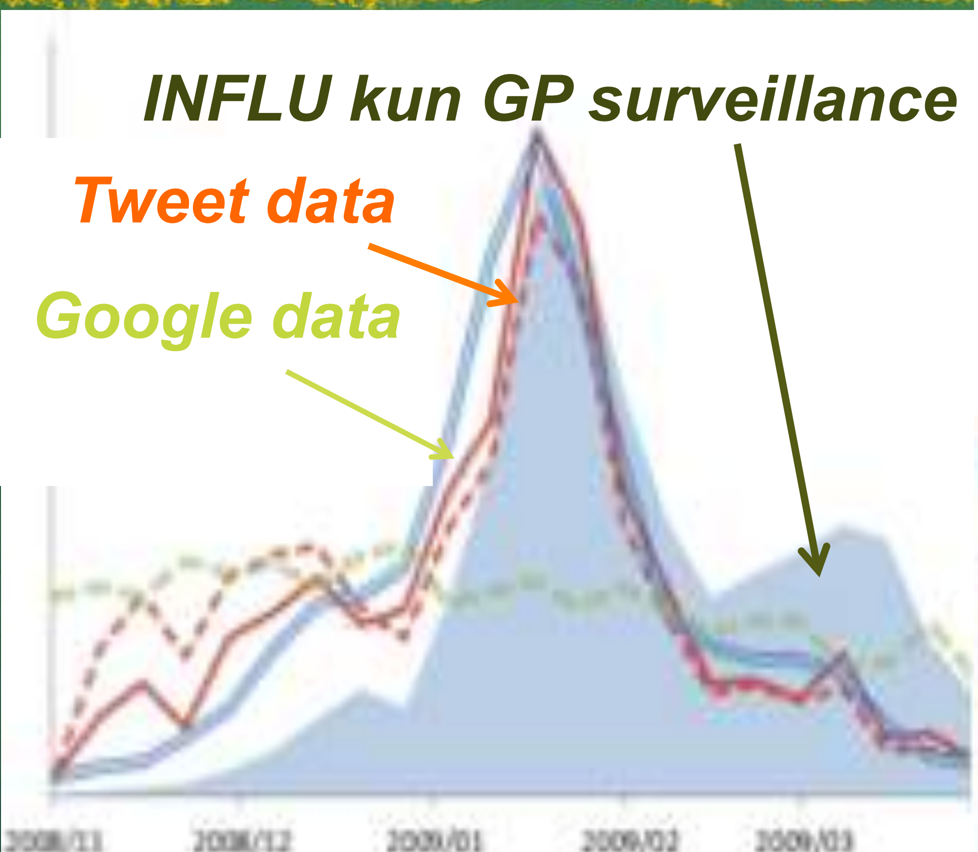
knowledge

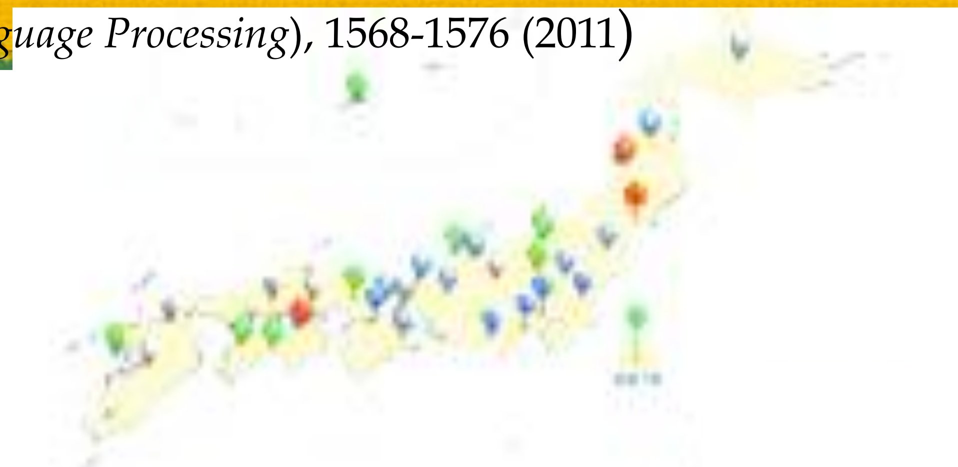
Influenza Web surveillance in Japan social networks

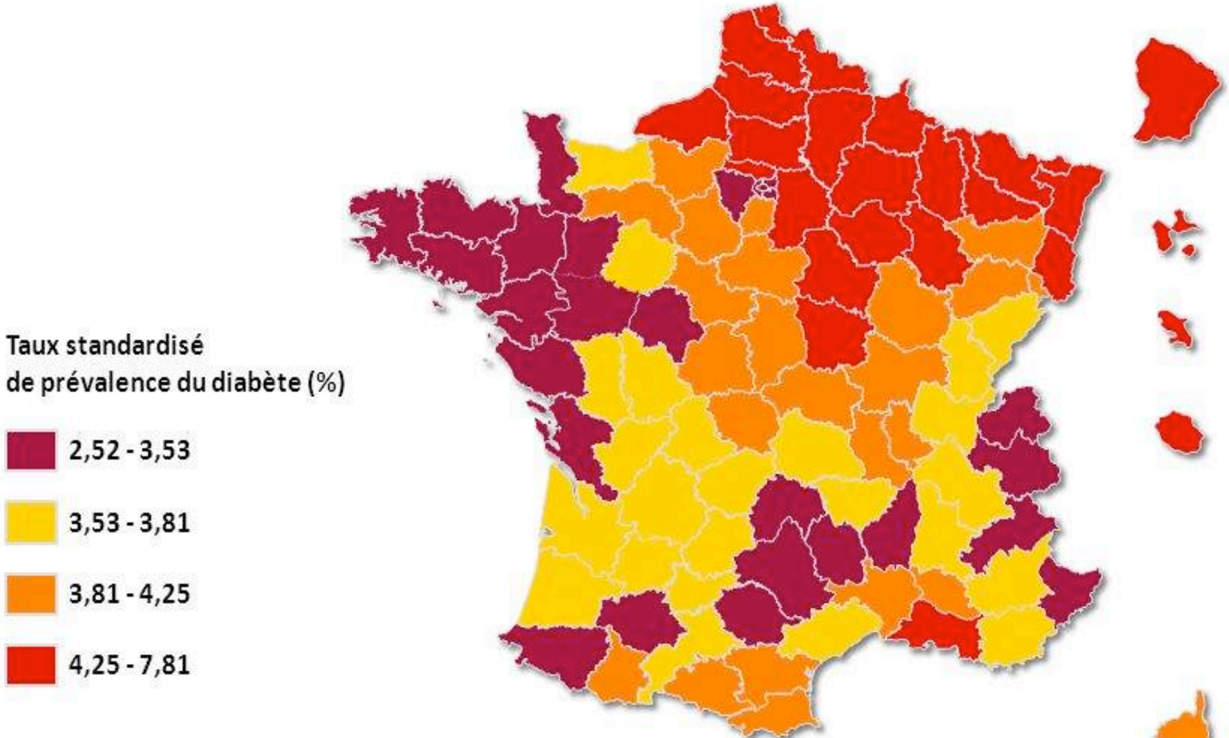
Eiji Aramaki, Sachiko Maskawa, Mizuki Morita

Twitter Catches The Flu: Detecting Influenza Epidemics using Twitter

EMNLP (*Empirical Methods on Natural Language Processing*), 1568-1576 (2011)

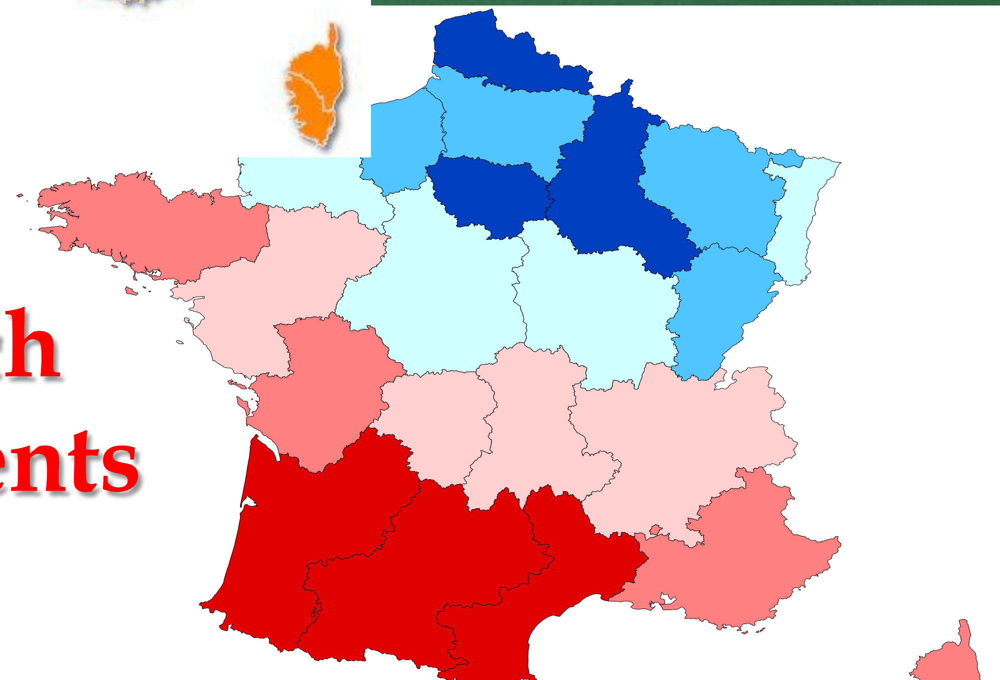






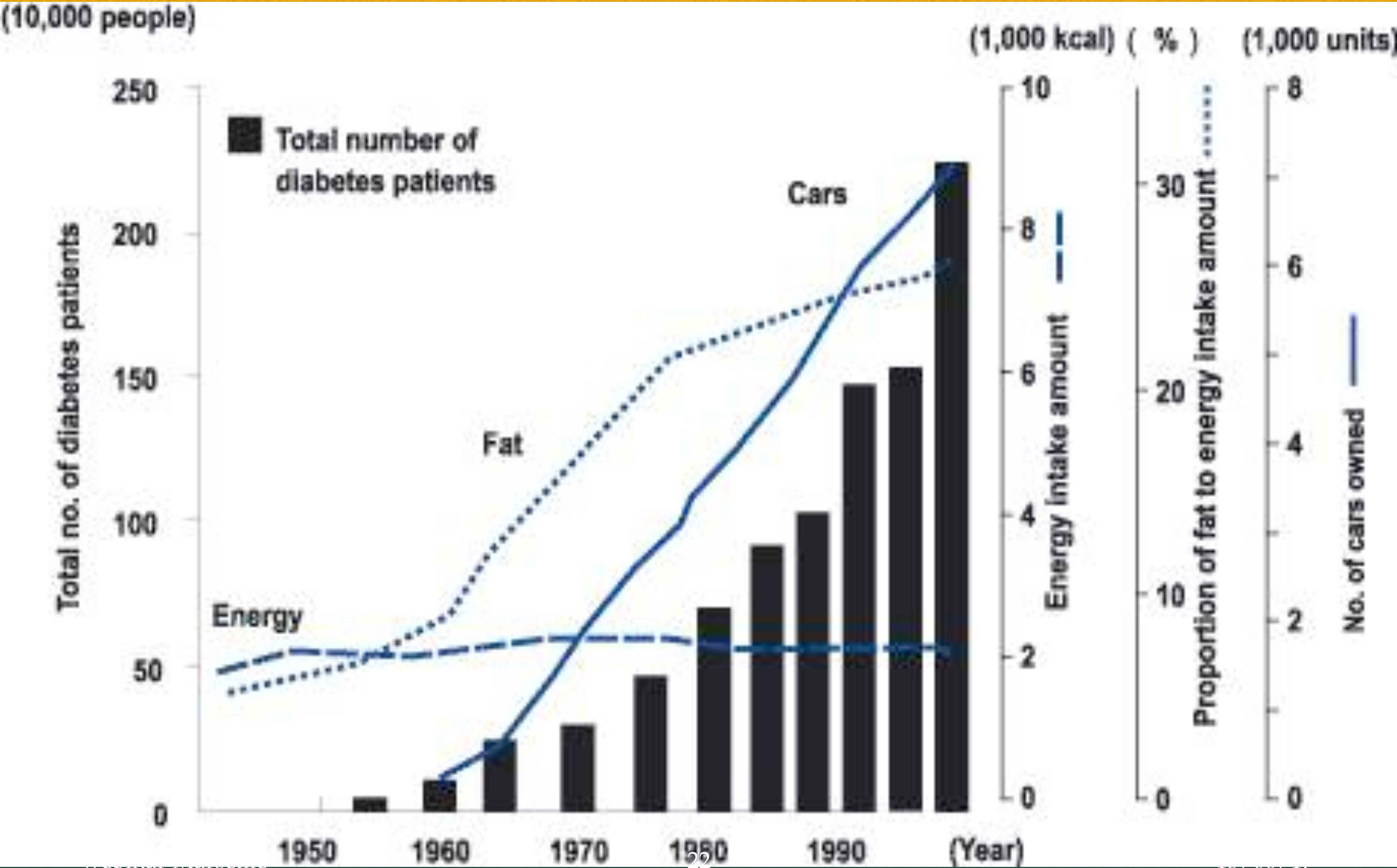
Type II diabetes prevalence OECD 2008

South-North French medico-social gradients



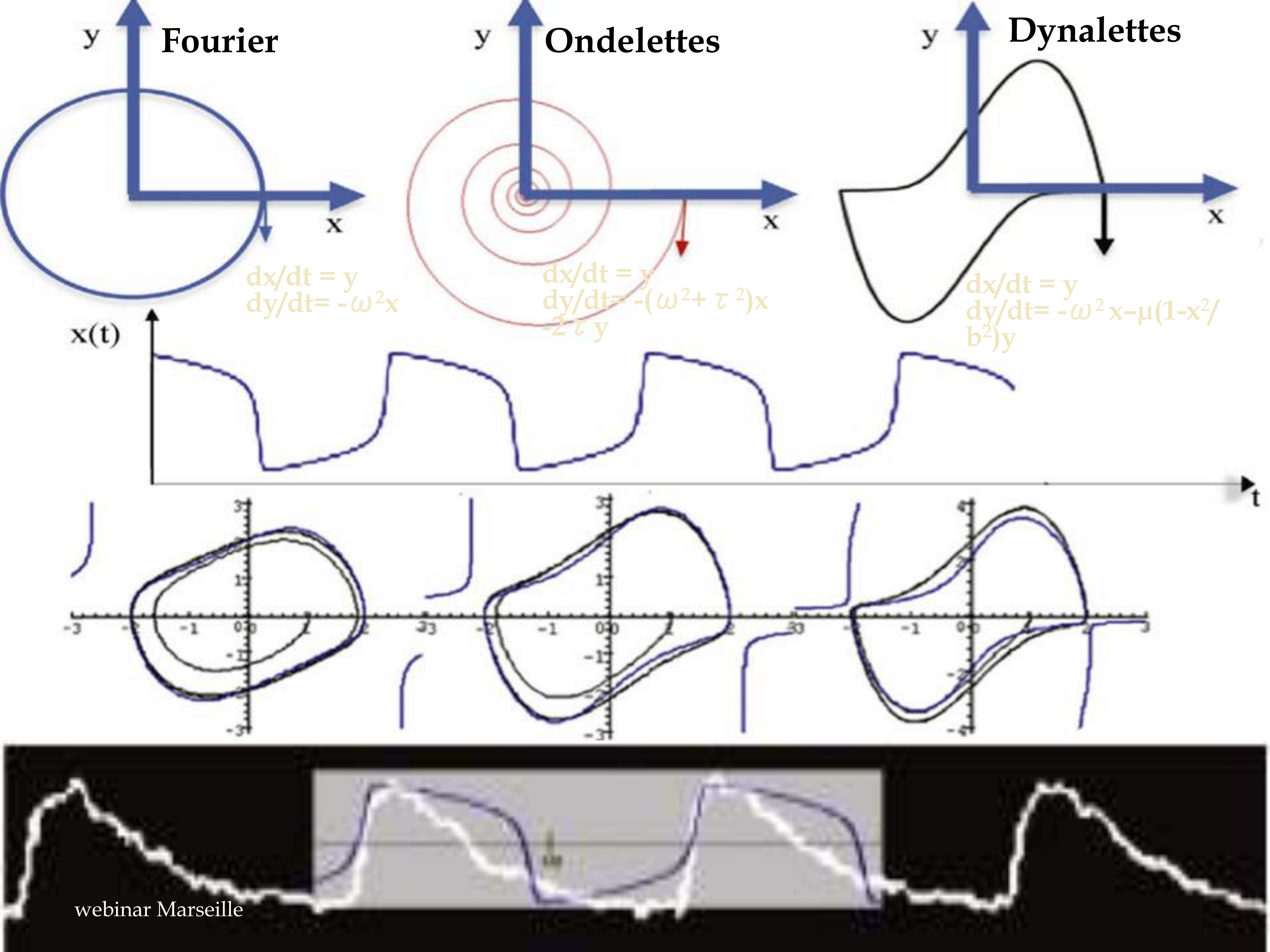
Attractivity of French regions INSEE 2008

Type II Diabetes increase in Japan



Model driven compression

Generalization of the Fourier & wavelets approach



Fourier
y
x
dx/dt = y
dy/dt= -ω²x
Ondelettes
y
x
dx/dt = y
dy/dt= -(ω²+ τ²)x -2τ y
Dynalettes
y
x
dx/dt = y
dy/dt= -ω² x-μ(1-x²/b²)y
x(t)
t
webinar Marseille

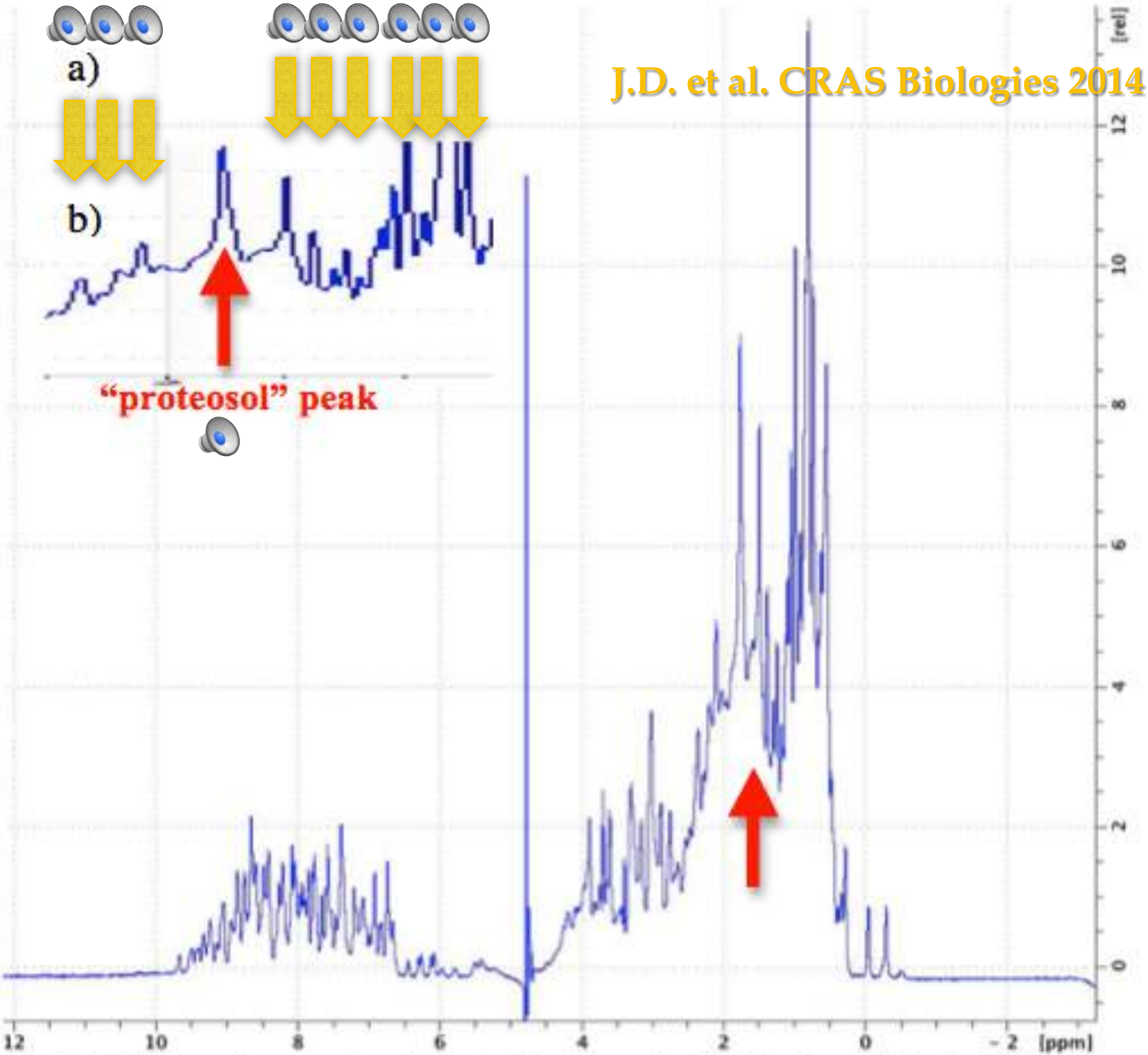


Fig. 1: a) original protein NMR spectroscopy signal; b) extraction of a peak called "proteosol" to be isolated and processed by the Dynalet transform (red arrow).

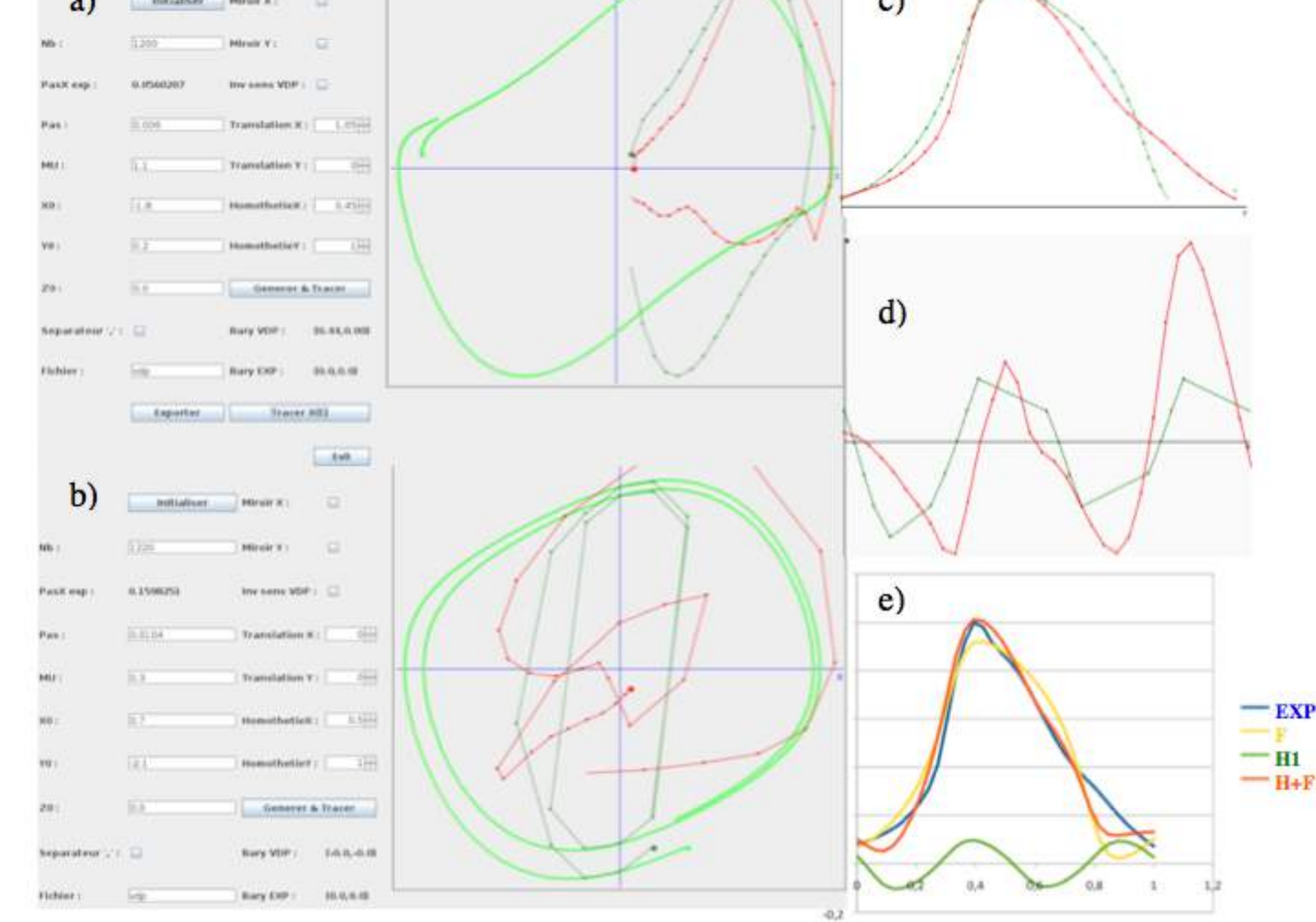


Fig. 2: a) original empirical protein signal of Fig. 1 (in red) matched with the van der Pol limit cycle (in dark green); b) first harmonic signal matched with the first harmonic of the van der Pol signal; c) fundamental temporal original signal (in red) matched with the van

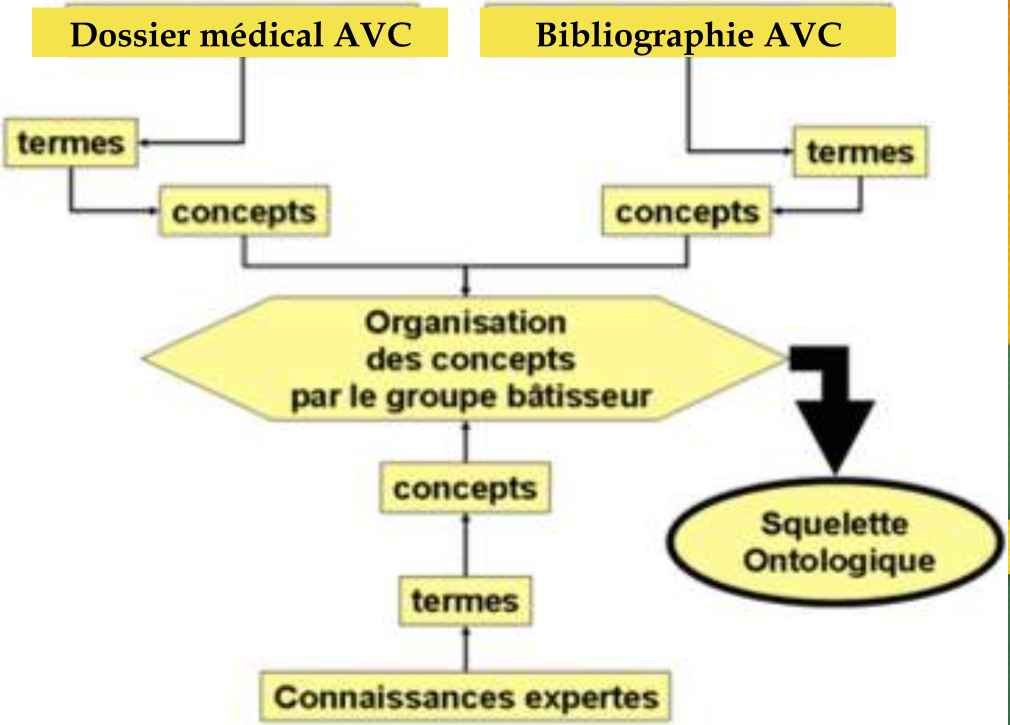
Qualitative ontologic compression

into logic representations

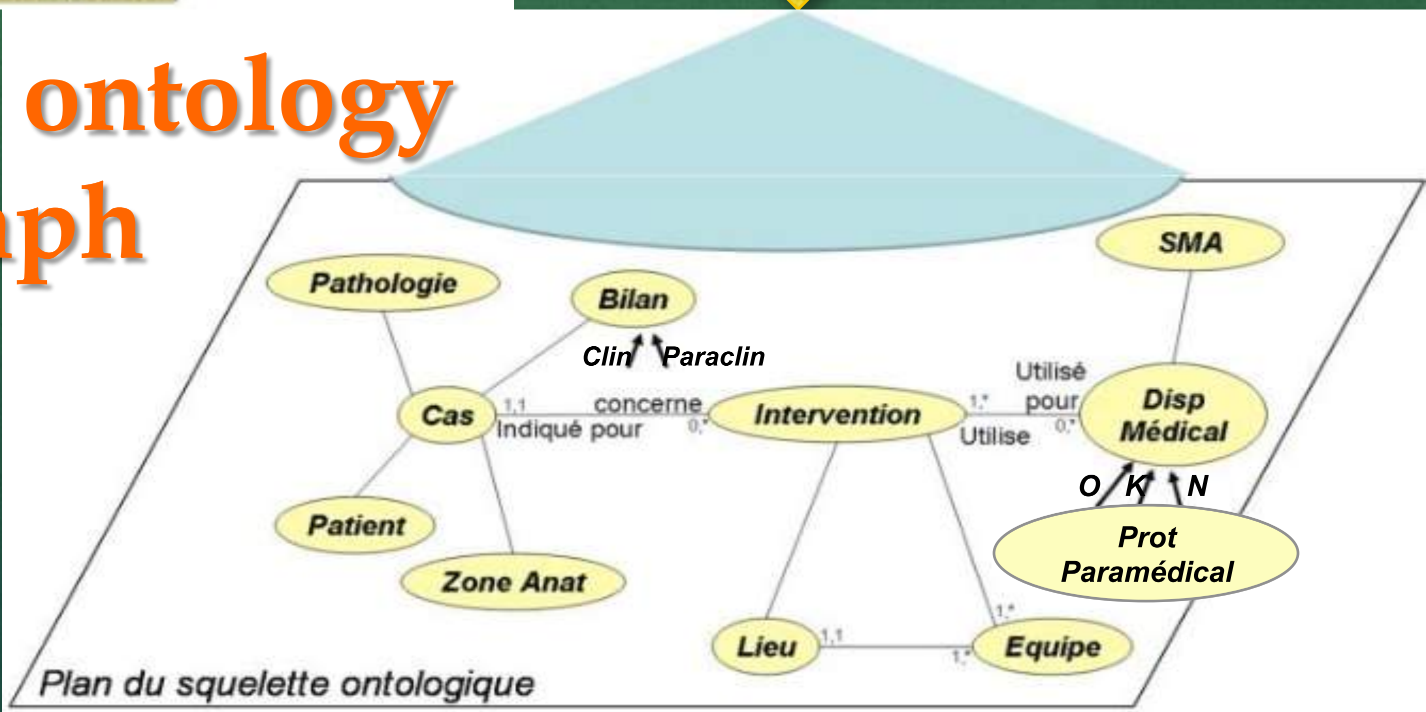
based on

semantic analysis and

ontologic representation

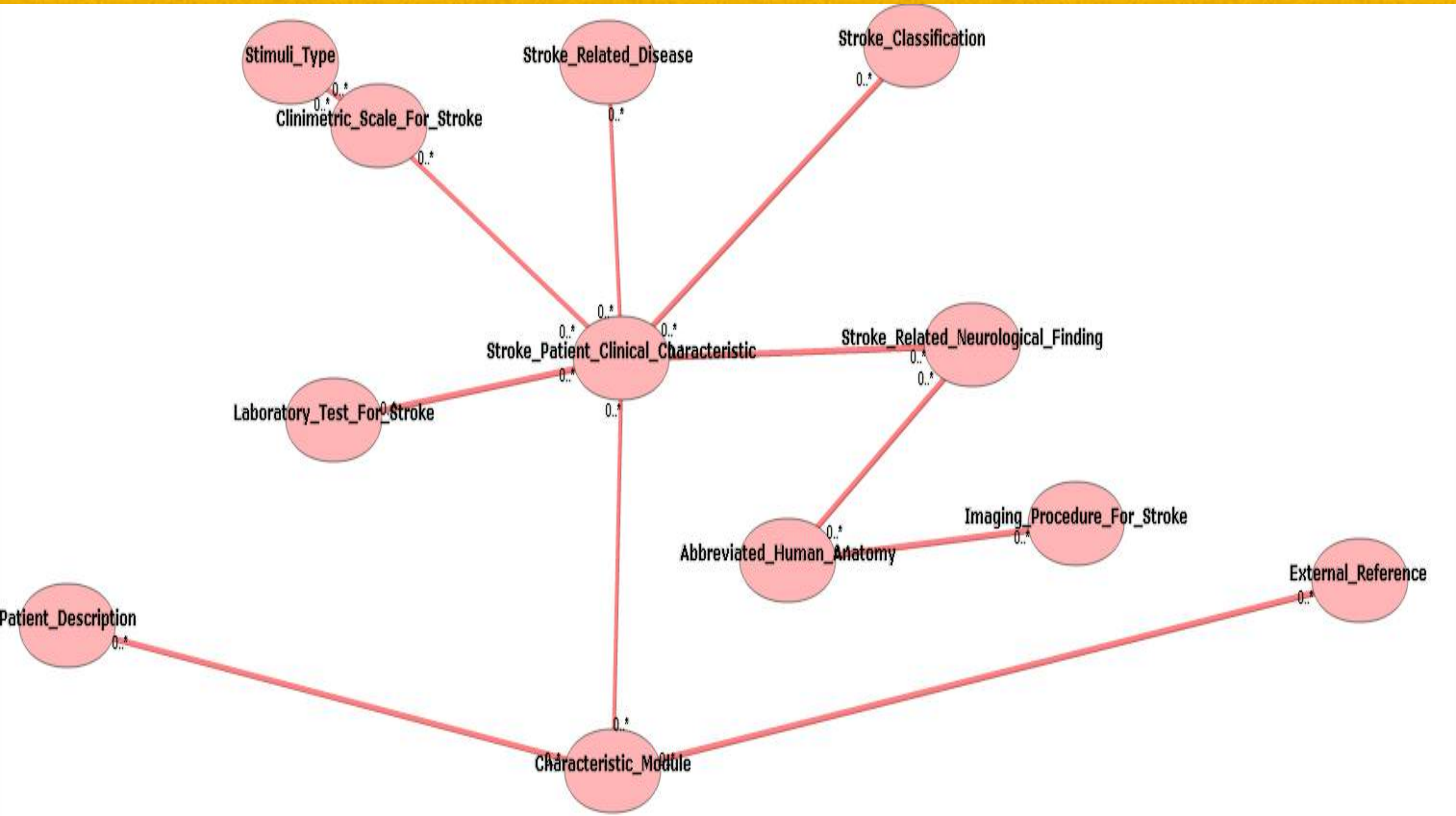


Concepts ontology graph



webinar Marseille

Ontology building tools



Methods of interpretation

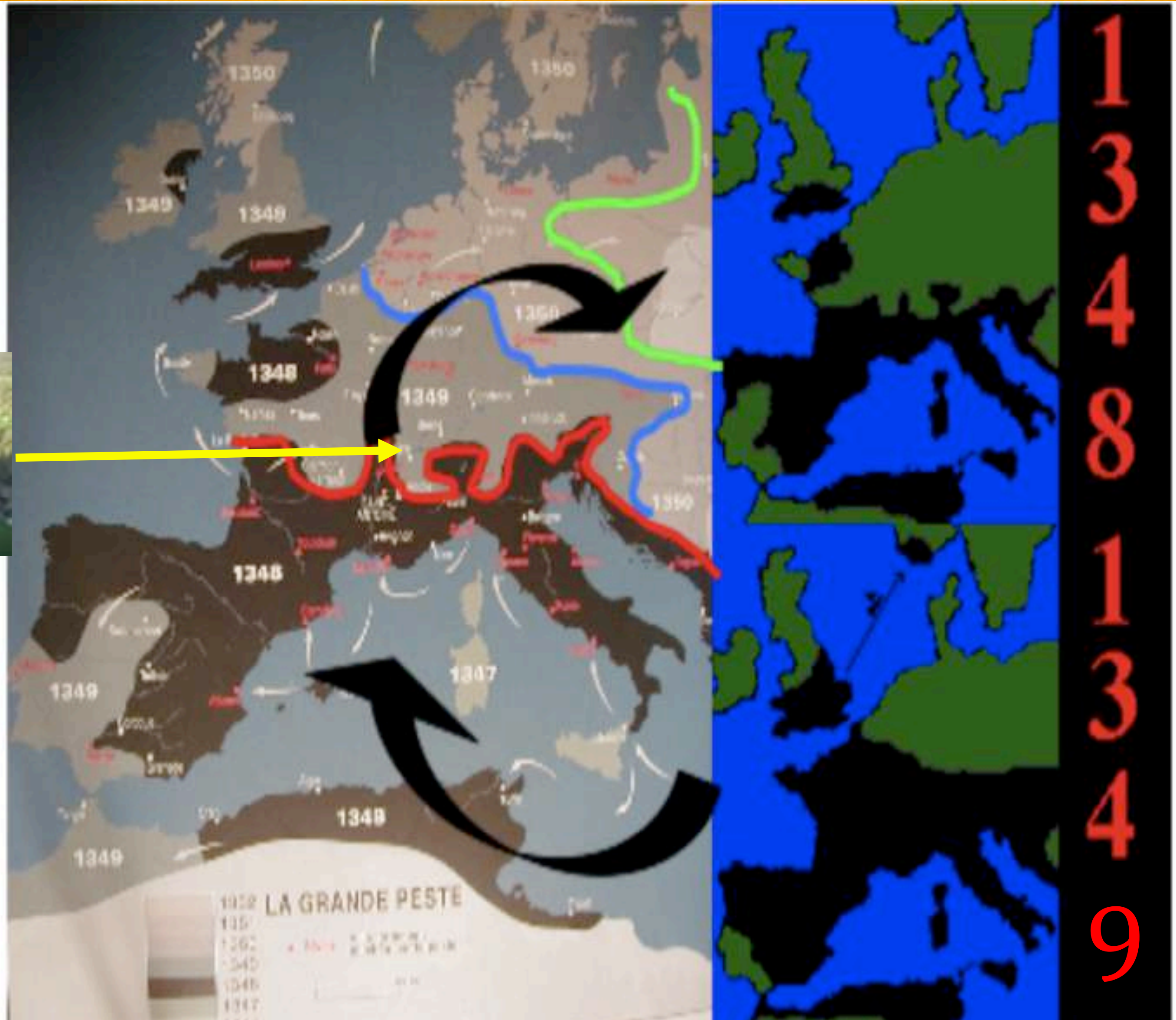
- Differential modelling
- Automata & Multi-Agent modelling
- Functional modelling

Differential modelling

Voies commerciales

Webinar Marseille

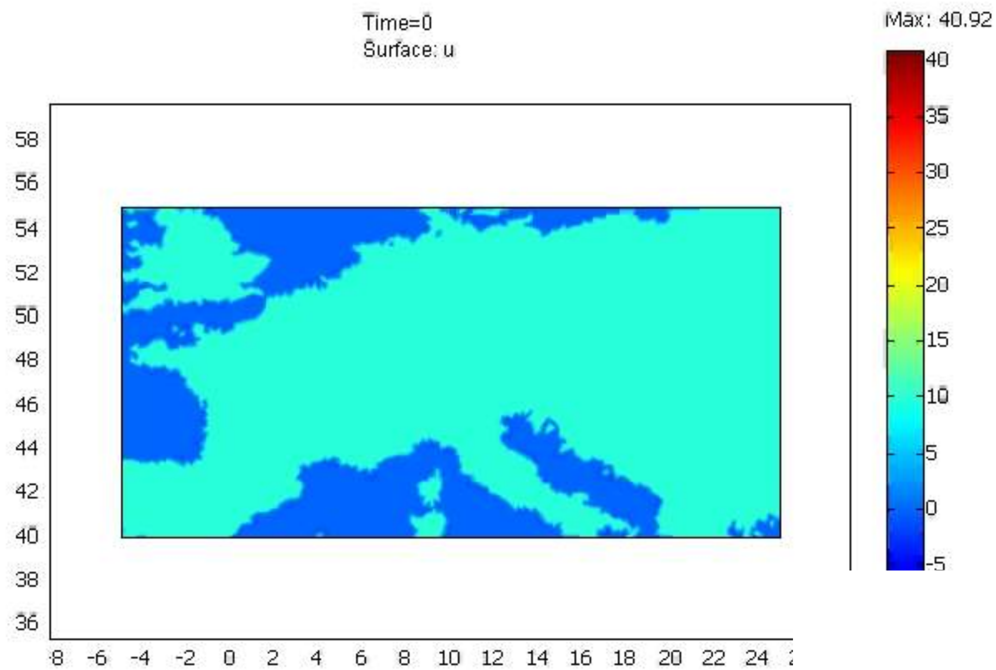
St Anton Abbey

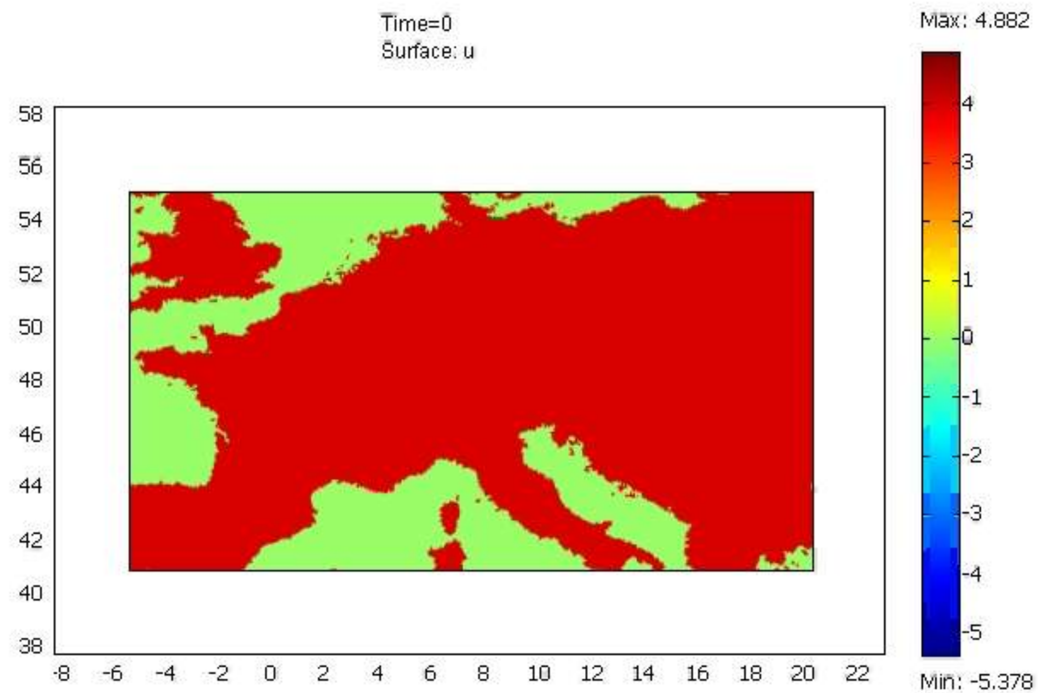


Daniel Bernoulli

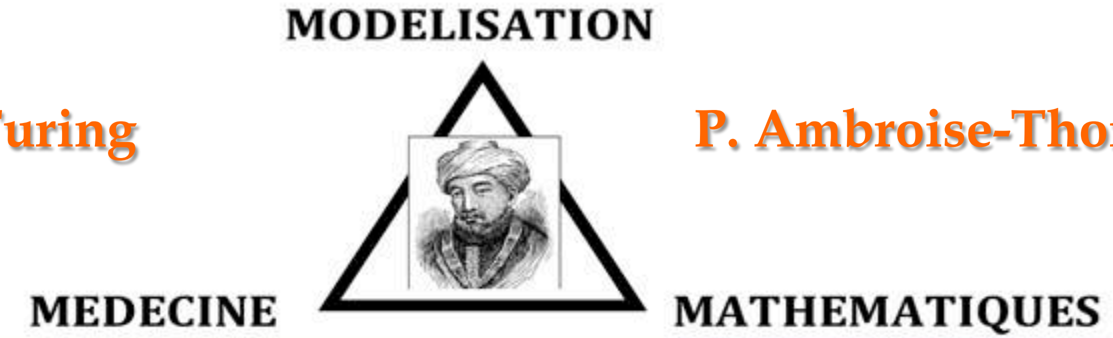
Jean d'Alembert

Ronald Ross


Time=0
Surface: u
Max: 40.92
58 56 54 52 50 48 46 44 42 40 38 36
-8 -6 -4 -2 0 2 4 6 8 10 12 14 16 18 20 22 24
40 35 30 25 20 15 10 5 0 -5

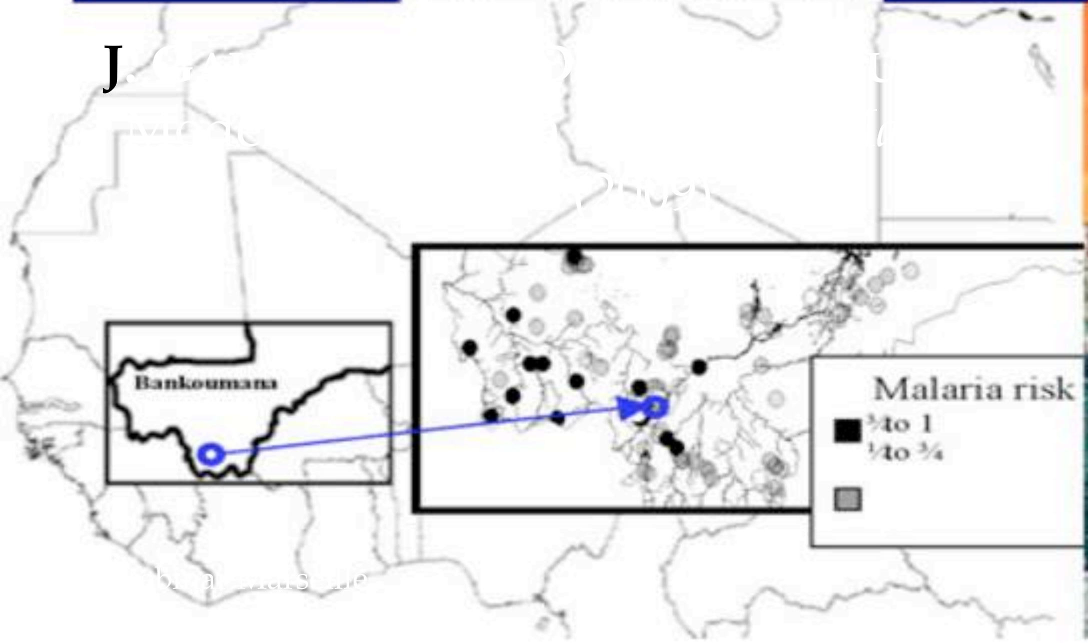

Time=0
Surface: u
Max: 4.882
58 56 54 52 50 48 46 44 42 40 38
-8 -6 -4 -2 0 2 4 6 8 10 12 14 16 18 20 22
4 3 2 1 0 -1 -2 -3 -4 -5
Min: -5.378

A. Turing

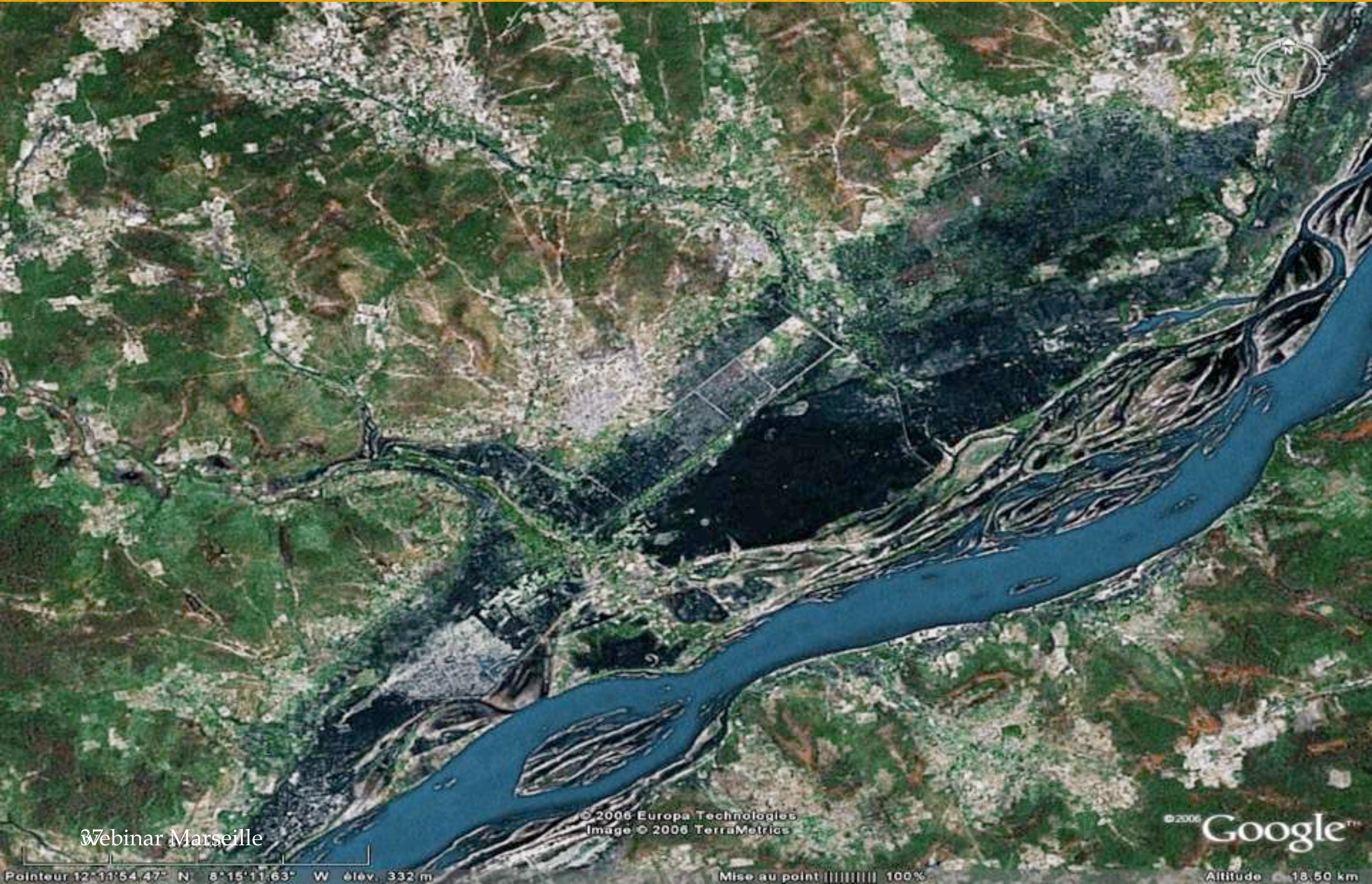


P. Ambroise-Thomas

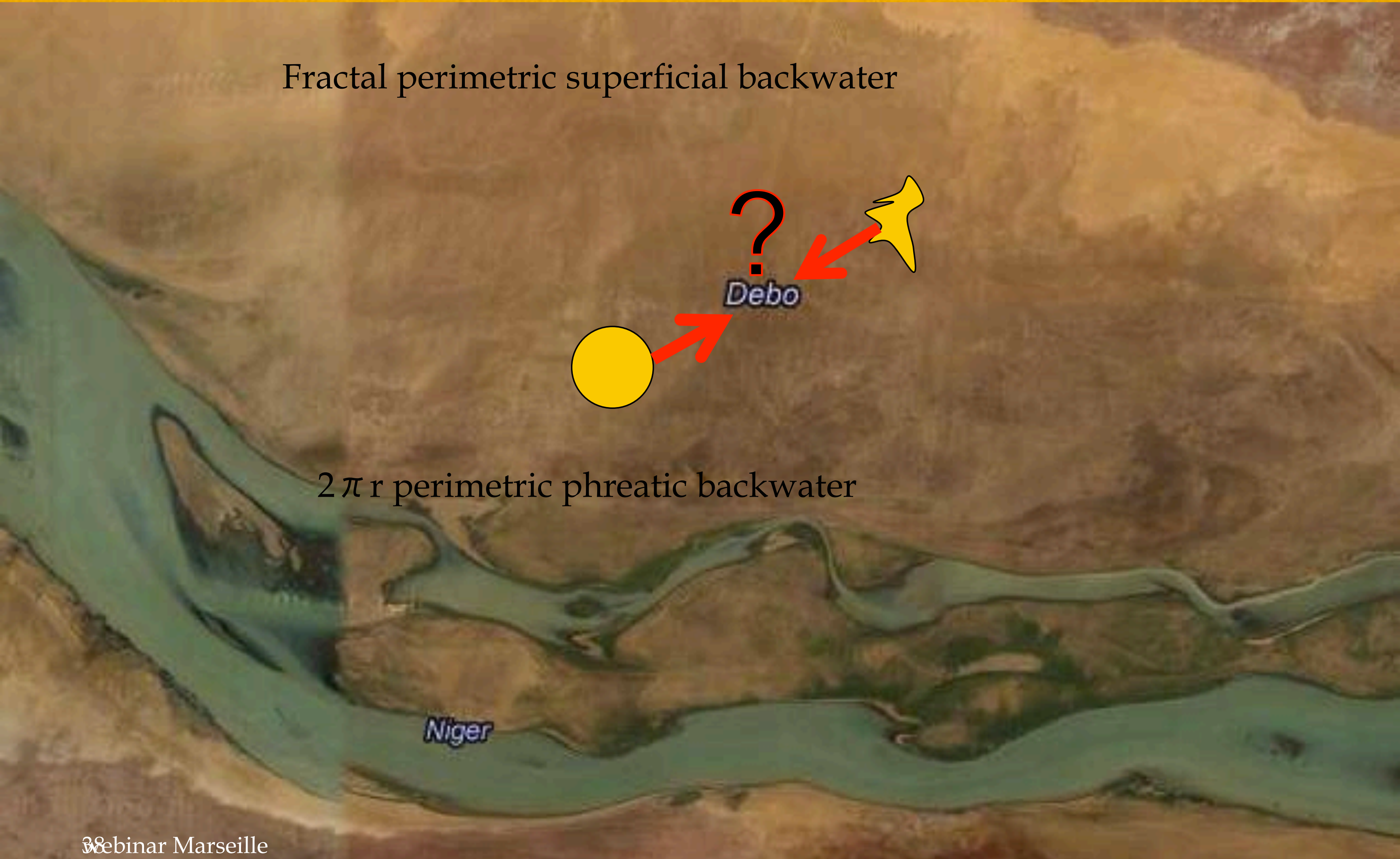
À Alger, j'ai abandonné
mes premières amours
les mathématiques dites pures
pour « entrer en médecine»



Bancoumana (Mali): selon quel gradient progresse la maladie?



Spatial problem



Endemic prevalence

Prévalence = Proportion des personnes
ayant ou ayant eu la maladie dans un échantillon et une durée donnés

$$\frac{\mathrm{d}S(t)}{\mathrm{d}t} = -\mu\alpha\beta\ A_i(t)\ S(t) + \delta\ R(t)$$

$$\frac{\mathrm{d}I(t)}{\mathrm{d}t} = +\mu\alpha\beta\ A_i(t)\ S(t) - (\eta_1 + \gamma)\ I(t) + \eta_2\ G(t)$$

$$\frac{\mathrm{d}G(t)}{\mathrm{d}t} = +\eta_1\ I(t) - (\eta_2 + \gamma)\ G(t)$$

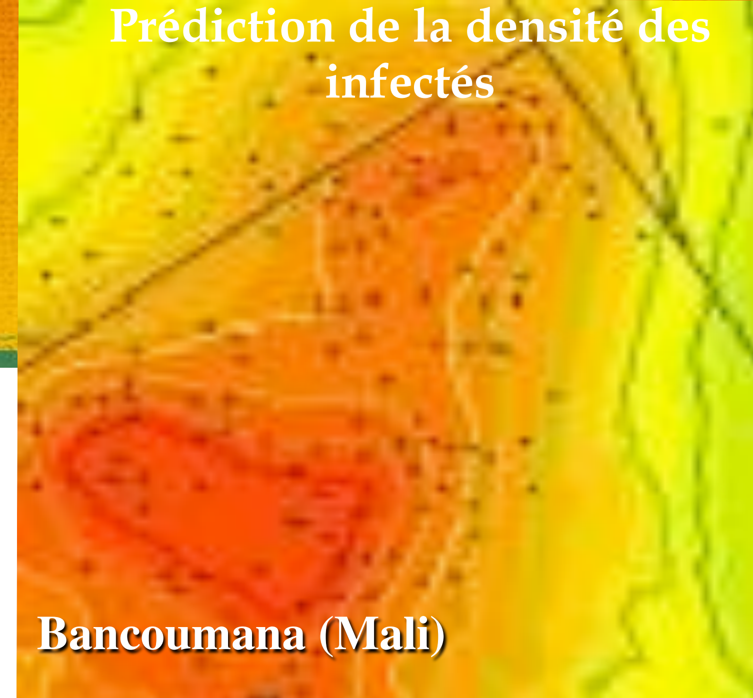
$$\frac{\mathrm{d}R(t)}{\mathrm{d}t} = +\gamma\ (I(t) + G(t)) - \delta\ R(t)$$

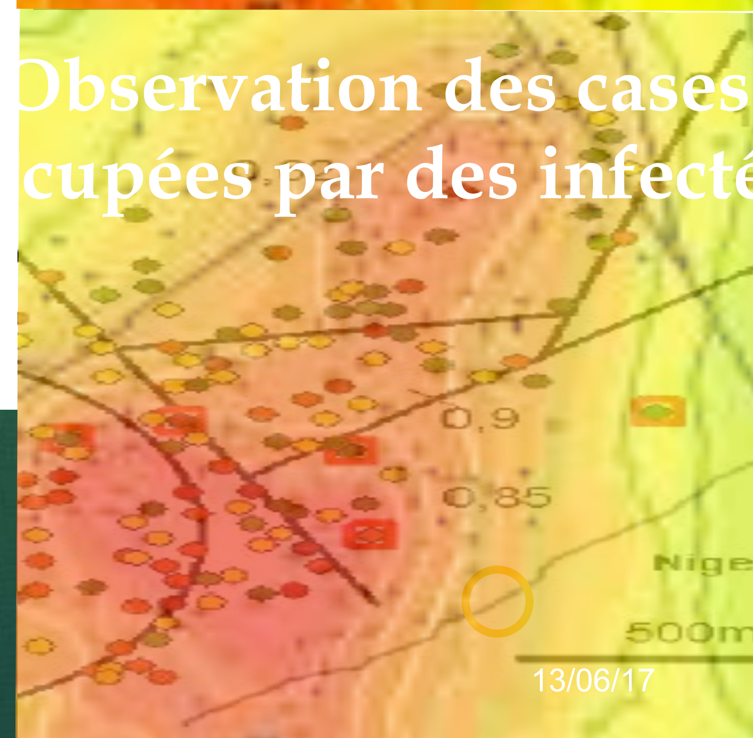
$$\frac{\partial A_s(x,t)}{\partial t} = \varpi - \alpha\zeta\ G(t)\ A_s(x,t) - \xi\ A_s(x,t) + D_i\ \Delta A_s(x,t)$$

$$\frac{\partial A_g(x,t)}{\partial t} = +\alpha\zeta\ G(t)\ A_s(x,t) - (\xi + \nu)\ A_g(x,t) + D_g\ \Delta A_g(x,t)$$

$$\frac{\partial A_i(x,t)}{\partial t} = \nu\ A_g(x,t) - \xi\ A_i(x,t) + D_s\ \Delta A_i(x,t)$$

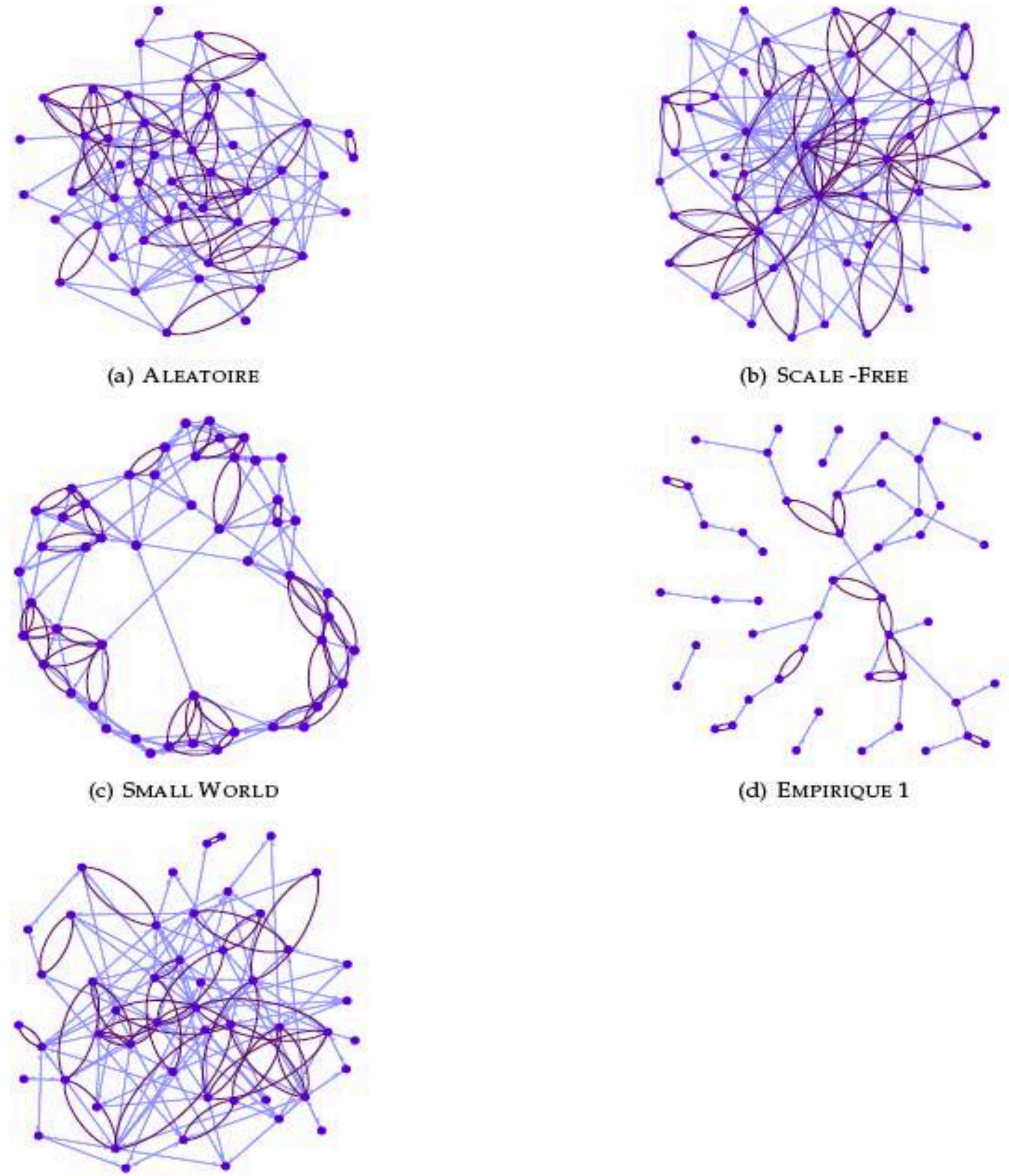
J. GAUDART,…, J. D. & O.K. DOUMBO, Modelling malaria incidence. *Malaria J.*, **8**, 61 (2009).





Automata modelling

Social networks



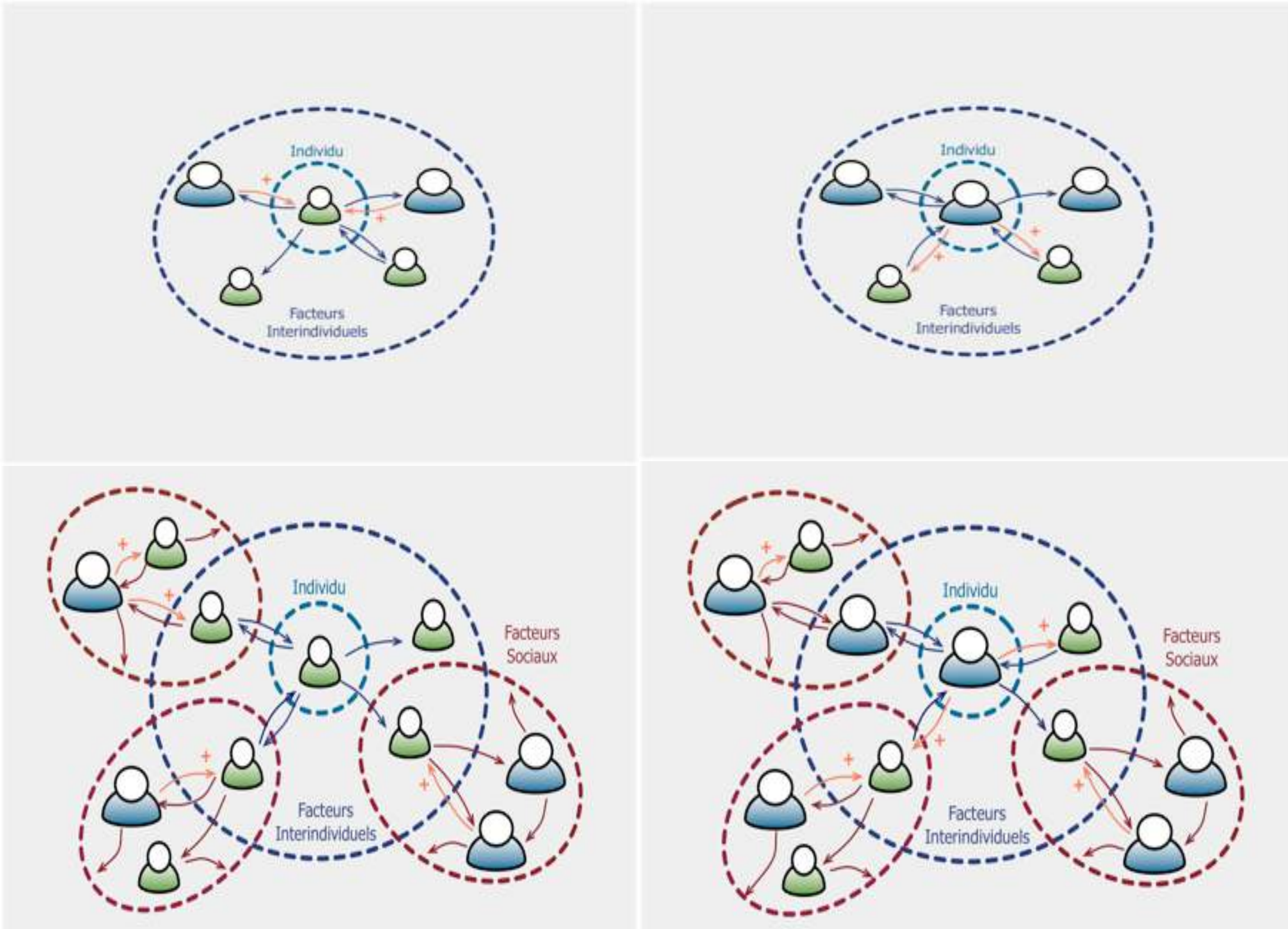
(a) ALEATOIRE

(b) SCALE -FREE

(c) SMALL WORLD

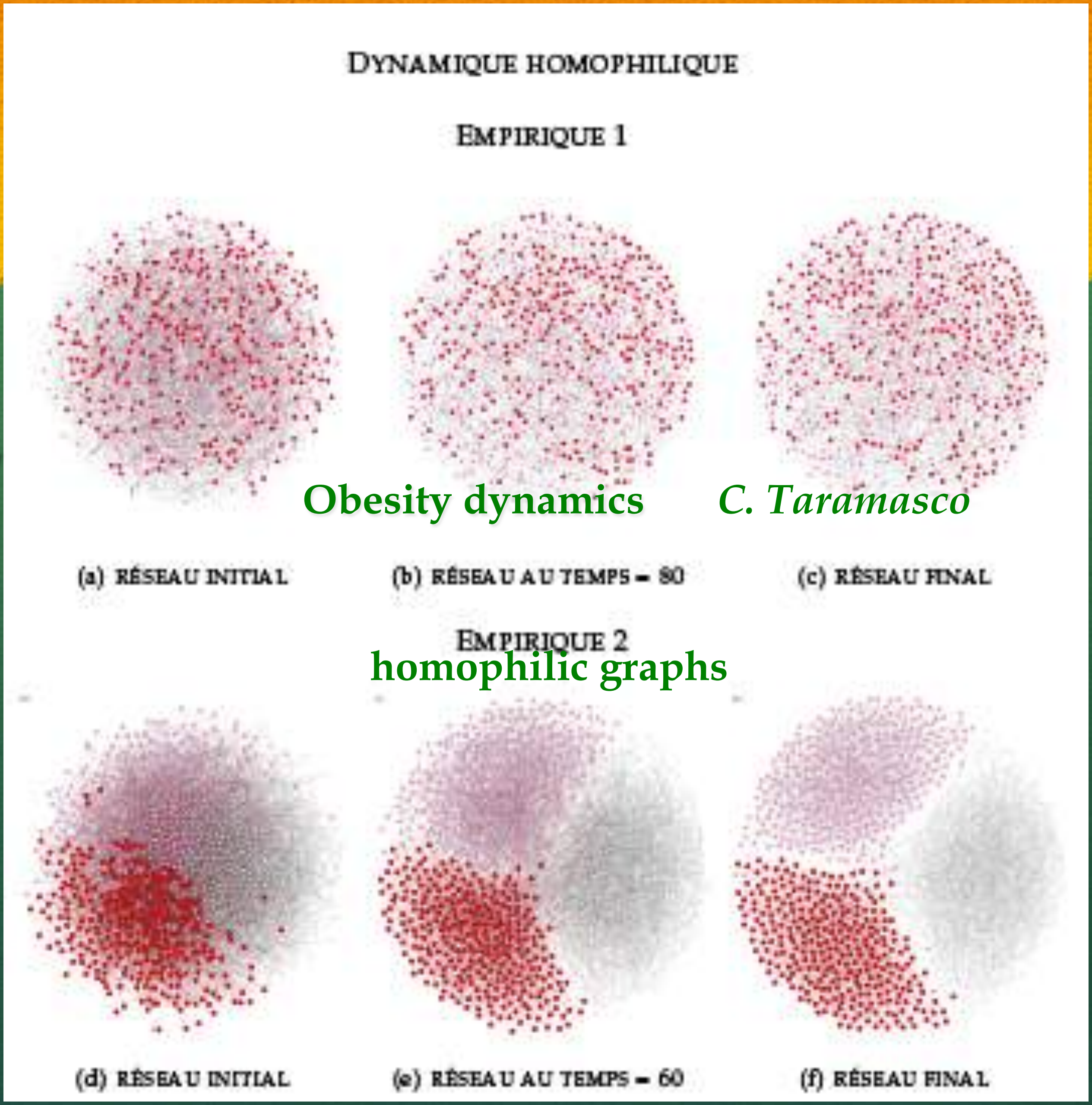
(d) EMPIRIQUE 1

(e) EMPIRIQUE 2



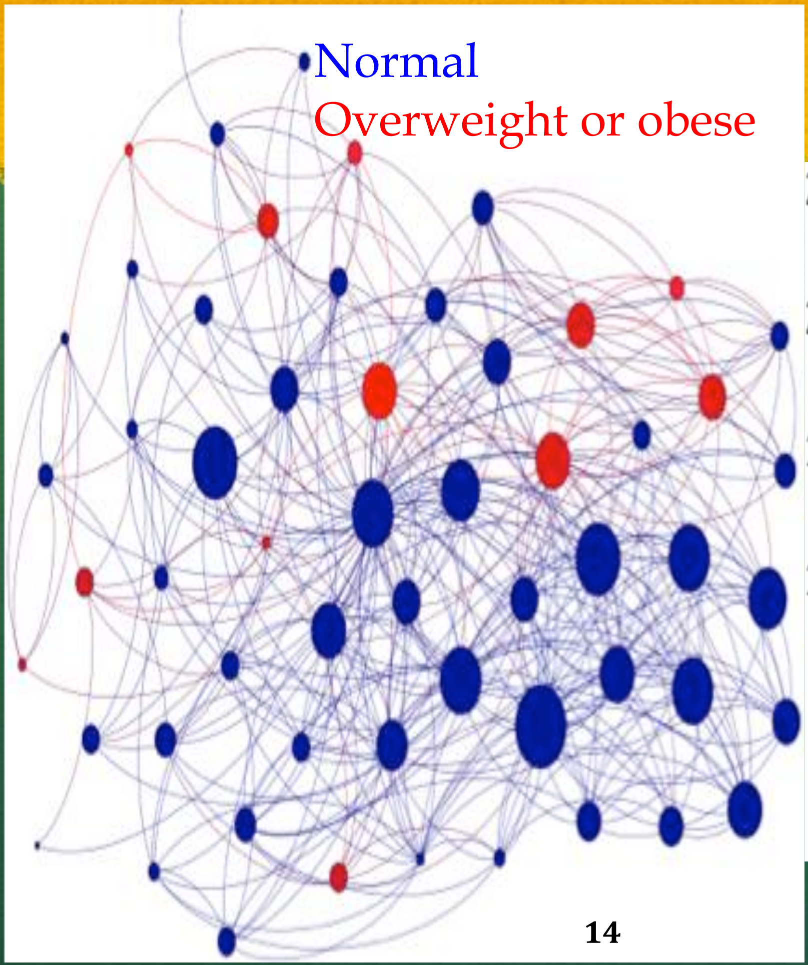
Empirical networks

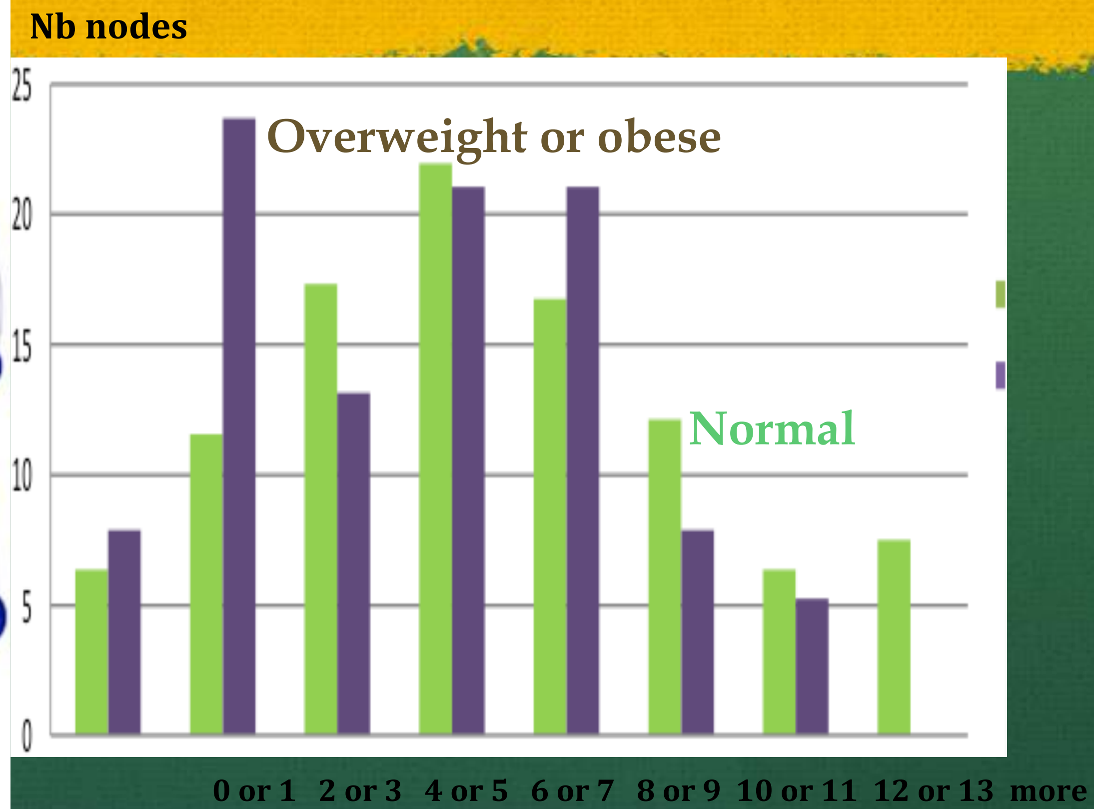
Obesity



DYNAMIQUE HOMOPHILIQUE
EMPIRIQUE 1
(a) RÉSEAU INITIAL
(b) RÉSEAU AU TEMPS = 80
(c) RÉSEAU FINAL
EMPIRIQUE 2
(d) RÉSEAU INITIAL
(e) RÉSEAU AU TEMPS = 60
(f) RÉSEAU FINAL
Obesity dynamics
C. Taramasco
homophilic graphs

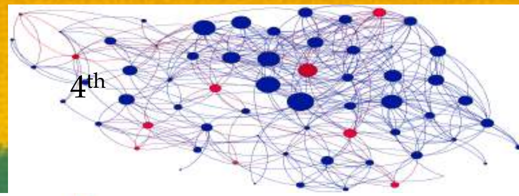
Friendship network at highschool

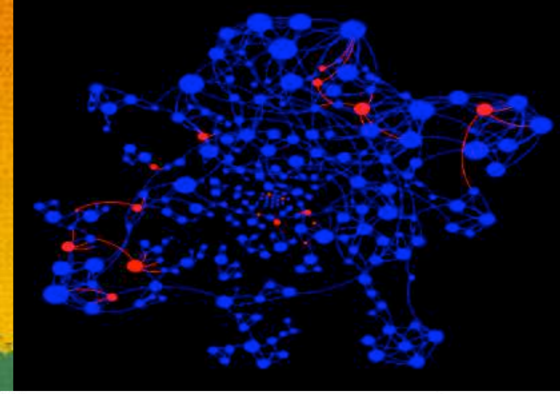


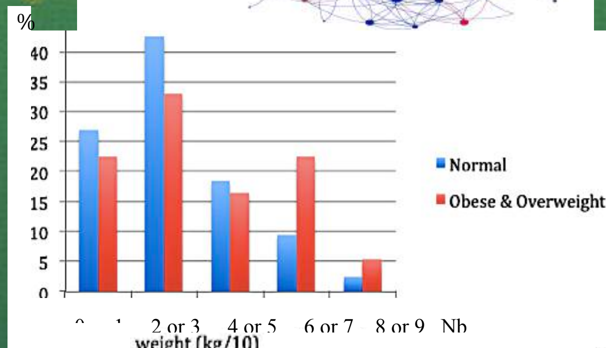


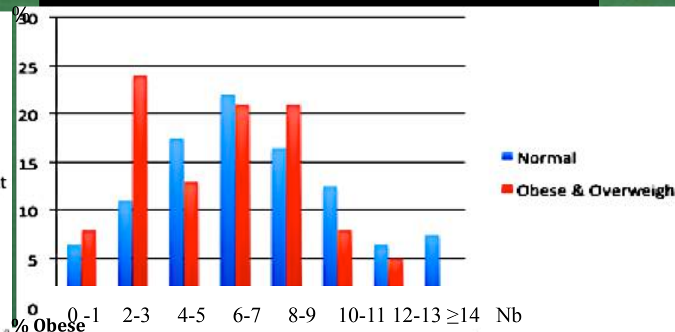
Bimodal distribution of friend's numbers in abnormal population of highschool


STEAMSURFING
BY VITALISEUR DE MARION
S'INSCRIRE
SE CONNECTER
FR
EN
Créez, partagez, vivez
vos exigences alimentaires

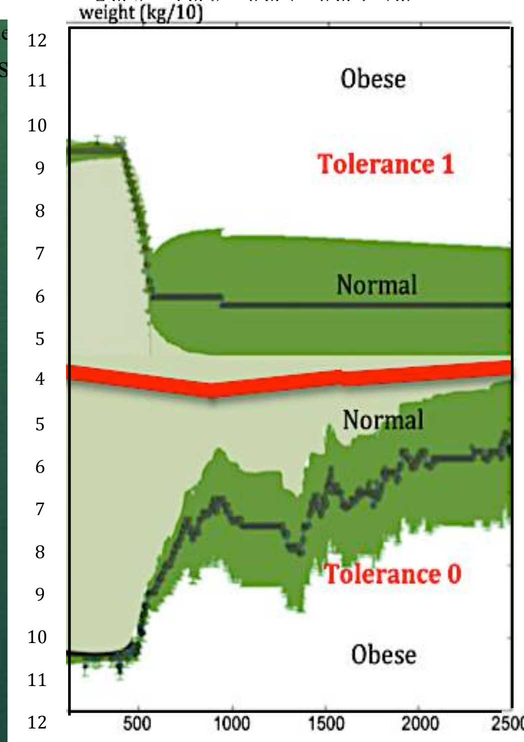


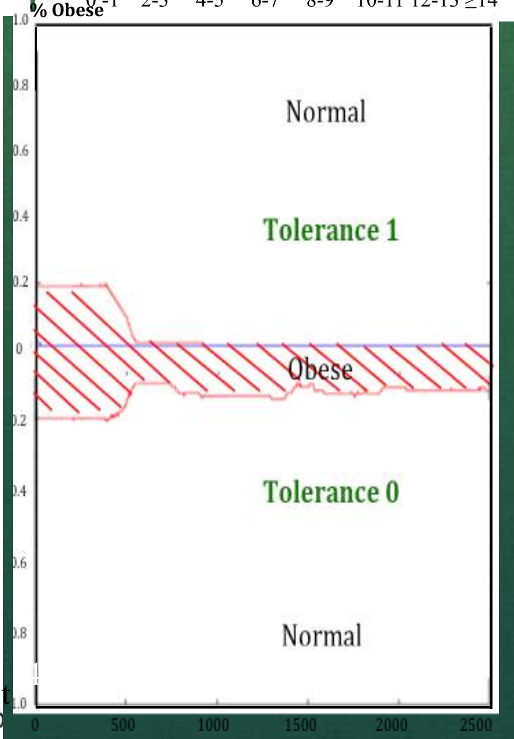






Suivi de l'obésité au collège





Functional modelling

For reducing big data related to

a metabolic pathway

into a few number of attractors

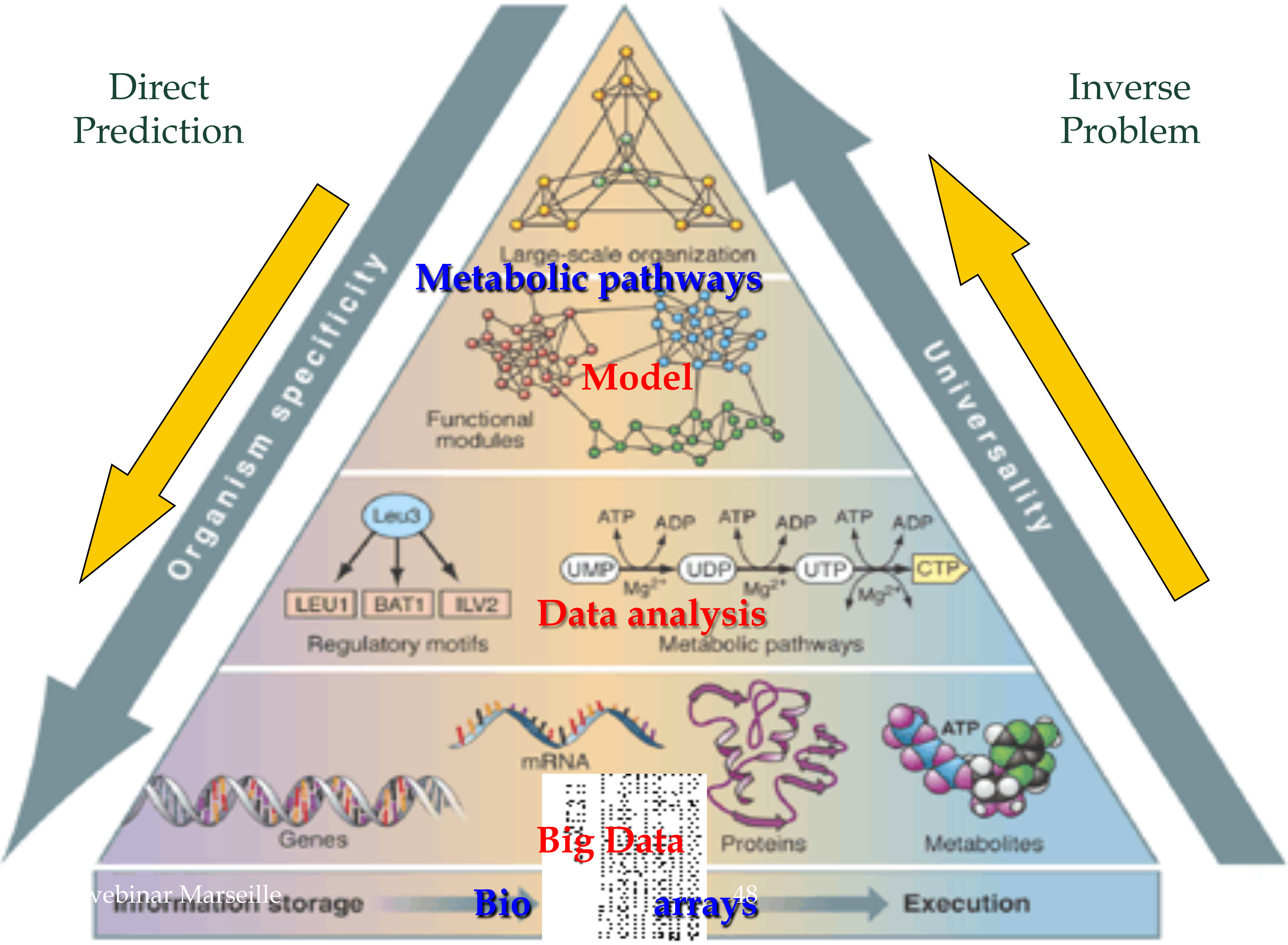
related to the circuits of its

interaction graph

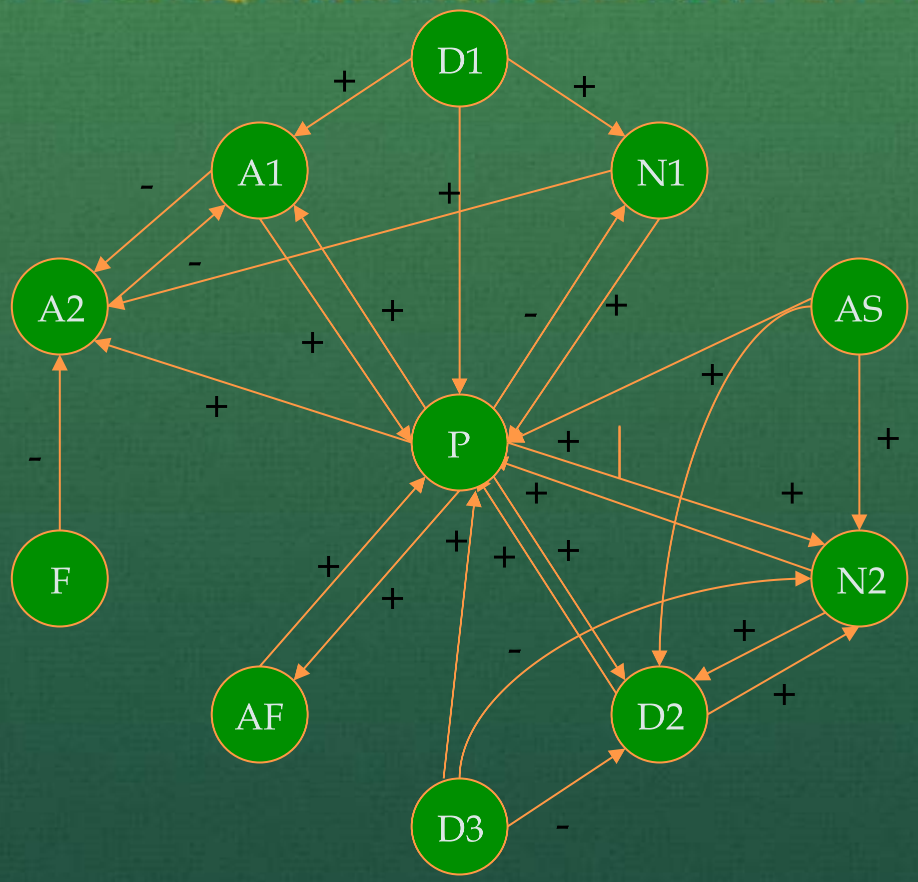
Life's complexity: from bio-arrays to metabolic pathways

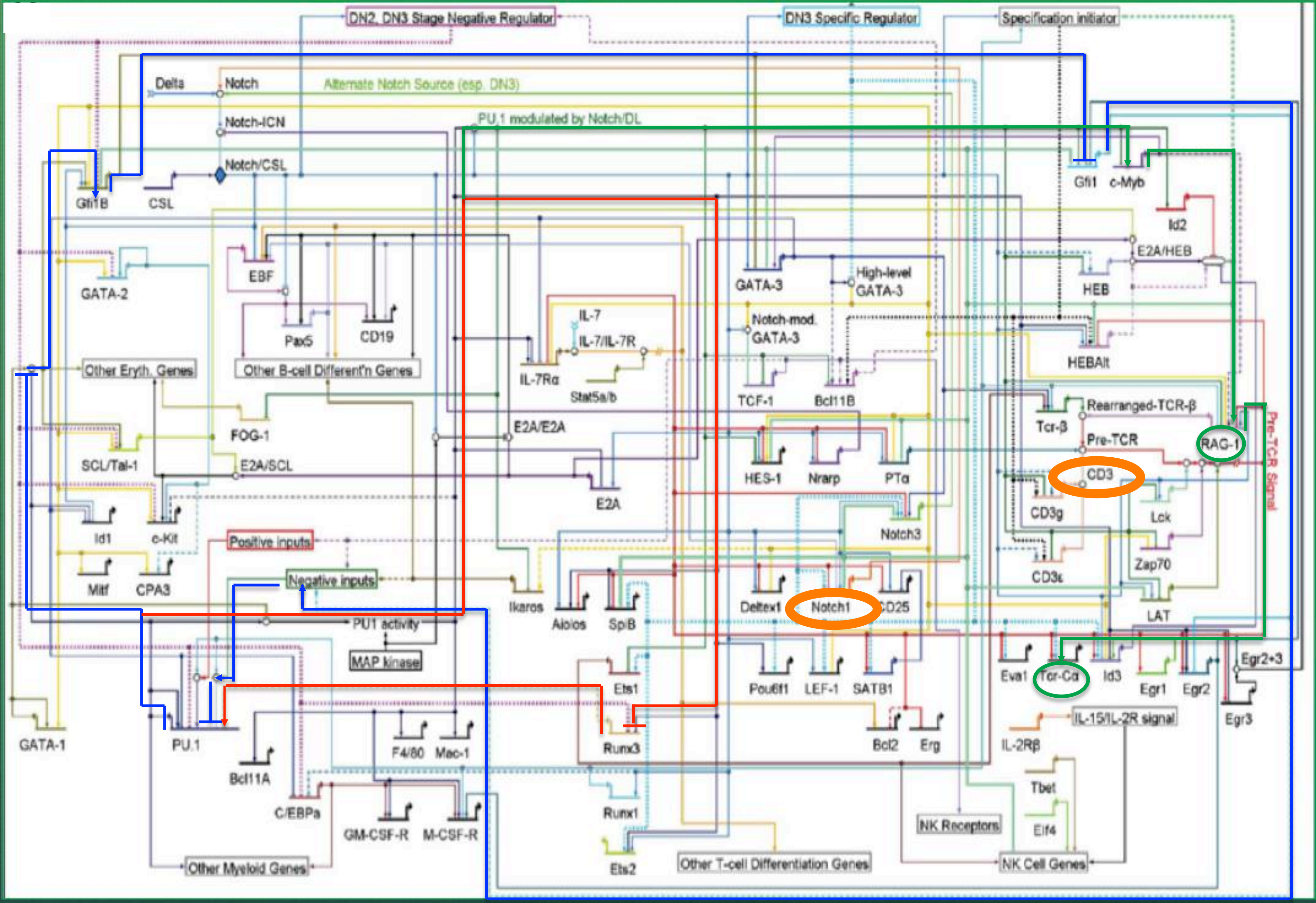
Direct Prediction

Inverse Problem



Social Interaction network





RAG control pathway

C. Georgescu, PNAS 2008; JD, A. Elena, M. Noual, S. Sené & F. Thuderoz J. Theor. Biol. 2011

Attractor Nb of positive circuits of length *r* tangent to negative of length *l*

$\ell \backslash r$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	T_ℓ^-
1	1																1
2	1	1															1
3	1	1	2														2
4	1	2	1	2													2
5	2	1	2	2	4												4
6	1	1	3	3	2	6											6
7	2	2	3	2	4	3	10										10
8	2	3	2	8	3	4	6	16									16
9	3	2	2	3	5	9	7	7	30								30
10	2	4	3	4	17	7	7	10	11	52							52
11	4	3	5	6	7	7	11	11	16	19	94						94
12	3	4	9	2	7	42	11	33	17	23	28	172					172
13	5	6	7	7	11	11	16	19	24	28	39	46	316				316
14	6	7	7	10	11	17	105	23	28	38	46	60	75	586			586
15	7	7	10	11	4	17	24	28	44	125	60	66	97				1096
16	7	10	11	33	19	23	28	278	46	60	75	88					2048
17	11	11	16	19	24	28	39	46	60	75	97						3856
18	11	17	17	23	28	6	46	60	729	96							7286
19	16	19	24	28	39	46	60	75	97								13798
20	19	23	28	32	125	60	75	88									26216
21	24	28	44	46	60	66	10										49940
22	28	38	46	60	75	96											95326
23	39	46	60	75	97												182362
24	46	60	66	88													349536
25	60	75	97														671092
26	75	96															1290556
27	97																2485534

$gcd(\ell, r) =$ 1, 2, 3, 4, 5, 6, 7, 8, 9

$\ell = r$

$\ell = 3 \times r / 2$

$\ell = 2 \times r$

$\ell = 4 \times r$

JD, M. NOUAL, S. SENE
Discrete Applied Maths 2014

Conclusion

- Toward a personalized patient time recording
- and at the population level, a spatiotemporal observation
- followed by the steps of big data compression, modelling and interpretation
- with restitution of information to patient and to population