

Multilevel Analysis of Co-authorship Networks

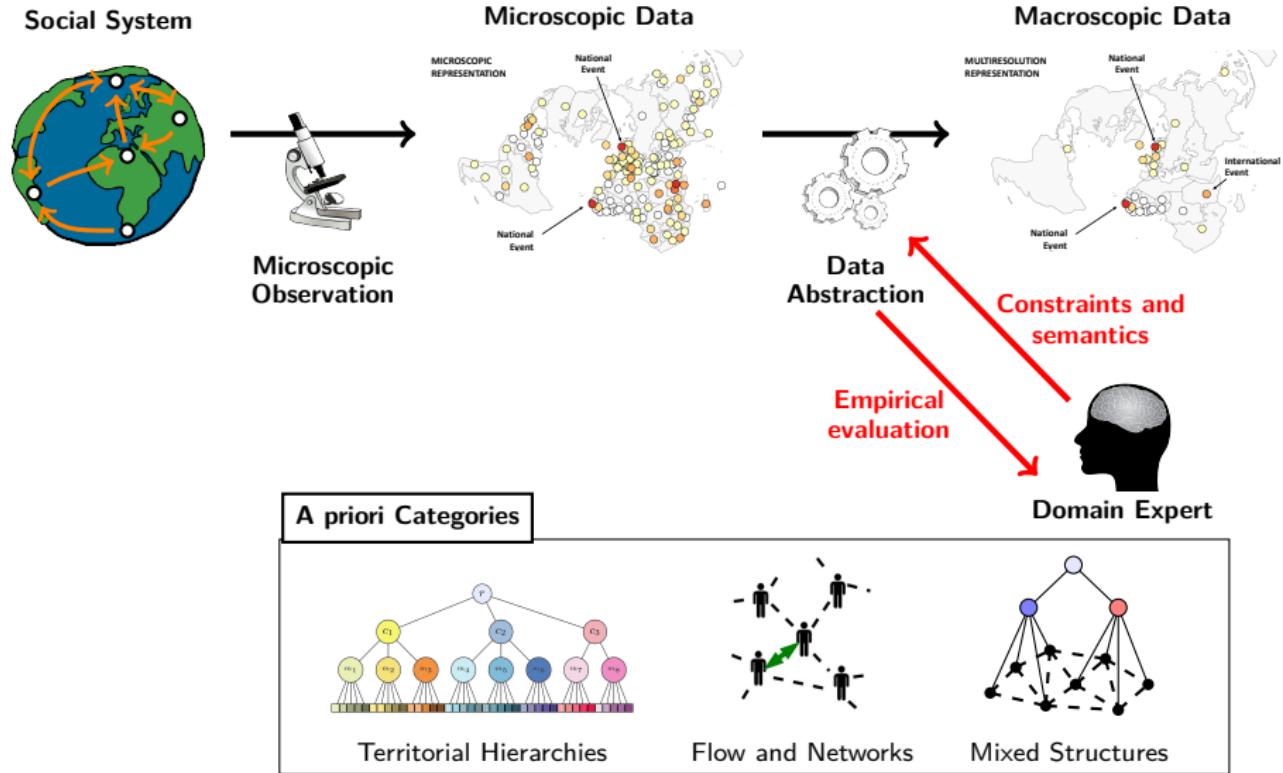
Robin Lamarche-Perrin, Lionel Tabourier, Fabien Tarissan



complexnetworks.fr



Multilevel Analysis of Social Systems



Data Abstraction

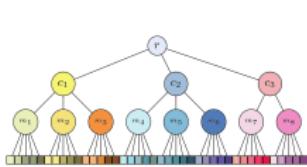
Constraints and semantics

Empirical evaluation

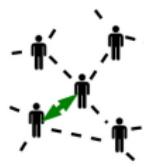


Domain Expert

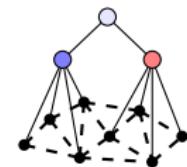
A priori Categories



Territorial Hierarchies



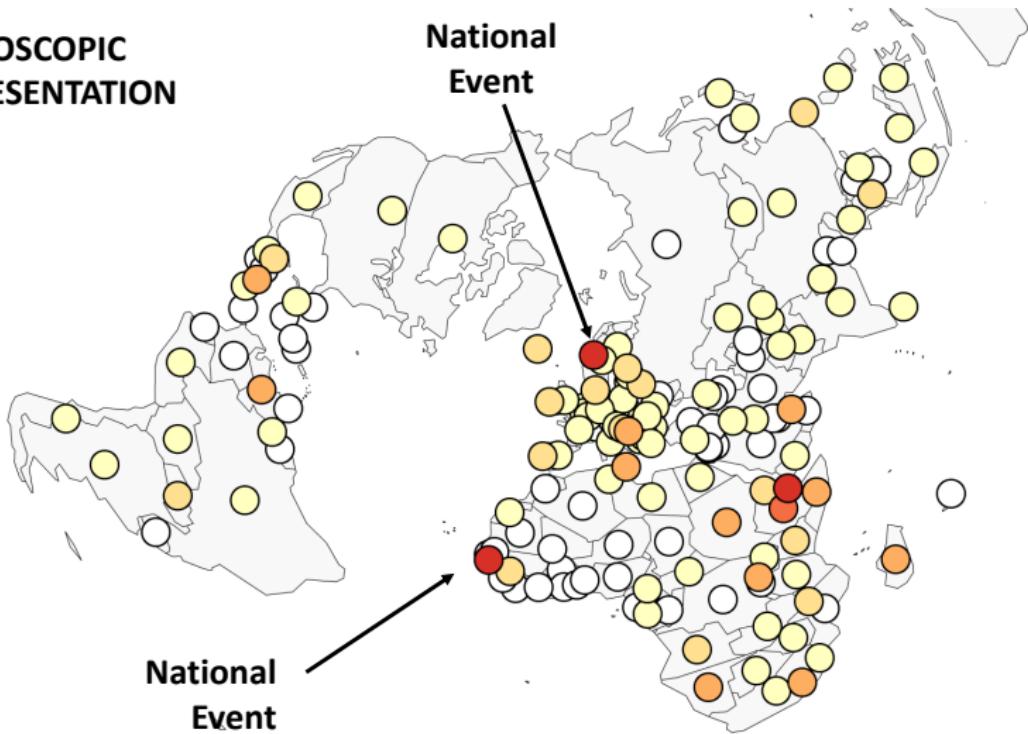
Flow and Networks



Mixed Structures

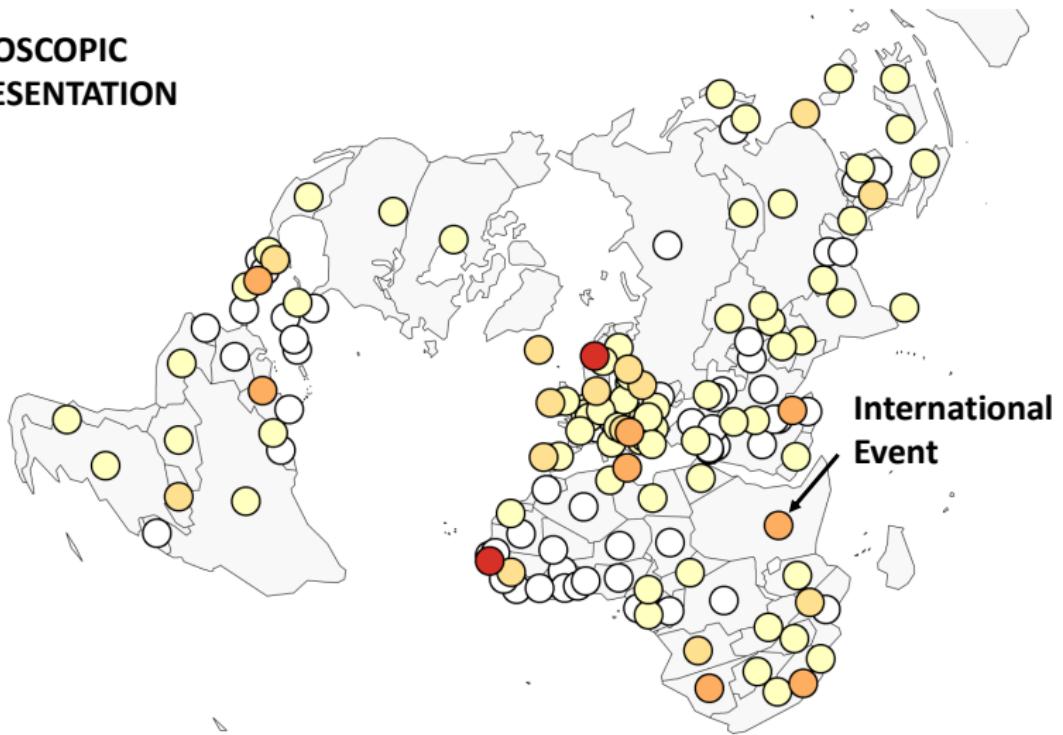
Example: Looking for Geographical Aggregates...

MICROSCOPIC
REPRESENTATION



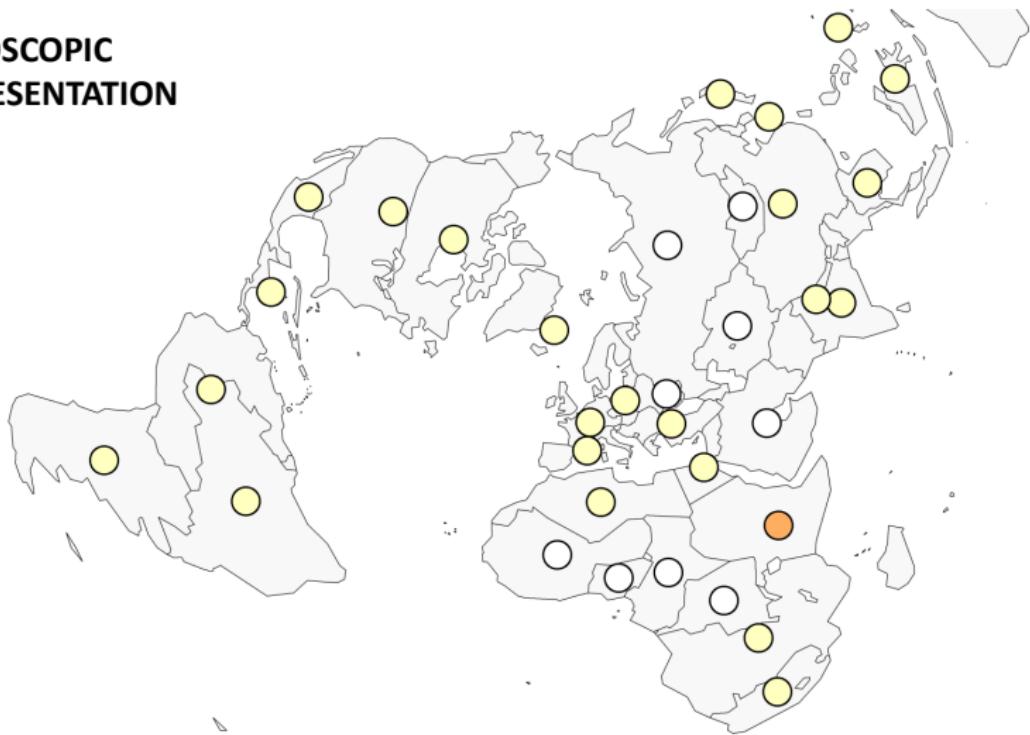
Example: Looking for Geographical Aggregates...

MICROSCOPIC REPRESENTATION



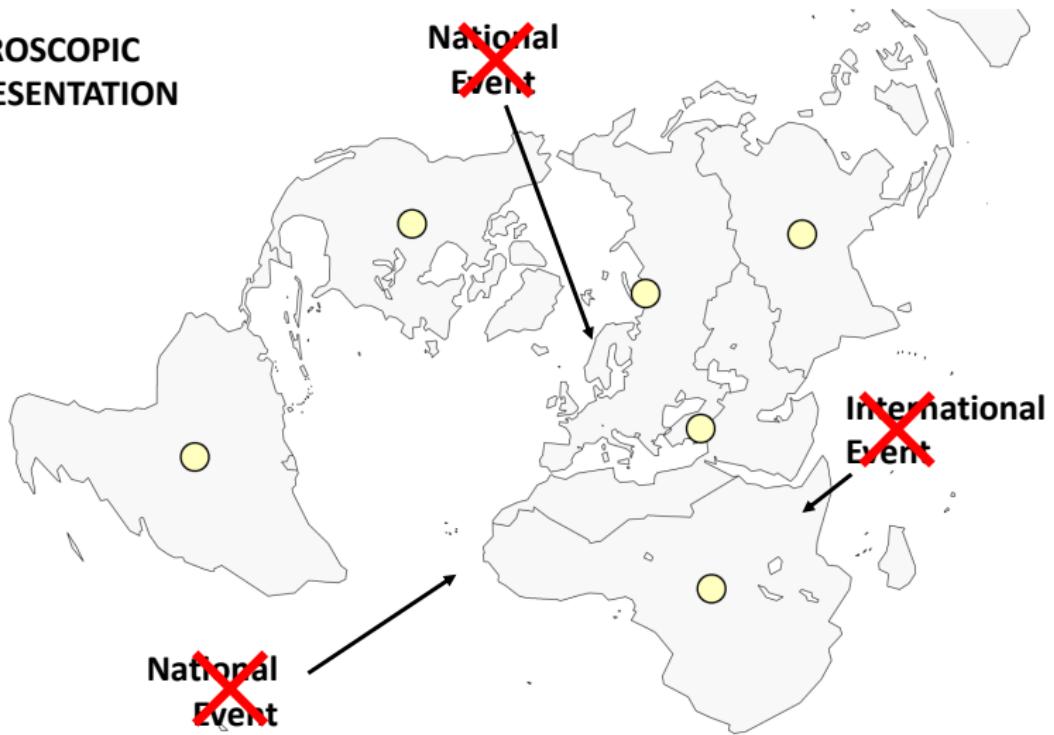
Problem I: Information of Geographical Aggregates

MESOSCOPIC REPRESENTATION



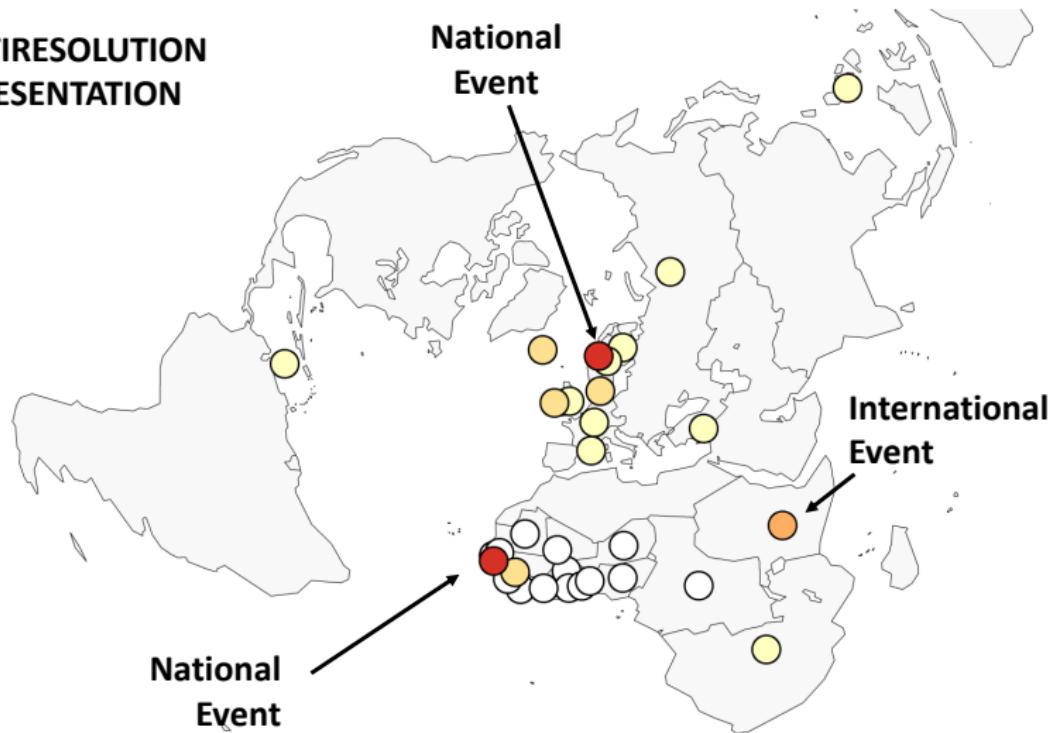
Problem I: Information of Geographical Aggregates

**MACROSCOPIC
REPRESENTATION**

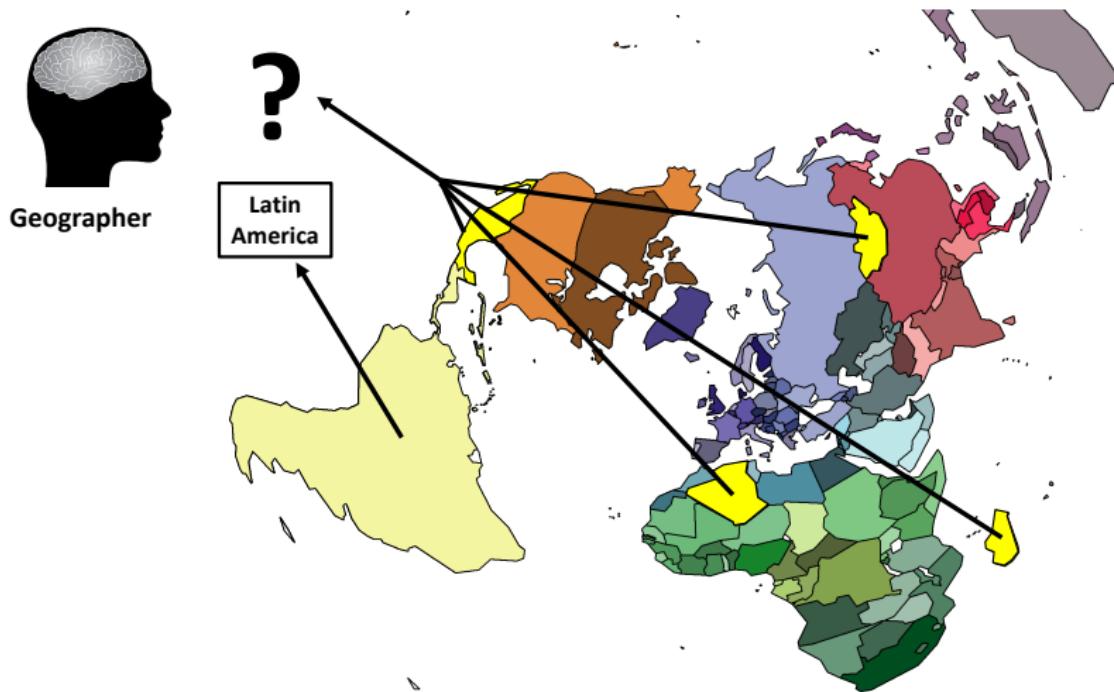


Problem I: Information of Geographical Aggregates

MULTIRESOLUTION
REPRESENTATION



Problem II: Semantics of Geographical Aggregates



Outline

Method:

Problem I. Measuring scales with information theory

Problem II. Formalising semantics with combinatorial constraints

Objective:

Data aggregation for the multiscale analysis of **social interactions** with applications to the study of scientific research

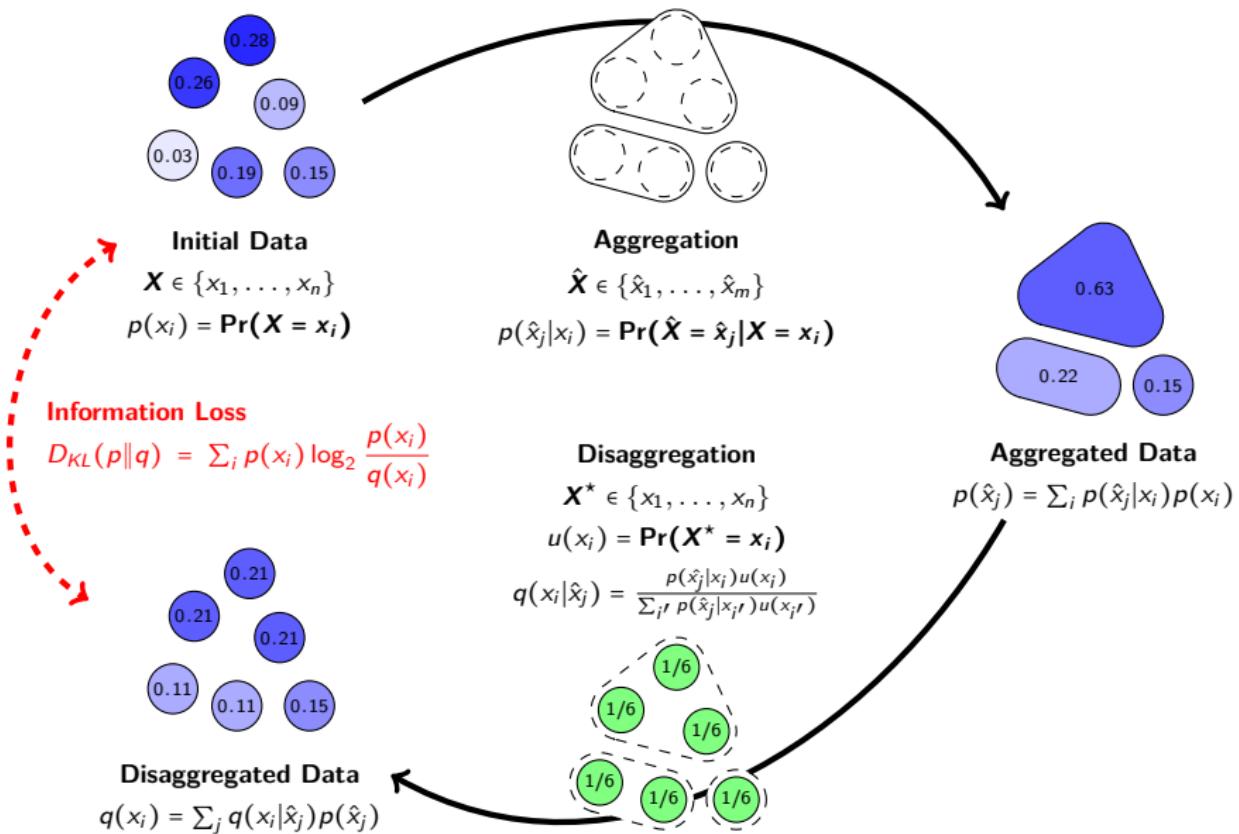
Discussion:

Evaluating *a priori* categories provided by the expert domain

Part I

Measuring scales with information theory

Data Aggregation and Information Loss

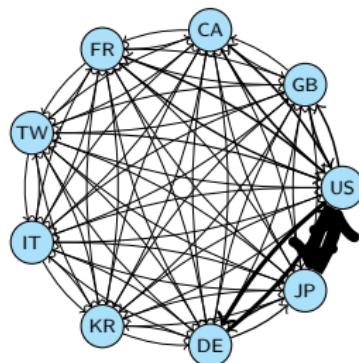


Aggregation of Social Relations

Adjacency matrix

	GB	CA	FR	TW	IT	KR	DE	JP	US
GB	3	5	1	2	0	11	23	82	
CA	3		2	1	0	6	15	89	
FR	5	3		1	3	1	14	28	83
TW	2	3	2		1	3	4	22	62
IT	2	1	3	1		0	7	12	31
KR	2	1	2	2	1		3	47	44
DE	11	6	12	2	6	1		78	167
JP	24	14	23	9	9	14	66		504
US	86	87	75	37	29	16	161	519	

Weighted directed graph



National Patent Citations. Unit: 100 patents; Period: 1990–1999;

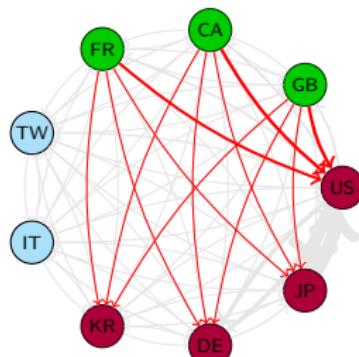
Source: NBER U.S. Patent Citations Data File (<http://www.nber.org/patents/>)

Aggregation of Social Relations

Adjacency matrix

	GB	CA	FR	TW	IT	KR	DE	JP	US
GB		3	5	1	2	0	11	23	82
CA	3		3	2	1	0	6	15	89
FR	5	3		1	3	1	14	28	83
TW	2	3	2		1	3	4	22	62
IT	2	1	3	1		0	7	12	31
KR	2	1	2	2	1		3	47	44
DE	11	6	12	2	6	1		78	167
JP	24	14	23	9	9	14	66		504
US	86	87	75	37	29	16	161	519	

Weighted directed graph



National Patent Citations. Unit: 100 patents; Period: 1990–1999;

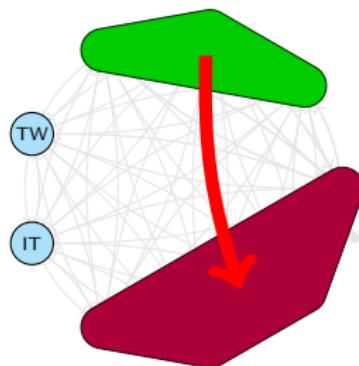
Source: NBER U.S. Patent Citations Data File (<http://www.nber.org/patents/>)

Aggregation of Social Relations

Adjacency matrix

	GB	CA	FR	TW	IT	KR	DE	JP	US
GB		3	5	1	2				
CA	3			3	2	1			
FR	5	3			1	3			
TW	2	3	2		1	3	4	22	62
IT	2	1	3	1		0	7	12	31
KR	2	1	2	2	1		3	47	44
DE	11	6	12	2	6	1		78	167
JP	24	14	23	9	9	14	66		504
US	86	87	75	37	29	16	161	519	

Weighted directed graph



National Patent Citations. Unit: 100 patents; Period: 1990–1999;

Source: NBER U.S. Patent Citations Data File (<http://www.nber.org/patents/>)

Aggregation of Social Relations

	GB	CA	FR	TW	IT	KR	DE	JP	US
GB						31	88	316	
CA		22							
FR				19					
TW									
IT	18						12	31	
KR					14			91	
DE		38					78	167	
JP	61	18	66						
US	248	66	30	161			1023		

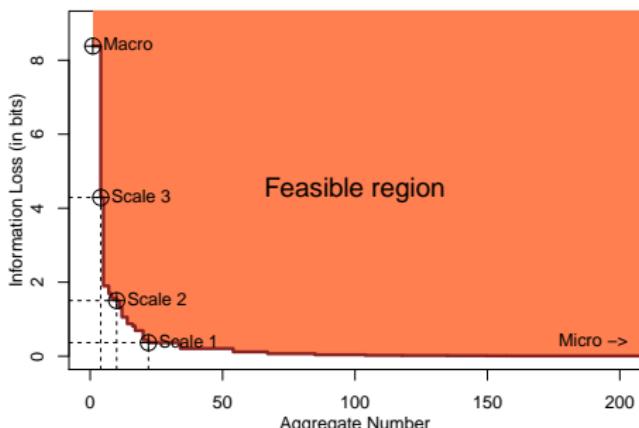
Scale 1: 22 aggregates
 → 96% of micro-information

	GB	CA	FR	TW	IT	KR	DE	JP	US
GB									
CA									
FR									
TW									
IT									
KR									
DE									
JP									
US									

Scale 2: 10 aggregates
 → 82% of micro-information

	GB	CA	FR	TW	IT	KR	DE	JP	US
GB									
CA									
FR									
TW									
IT									
KR									
DE									
JP									
US									

Scale 3: 4 aggregates
 → 49% of micro-information

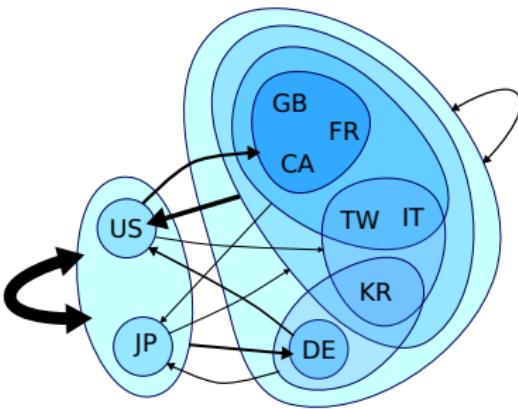


Aggregation of Social Relations

Adjacency matrix

	GB	CA	FR	TW	IT	KR	DE	JP	US
GB									
CA									
FR									
TW				142					
IT									
KR						125			
DE							167		
JP			93						
US	248	82			227		1023		

Weighted directed graph



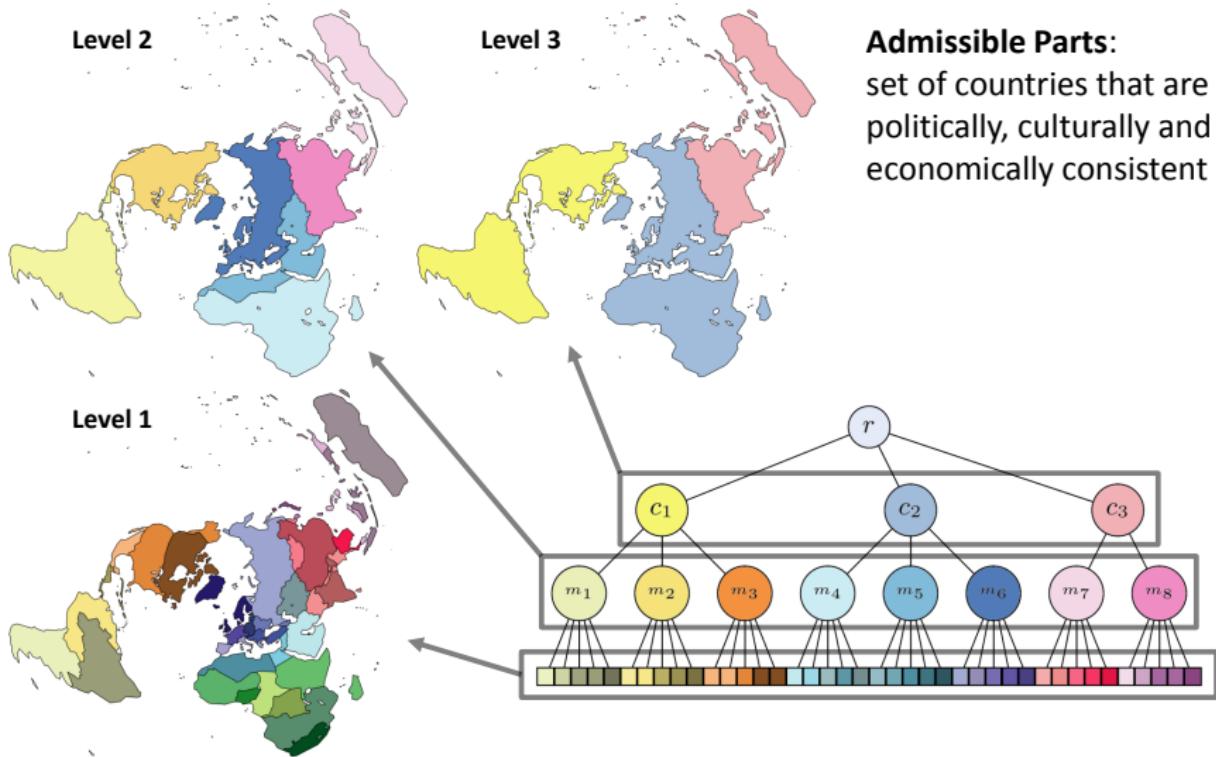
National Patent Citations. Unit: 100 patents; Period: 1990–1999;

Source: NBER U.S. Patent Citations Data File (<http://www.nber.org/patents/>)

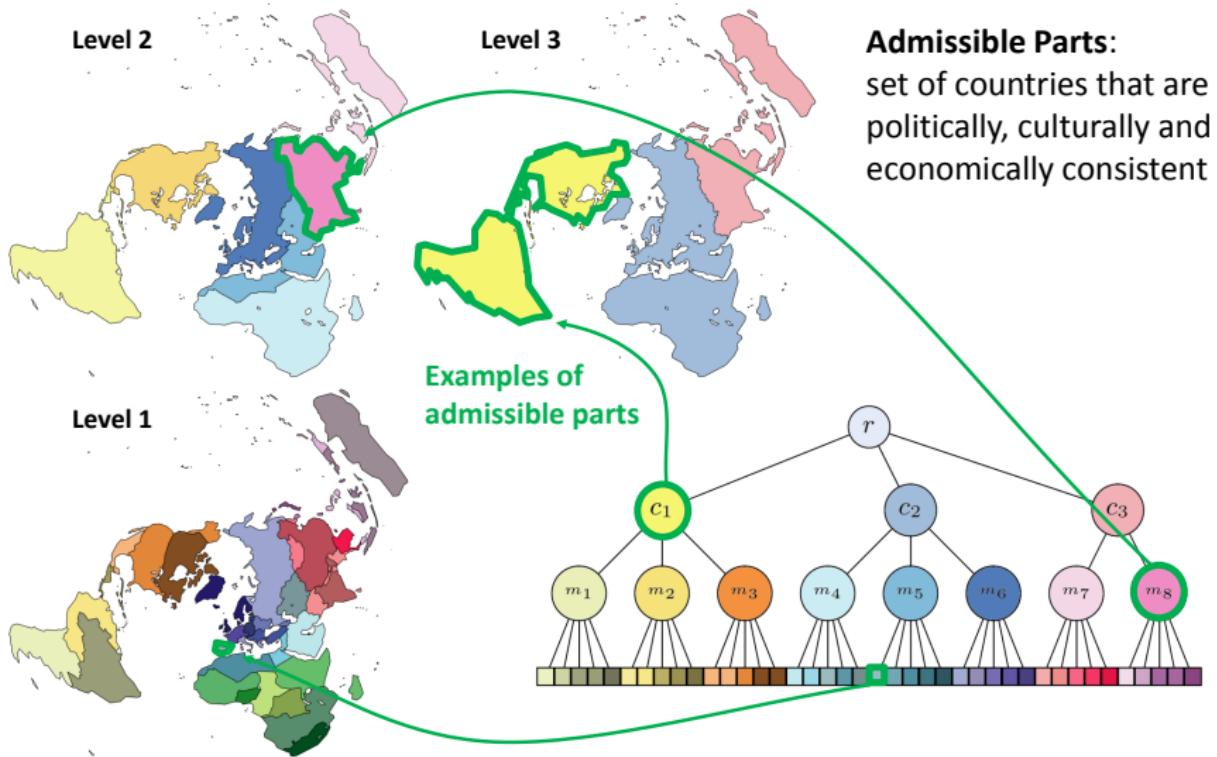
Part II

Formalising semantics
with combinatorial constraints

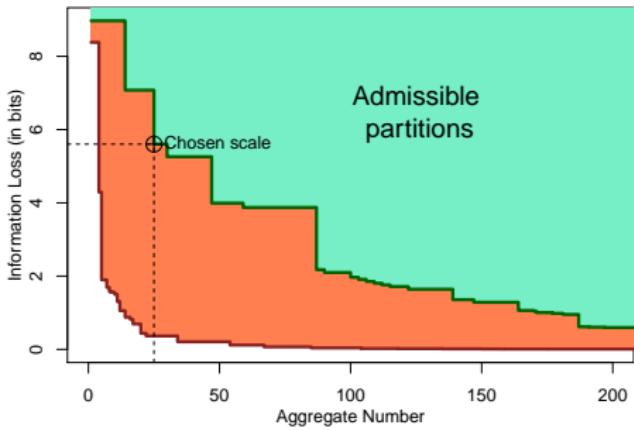
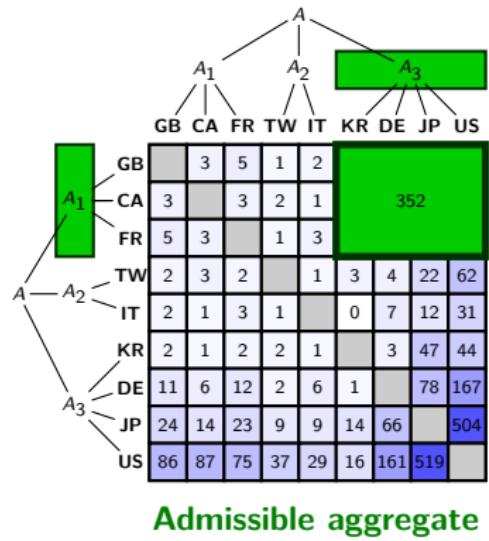
Hierarchical Structures in Geography



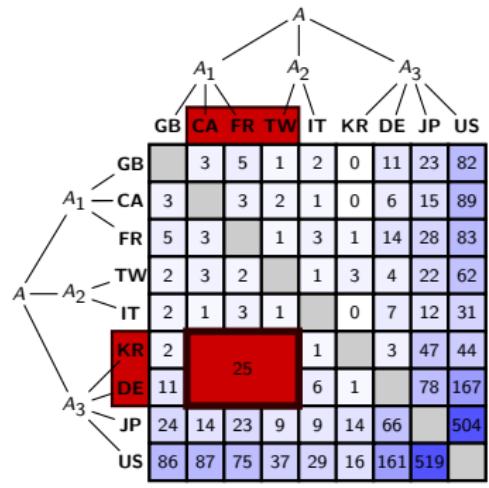
Hierarchical Structures in Geography



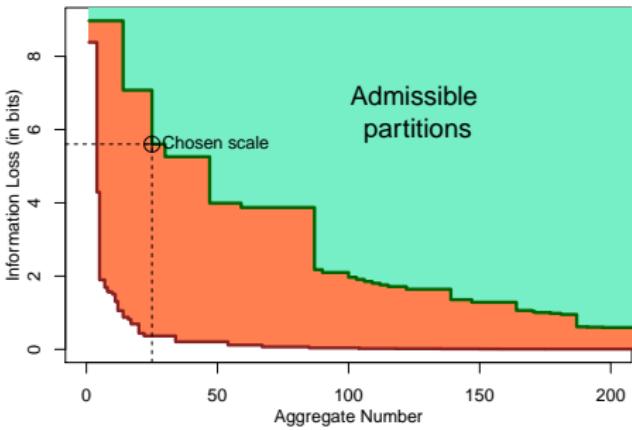
Hierarchical Aggregation of Social Relations



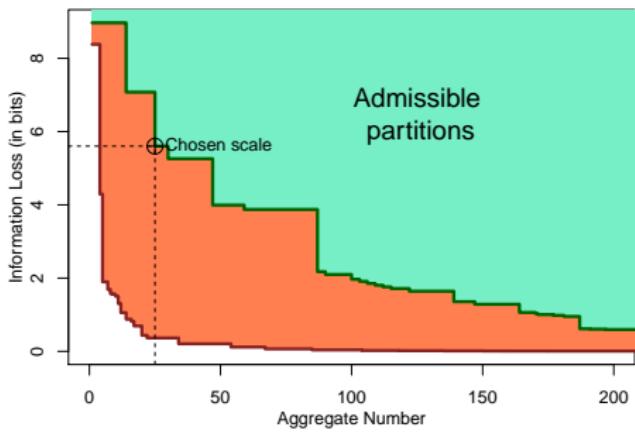
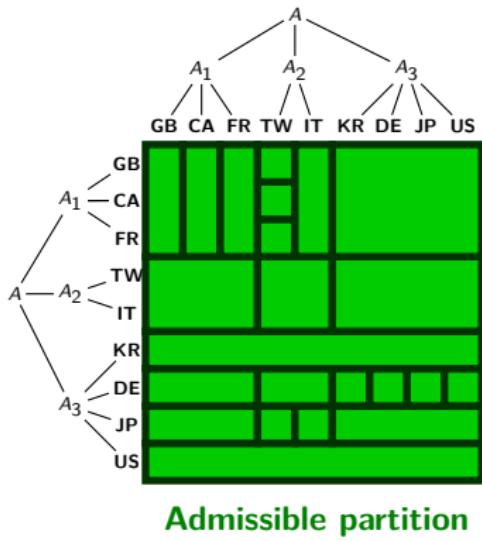
Hierarchical Aggregation of Social Relations



Unauthorised aggregate



Hierarchical Aggregation of Social Relations



Geographic Aggregation of Social Relations

Dataset:

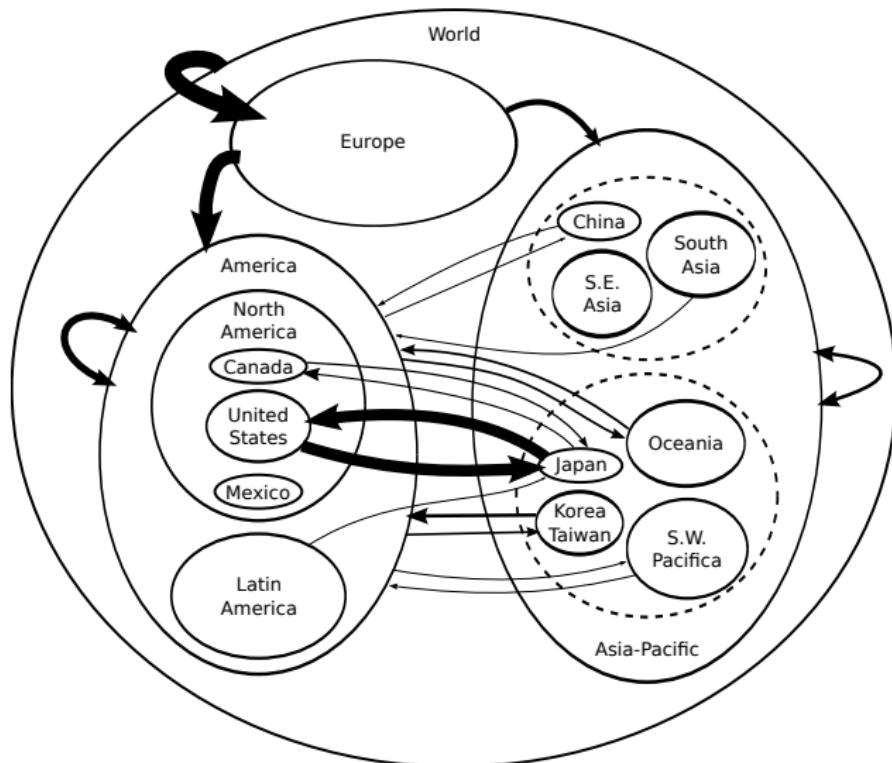
- 1,593,781 patents (1975–1999)
- 2,616,237 citations

Citation matrix:

- 139 countries
- 19,182 possible interactions
- 2,137 observed interactions
- Sparsity: 11.1 %

Hierarchy:

- 5 levels (WUTS)
- 186 admissible 1D-aggregates
- 34,596 admissible 2D-aggregates
- So many admissible partitions...



Disclaimer!

① Epistemic bias:

What do citation networks really tell us about the social practice of patent granting?

② Data bias:

Is our citation dataset in any way representative of the domain?

③ Method bias:

What does the aggregation method really tell us about the data?

④ Algorithmic bias:

Is the aggregation algorithm neutral with respect to the method?

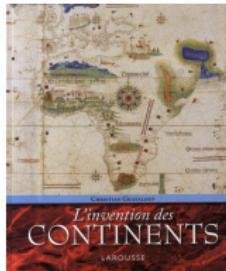
⑤ Semantic bias:

What is the influence of the chosen semantic structure on the results given by the algorithm?

Part III

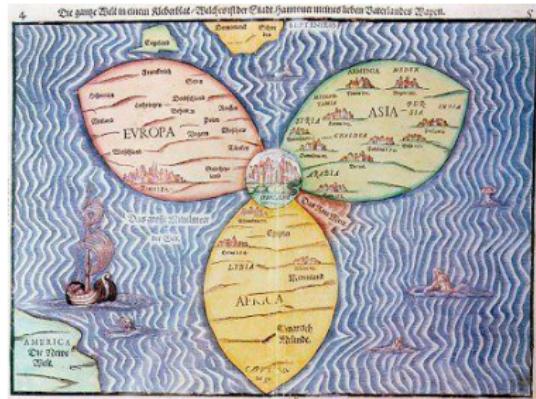
Discussion

Toward new geographical categories

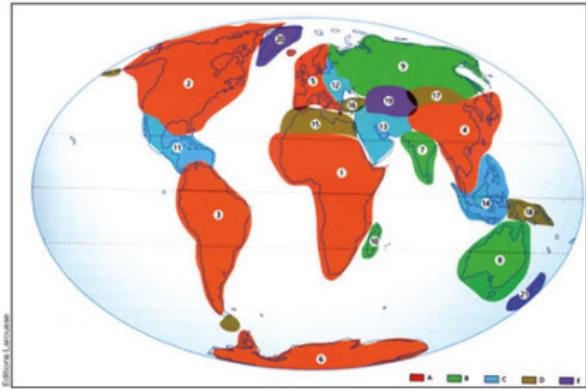


Christian Grataloup. 2009. *L'invention des continents : comment l'Europe a découpé le monde*. Paris, Larousse.

Outdated categories...



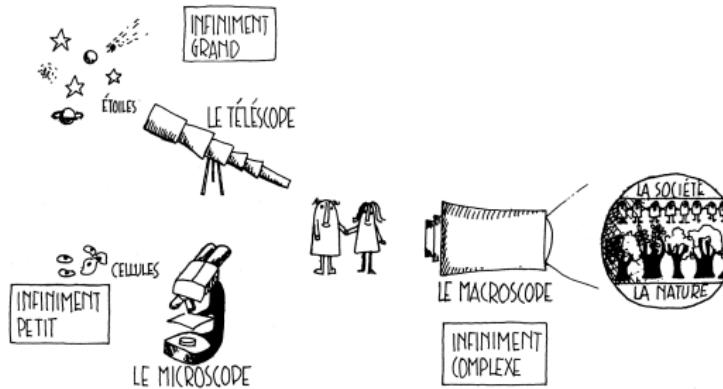
Building new ones?



Thank you for your attention

Mail: Robin.Lamarche-Perrin@lip6.fr

Web: www-complexnetworks.lip6.fr/~lamarche/



« Aujourd'hui, nous sommes confrontés à un autre infini : l'infiniment complexe. Mais cette fois, plus d'instrument. »

Joël de Rosnay. 1975. *Le Macroscope*.