Complex systems science: dream of universality, reality of interdisciplinarity

Sébastian Grauwin, Guillaume Beslon, Eric Fleury, Sara Franceschelli, Céline Robardet, Jean-Baptiste Rouquier & Pablo Jensen

Les Houches, Sept 2010

ENS DE LYON

Institut des Systèmes Complexes
Complex Systems Institute
Rhône-Alpes
IXXI

cnrs

INSA LYON

INRIA

Complex Systems "science" (CSS) ?

An epistemologic motivation: what IS the CSS?

- Reductionnism fruitful IF effective micro to macro link
- CSS micro to macro + universal principles?
- Look for "universal principles applying to systems in general" (von Bertalanffy, 1968)
- "[Complex networks science] suggests that nature has some universal organizational principles that might finally allow us to formulate a general theory of complex systems." (Solé, 2000)

Complex Systems "science" (CSS) ?

- Have universal principles been discovered?
- CSS able to link sociology, biology, physics and computer science - very different disciplines ?

- Use CSS tools (networks) to study the CSS field **objectively**!
  - Detect and characterize communities from CSS bibliographic data
  - Study these communities' characteristics, connections, evolution
- $\Rightarrow$ **Quantitative approach** renewing the "classical" qualitative / famous-author centered approaches

Outline

- Extraction of the dataset
- How do we build a “good” CSS community network?
- Quantitative tools to analyze the properties of this network
- Evolution of the CSS community network

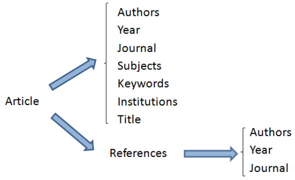
Dataset: articles from the "complex systems" science field

Data extracted from WoS with a **large** choice of Topic Keywords:

self organ*, complex network*, dynamical system, econophysics, strange attractor, synergetics, complex* and:

adaptive system*, artificial intelligence, attractor, bifurcation, chaos, control, criticality, ecology, economics, epistemology, far from equilibrium, feedback, fractal, ising, multi agent, multiagent, multi scale, multifractal, multiscale, neural network*, nonlinear dynamic*, non linear system*, nonlinear dynamic*, nonlinear system*, phase transition, plasticity, random walk, robustness, scaling, social system*, spin glass*, stability AND lyapunov, stochastic, synchronization, turbulence, universality, cell* automat*

Structure of the recorded data:



215 305 articles

4 050 318 distinct references

Building and analyzing the CS network

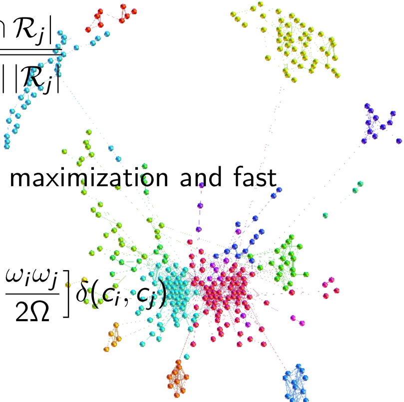
- Nodes = articles from a 10-years period
- Links by **bibliographic coupling**, weighted by Kessler (1963)]'s cosine similarity:

$$\omega_{ij} = \frac{|\mathcal{R}_i \cap \mathcal{R}_j|}{\sqrt{|\mathcal{R}_i|\,|\mathcal{R}_j|}}$$

  $\approx$ fraction of shared references.

- Community detection by modularity maximization and fast Louvain Algorithm (Blondel, 2008)

$$Q^\omega = \frac{1}{2\Omega} \sum_{ij} \left[\omega_{ij} - \frac{\omega_i \omega_j}{2\Omega} \right] \delta(c_i, c_j)$$

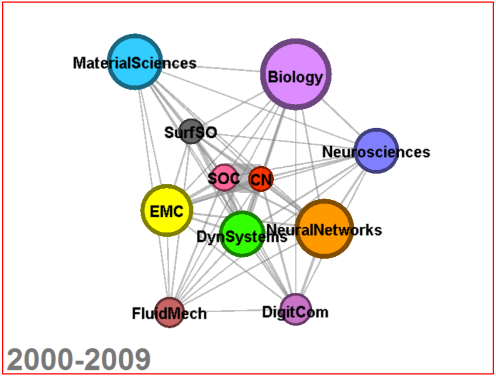


Discipline level

Some statistics

Decade	N	$< k >$	$< \omega_{ij} >^{-1}$	Q
1985-1994	20286	49.62	4484	0.7421
1990-1999	62040	99.61	7283	0.75
1995-2004	109458	115.68	11299	0.7592
2000-2009	141098	138.94	10848	0.7128

Discipline level



- 11 communities with more than 300 articles
- Surface nodes ∼ number articles in community
- Width links ∼ mean weight between articles pairs

Community characterization - an example

Complex Networks (CN)
2000-2009, down community $N = 3684, < \omega >^{-1} = 21.87$

Keyword	prop	σ
complex networks	0.53	260.804
dynamics	0.17	48.214
small-world networks	0.13	129.264
model	0.11	23
internet	0.1	96.383
networks	0.08	41.746
evolution	0.08	30.27
scale-free networks	0.07	98.245
systems	0.07	13.515
organization	0.07	44.6

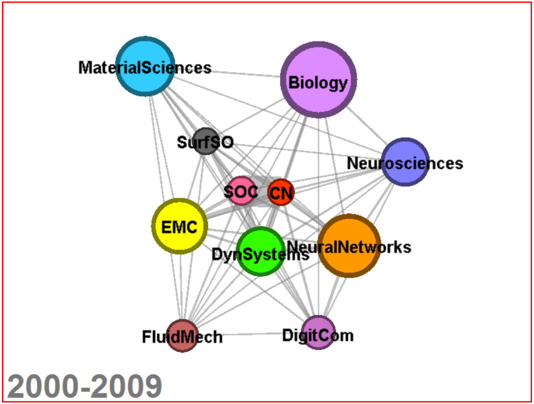
Refs	Times cited
Albert R, 2002, REV MOD PHYS (74), 47	2006
Barabasi AL, 1999, SCIENCE (286), 509	1659
Watts DJ, 1998, NATURE (393), 440	1556
Newman MEJ, 2003, SIAM REV (45), 167	1288
Strogatz SH, 2001, NATURE (410), 268	877
Dorogovtsev SN, 2002, ADV PHYS (51), 1079	671
Albert R, 2000, NATURE (406), 378	635
Jeong H, 2000, NATURE (407), 651	486
Albert R, 1999, NATURE (401), 130	445
Boccaletti S, 2006, PHYS REP (424), 175	420

Journal	prop	σ
PHYS REV E	0.331	119.35
PHYSICA A	0.217	91.76
LECT NOTE COMPUT SCI	0.082	15.05
PHYS REV LETT	0.067	37.27
EUR PHYS J B	0.065	64.21
PROC NAT ACAD SCI USA	0.041	16.63
INT J MOD PHYS C	0.035	39.6
CHIN PHYS LETT	0.031	40.65
CHAOS	0.029	28.95
IEEE INT SYMP CIRC SYST PROC	0.024	17.26

Refs (Journals)	Times cited
PHYS REV LETT	9246
NATURE	9101
PHYS REV E 2	8864
SCIENCE	5482
P NATL ACAD SCI USA	4450
PHYSICA A	3563
REV MOD PHYS	2240
PHYS REV E	2047
EUR PHYS J B	1707
SIAM REV	1401

$\Rightarrow$ Automatic labels!

Discipline level (2)



- Subject analysis: Center = theory / Periphery = "applications"
- Universal explanation ? Links between discipline?
- ⇒ Need to go deeper!

Discipline level (3)

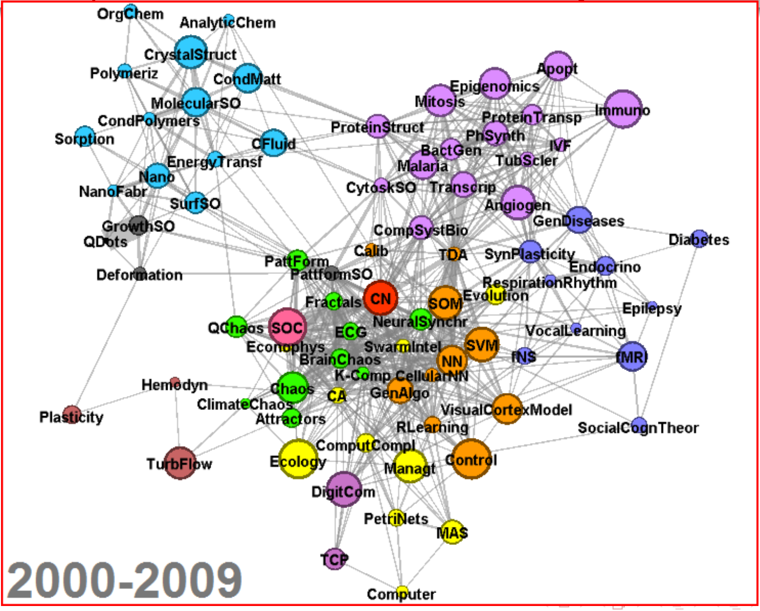
Community	N	$<\omega>^{-1}$	Q^i
Complex Networks (CN)	3684	21.87	0.087*
Neural Networks (NN)	21913	1759.12	0.532
Cellular and Molecular Biology (Biology)	32107	4852.62	0.615
Material Sciences	19531	3148.04	0.656
Dynamical Systems (DynSystems)	13115	1661.96	0.571
Self-Organized Criticality (SOC)	4447	199.3	0.202*
Ecology - Management - Computational Model (EMC)	16948	5560.05	0.708
Neurosciences	12567	3231.99	0.708
Control	6094	988.63	0.579
Surface Self-Organization (SurfSO)	3809	493.33	0.517
Fluid Mechanics (FluidMech)	5534	2689.56	0.685

Dense homogeneous communities

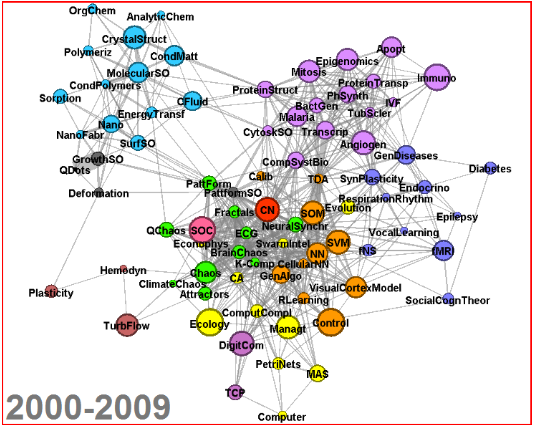
vs

sparser heterogeneous ones

Sub-discipline level : recursive modularity max



Sub-discipline level



- Interdisciplinary links through “trading zones”
- Self-organization everywhere?

Communities characterization

Neural Synchronization (NeuralSynchr)

$N =, < \omega_{in} >^{-1} = 453.59$

Keyword	prop	σ	Refs	Times cited
synchronization	0.11	50.301	Hodgkin AL, 1952, J PHYSIOL (117), 500	124
model	0.1	14.266	Gammaitoni L, 1998, REV MOD PHYS (70), 223	106
dynamics	0.08	12.855	Kelso JAS, 1995, DYNAMIC PATTERNS SEL	82
oscillations	0.06	37.556	Fitzhugh R, 1961, BIOPHYS J (1), 445	67
stochastic resonance	0.05	69.539	Gray CM, 1989, NATURE (338), 334	61
EEG	0.05	41.19	Thelen E, 1994, DYNAMIC SYSTEMS APPR	58
systems	0.05	3.526	Varela F, 2001, NAT REV NEUROSCI (2), 229	58
chaos	0.04	9.972	Singer W, 1995, ANNU REV NEUROSCI (18), 555	55
noise	0.04	21.803	Haken H, 1985, BIOL CYBERN (51), 347	53
self-organization	0.04	5.657	Rodriguez E, 1999, NATURE (397), 430	51

Journal	prop	σ	Refs (Journals)	Times cited
PHYS REV E	0.04	8.96	J NEUROPHYSIOL	1666
LECT NOTE COMPUT SCI	0.028	0.64	J NEUROSCI	1520
J NEUROPHYSIOL	0.017	14.15	NATURE	1231
NEUROCOMPUTING	0.014	9.69	SCIENCE	1168
J NEUROSCI	0.013	6.55	PHYS REV LETT	886
NEURAL COMPUT	0.013	13.71	P NATL ACAD SCI USA	853
AIP CONF PROC	0.013	4.75	J PHYSIOL-LONDON	739
PHYSICA A	0.012	1.14	BIOL CYBERN	649
CHAOS	0.011	9.29	ELECTROEN CLIN NEURO	513
CLIN NEUROPHYSIOL	0.011	14.72	NEUROSCIENCE	480

$\Rightarrow$ labels $\sim$ automatic at this level

Trading zones (1)

Exchanges between disciplines : analogy with pidgin (Galison, 1997)

Formalization of the concept:

- Trading force = fraction of the total weight between communities $\bar{I}$ and $\bar{J}$ which goes through sub-community I.

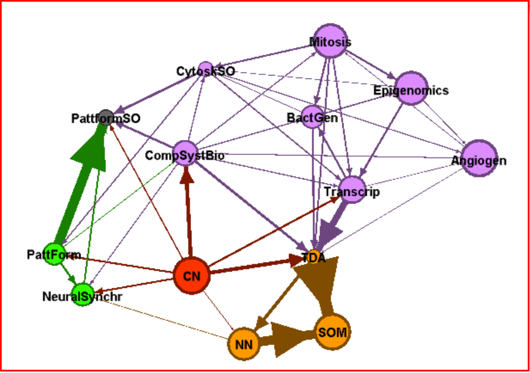
$$T_{I,(\bar{I},\bar{J})} = \frac{\sum_{i \in I, j \in \bar{J}} \omega_{ij}}{\sum_{i \in \bar{I}, j \in \bar{J}} \omega_{ij}}$$

- Tested against its expected value $T^{exp}_{I,(\bar{I},\bar{J})} = N_I / N_{\bar{I}}$

$\Rightarrow$ automatic detection of trading zones!

Trading zones (2)

Community	Bridge	T (%)	T/T^{exp}
TDA	bio/nn	46.355	20.51
CompSystBio	bio/sd	49.704	8.87
CompSystBio	bio/soc	49.255	8.79
ProteinStruct	bio/chem	47.442	8.32
CompSystBio	bio/cn	42.931	7.66
TDA	cn/nn	16.543	7.32
Hemodyn	ns/turb	54.552	7.22
NeuralSynchr	ns/sd	59.788	5.2
Hemodyn	bio/turb	39.113	5.18
CytoskSO	bio/soc	9.561	4.71
CytoskSO	bio/sd	9.522	4.69
Calib	chem/nn	8.717	4.51
CellularNN	cn/nn	9.605	4.3
Transcrip	bio/cn	25.806	4.05



- Transcriptomics data analysis : uses SOM tools, data interesting for biologists
- Computational systems biology : more speculative !
- Claims to universality : rhetorics to legitimate simple models

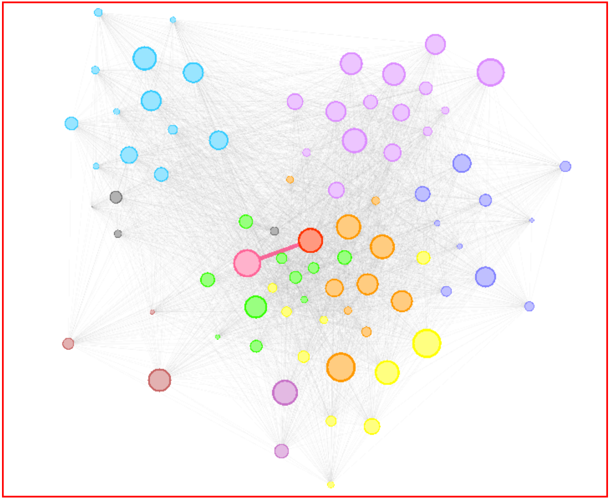
Coherence of the whole CSS ?

Detecting references shared by different communities

- Each community-community link is normalized to 1
- Count the fraction of article-article links in which reference r is present
- $\Rightarrow$ Local "gluing force" $g_{IJ}(r)$

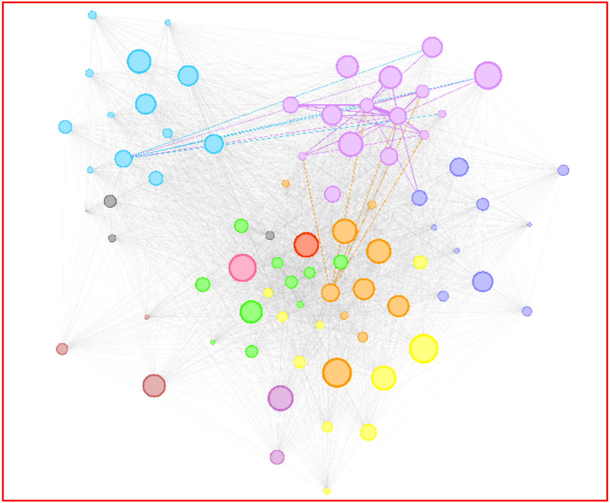
Local gluing force - $g_{IJ}(r) > 0$

Albert & Barabasi, 2002, REV MOD PHYS



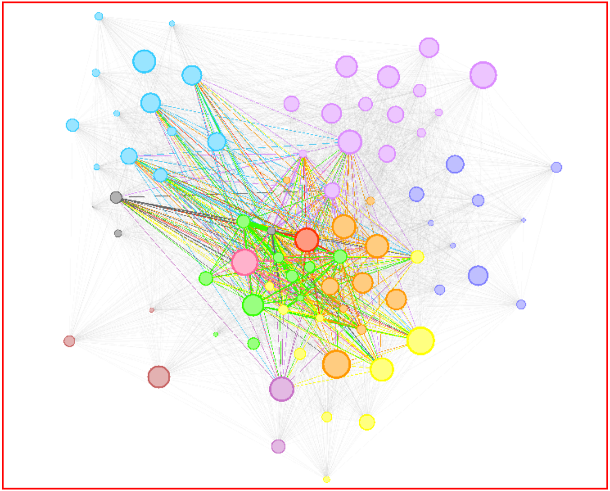
Local gluing force - $g_{IJ}(r) > 0$

Laemmli, 1977, NATURE



Local gluing force - $g_{IJ}(r) > 0$

Nicolis & Prigogine, 1977, Self Orga in Non-Equilibrium Systems



Local gluing force - $g_{IJ}(r) > 0$

Press WH, Numerical Recipes

Coherence of the whole CSS ?

The 20 most gluing references

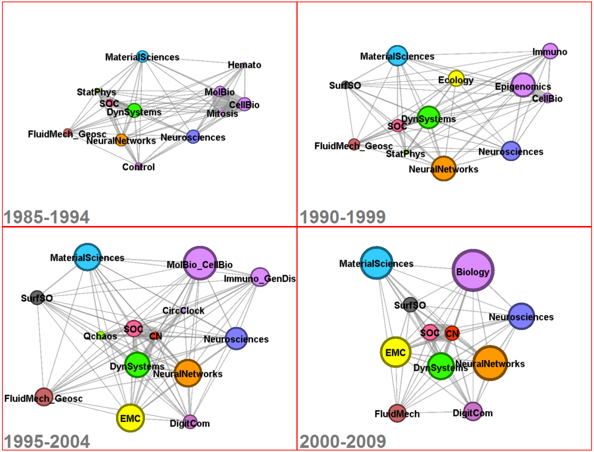
$$Glue(r) = \frac{1}{Z} \sum_{I \neq J} g_{IJ}(r)$$

Reference	*Glue* (%)
Press WH, NUMERICAL RECIPES*	1.250
Metropolis N, 1953, J CHEM PHYS, 21, 1087	0.509
Shannon CE, 1948, BELL SYST TECH J, 27, 379	0.445
Nicolis G, 1977, SELF ORG NONEQUILIBR*	0.420
Kauffman SA, 1993, ORIGINS ORDER SELF O*	0.309
Hebb DO, 1949, ORG BEHAV	0.297
Alberts B, 1994, MOL BIOL CELL	0.288
Abramowitz M, HDB MATH FUNCTIONS*	0.269
Feller W, INTRO PROBABILITY TH*	0.268
Watson JD, 1953, NATURE, 171, 737	0.250
Lakowicz JR, 1999, PRINCIPLES FLUORESCE	0.249
Turing AM, 1952, PHILOS T ROY SOC B, 237, 37	0.237
Witten TA, 1981, PHYS REV LETT, 47, 1400	0.234
Cohen J, 1988, STAT POWER ANAL BEHA	0.223
Hopfield JJ, 1982, P NATL ACAD SCI USA, 79, 2554	0.217
Stanley HE, 1971, INTRO PHASE TRANSITI	0.202
Whitesides GM, 2002, SCIENCE, 295, 2418	0.188
Marquardt DW, 1963, J SOC IND APPL MATH, 11, 431	0.174
Chomczynski P, 1987, ANAL BIOCHEM, 162, 156	0.167
Shannon CE, MATH THEORY COMMUNIC*	0.162

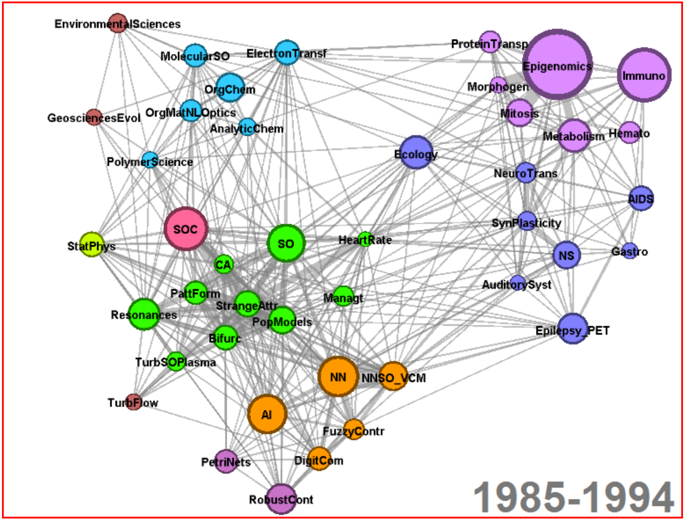
- Computer recipes ! CSS : transforming complex systems into complicated ones? Engineering vision of society
- Self-organized micro-macro defines CSS ? Specific applications of SO!

Evolution of the CSS network: discipline level

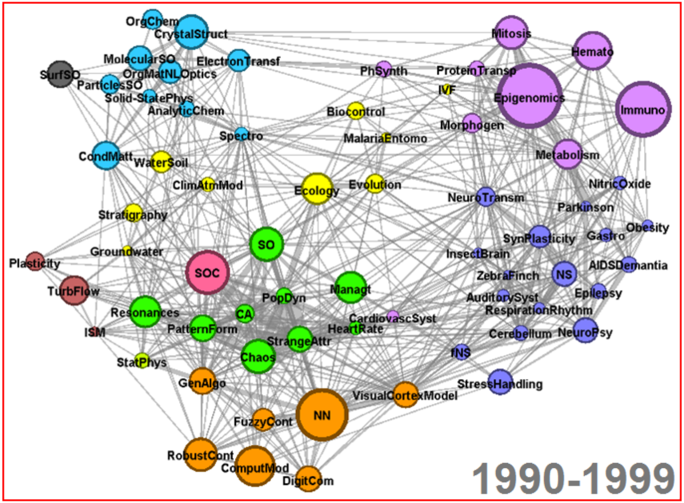
- History of CSS : where do CN come from?
- How do sub-disciplines evolve ?



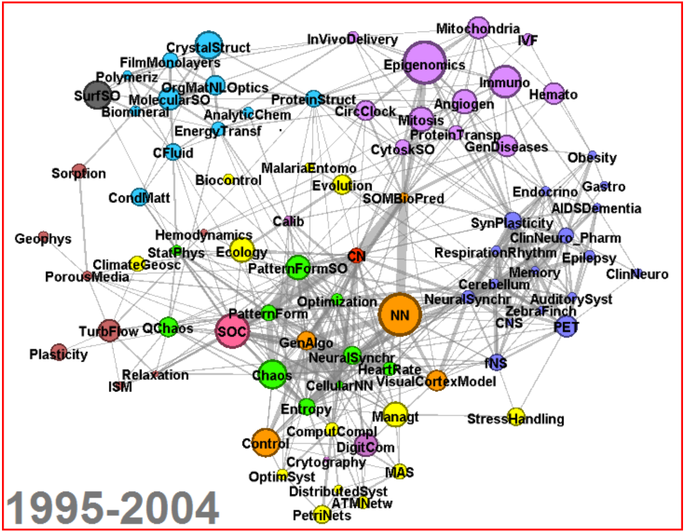
Evolution of the CSS network: subdiscipline level



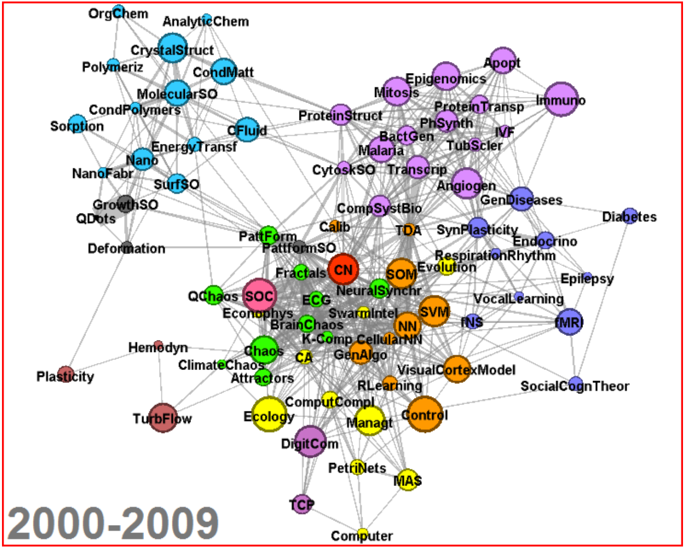
Evolution of the CSS network: subdiscipline level



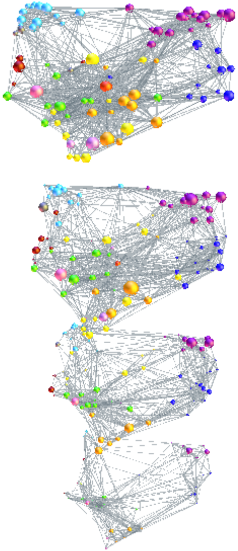
Evolution of the CSS network: subdiscipline level

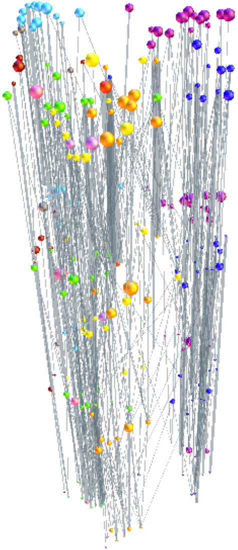


Evolution of the CSS network: subdiscipline level

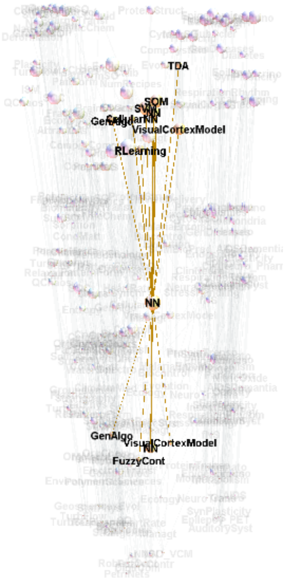


Inter-decade Bibliographic Coupling





Inter-decade Bibliographic Coupling



- Bibliographic coupling leads to the visualization of Normal Science (Kuhn)
- Allows to give the natural predecessors/successors of a community
- Inter-decade comparison can be achieved with other kinds of links

Summary

- Bibliometric coupling + community detection
  - provide for a global AND detailed view of the CSS field
  - represents “normal science”, giving equal weight to all published papers

- An OBJECTIVE view of the reality of CSS.
  - Theoretical (“imperialistic”) center vs application peripheral field
  - Knowledge transfer via “Trading zones” (TDA, Systems Bio, cytoskSO,...)
  - No universal principle: the CSS coherence is ensured by practical methods of numerical computation and specific application of SO.

Future work

- On the dynamic:
  - following each author carrier $\Rightarrow$ homonyms problems
  - analyzing the evolution of the topics of interest of a community via its keywords $\Rightarrow$ thesis for an historian of science
- Apply our methods to other themes / fields of science
- Compare our results with co-citation graphs $\Rightarrow$ "buzz science" *vs* "normal science".

The end

Thank you for your attention!

PS: come see me to check where you are on our maps...