

Multidimensional and Multilevel Analysis of International Media Flows

Robin Lamarche-Perrin, Claude Grasland



**Collaborators from
ANR GEOMEDIA:**

Yves Demazeau

Benjamin Loveluck

Hugues Pecout

Jean-Marc Vincent

What is the state of the world?

International sections of daily newspapers



Which event made the news?

- At a given date?
- Regarding a given country?
- According to a given newspaper?

⇒ **Geomedia agenda-setting**

ANR GEOMEDIA Project: Data and tools for a quantitative analysis of the multiple facets of international news at a worldwide scale

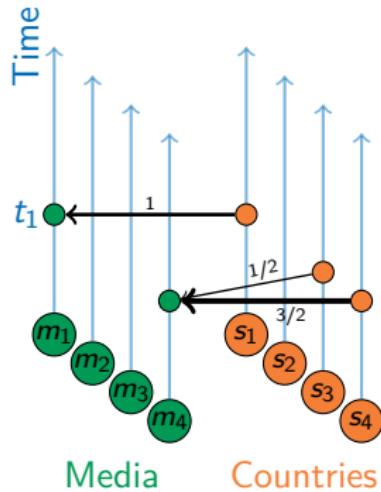
→ Media studies, political sciences, quantitative geography, computer science

The Three Dimensions of Media Flows

International Media Flows



Weighted temporal
bipartite graph

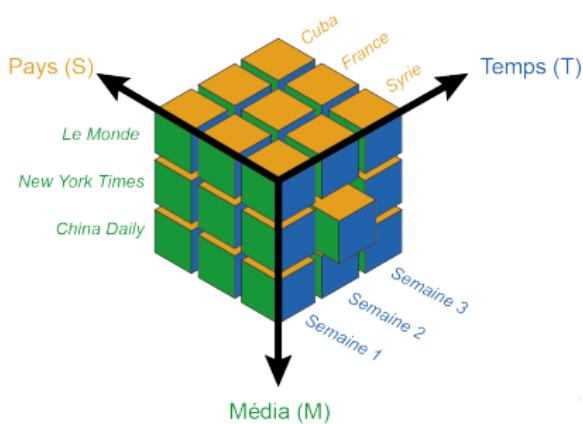


Example:

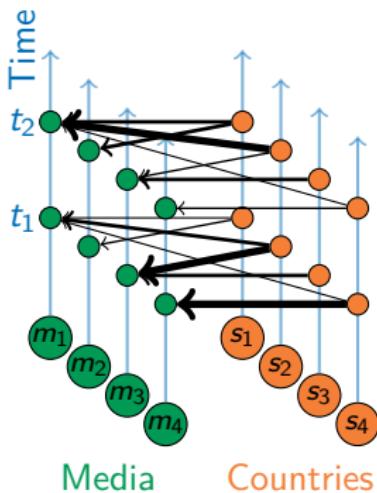
- 1 article published by media m_1 at time t_1 and citing country s_1
- 1 article published by media m_4 at time t_1 and citing country s_4
- 1 article published by media m_4 at time t_1 and citing countries s_3 and s_4

The Three Dimensions of Media Flows

Geomedia Cube
(media × space × time)



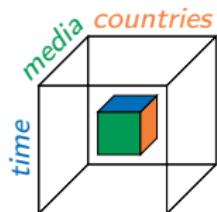
**Weighted temporal
bipartite graph**



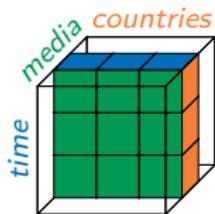
Data: 292 767 articles

published by **36 newspapers** (in 23 different states)
during **52 weeks** (from 28/04/2014 to 26/04/2015)
and citing **197 countries** (recognised by the UN)

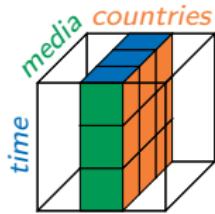
What can be said about one edge?



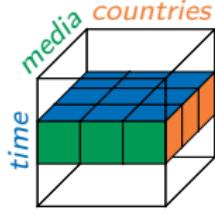
Cuba as been cited **16 times** by Le Monde during the week of December 15th, 2014.
→ Is that a lot?



Knowing that Le Monde made 276 citations **that week**, and that it usually devotes 0.55% of its citations to Cuba?



Knowing that Le Monde made 276 citations **that week**, and that 9.0% of citations of all media **that week** were dedicated to Cuba?



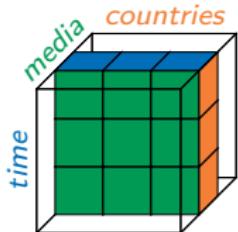
Knowing that Le Monde made 64 citations about Cuba within the whole corpus, and that 18.8% of all citations about Cuba where concentrated on **that particular week**?

Part I

The Multiple Facets of Media Flows

C. Grasland, R. Lamarche-Perrin, B. Loveluck, and H. Pecout. "L'agenda géomédiaque international : analyse multidimensionnelle des flux d'actualité". In *L'Espace Géographique*, vol. 45, issue 2016/1, p. 25-43. Éditions Belin, Paris, 2016.

ISTA Model: Internal Spatio-Temporal Agenda



Knowing that **Le Monde** made 276 citations **that week**, and that it usually devotes 0.55% of its citations to **Cuba**?

→ One would then expect **1.5 citations**, so **16** is a lot!

Given a media $m \in M$, detect spatio-temporal irregularities
 $(s_j, t_k) \in S \times T$.

m	s_1	s_2	s_3
t_1	3	4	3
t_2	3	5	4
t_3	2	7	1

→

m	s_1	s_2	s_3	
t_1				10
t_2				12
t_3				10

8 16 8 32

m	s_1	s_2	s_3
t_1	2.5	5	2.5
t_2	3	6	3
t_3	2.5	5	2.5

Raw values:

$$v(m, s_j, t_k)$$

Marginal values:

$$v(m, \cdot, t_k) = \sum_{s_j \in S} v(m, s_j, t_k)$$

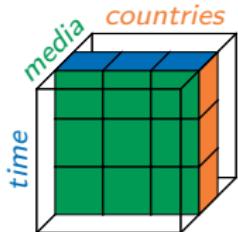
Expected values:

$$v^*(m, s_j, t_k) = \frac{v(m, s_j, \cdot) v(m, \cdot, t_k)}{v(m, \cdot, \cdot)}$$

$$v(m, s_j, \cdot) = \sum_{t_k \in T} v(m, s_j, t_k)$$

$$v(m, \cdot, \cdot) = \sum_{s_j \in S} \sum_{t_k \in T} v(m, s_j, t_k)$$

ISTA Model: Internal Spatio-Temporal Agenda



Knowing that **Le Monde** made 276 citations **that week**, and that it usually devotes 0.55% of its citations to **Cuba**?

→ One would then expect **1.5 citations**, so **16** is a lot!

Given a media $m \in M$, detect spatio-temporal irregularities
 $(s_j, t_k) \in S \times T$.

m	s_1	s_2	s_3
t_1	3	4	3
t_2	3	5	4
t_3	2	7	1

→

m	s_1	s_2	s_3
t_1	+.51	-.12	+.51
t_2	0	-.11	+.63
t_3	-.09	+.73	-.43

←

m	s_1	s_2	s_3
t_1	2.5	5	2.5
t_2	3	6	3
t_3	2.5	5	2.5

Raw values:

$$v(m, s_j, t_k)$$

Significativity:

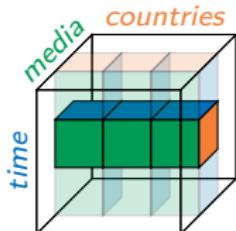
$$p(m, s_j, t_k) = \Pr(X \geq v(m, s_j, t_k)) \quad \text{with } X \sim \text{Pois}(v^*(m, s_j, t_k))$$

Expected values:

$$v^*(m, s_j, t_k) = \frac{v(m, s_j, .) \ v(m, ., t_k)}{v(m, ., .)}$$

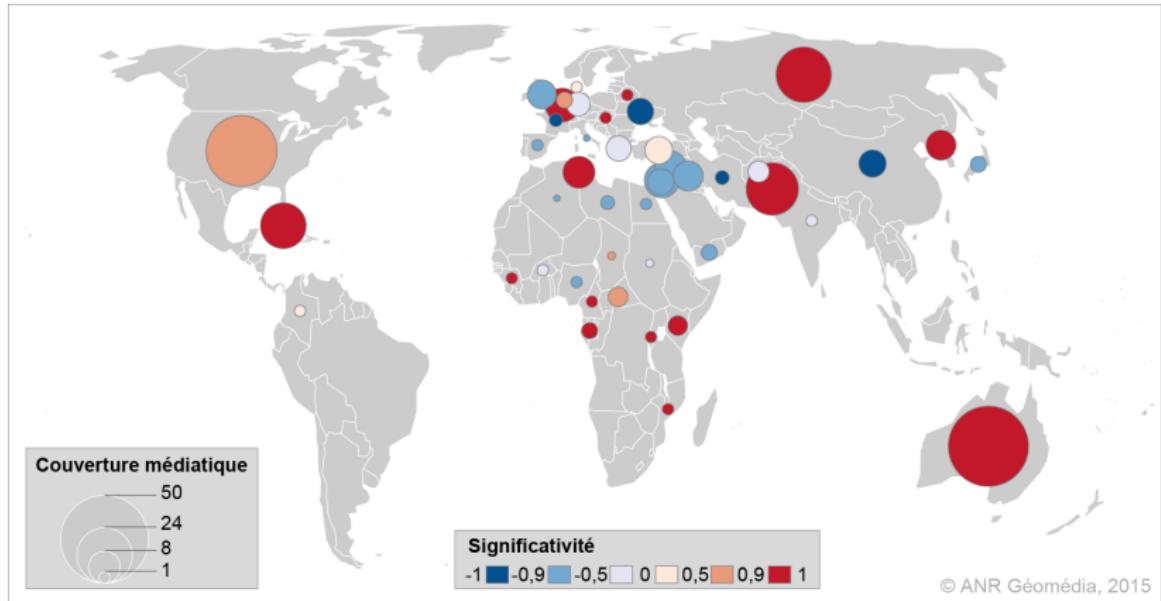
$$\sigma(m, s_j, t_k) = 2p(m, s_j, t_k) - 1 \in [-1, +1]$$

ISTA Model: Internal Spatio-Temporal Agenda

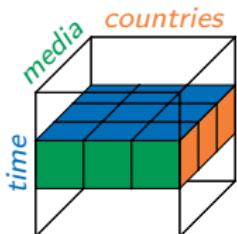


This data model measures significant divergences with respect to the **mean internal agenda** of a particular newspaper

(*Le Monde*, week of 15th December, 2014)



MSA Model: Mean Spatial Agenda



Knowing that **Le Monde** made 276 citations **that week**, and that 9.0% of citations of all media **that week** were dedicated to **Cuba**?

→ One would then expect **25 citations**, so **16** is actually not much...

Given a time $t \in \mathcal{T}$, detect spatio-media irregularities $(m_i, s_j) \in M \times S$.

	t	s_1	s_2	s_3
m_1	3	4	3	
m_2	3	5	4	
m_3	2	7	1	

→

	t	s_1	s_2	s_3	
m_1					10
m_2					12
m_3					10

8 16 8 32

	t	s_1	s_2	s_3
m_1	2.5	5	2.5	
m_2	3	6	3	
m_3	2.5	5	2.5	

Raw values:

$$v(m_i, s_j, t)$$

Marginal values:

$$v(m_i, ., t) = \sum_{m_i \in M} v(m_i, s_j, t)$$

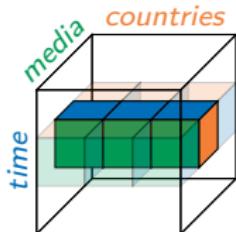
$$v(., s_j, t) = \sum_{s_j \in S} v(m_i, s_j, t)$$

$$v(., ., t) = \sum_{m_i \in M} \sum_{s_j \in S} v(m_i, s_j, t)$$

Expected values:

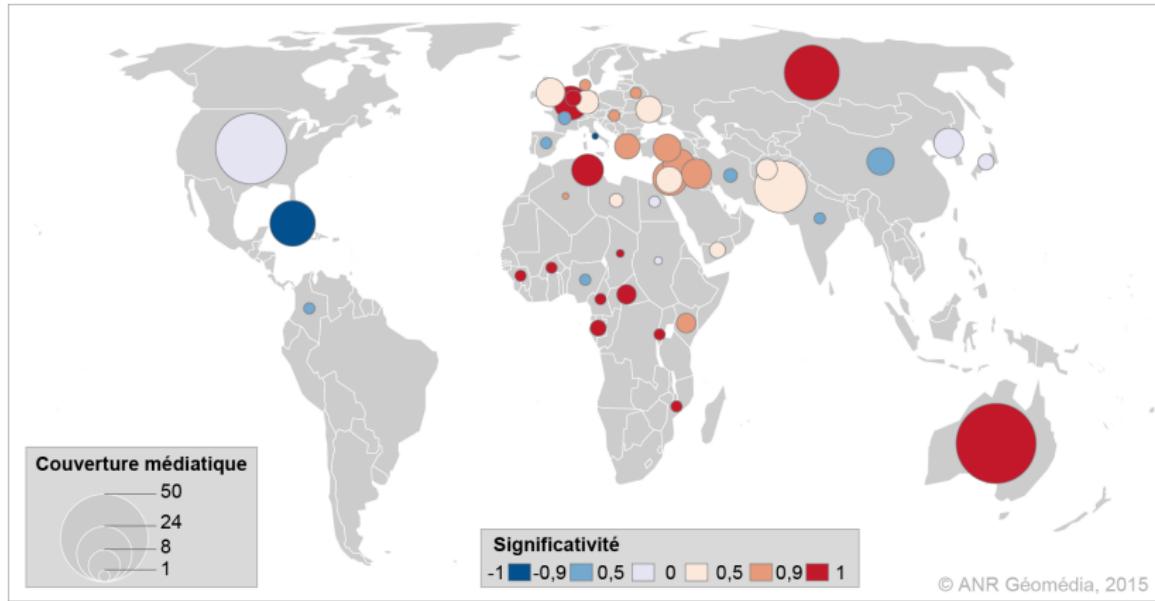
$$v^*(m_i, s_j, t) = \frac{v(m_i, ., t) v(., s_j, t)}{v(., ., t)}$$

MSA Model: Mean Spatial Agenda

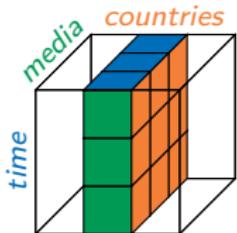


This data model measures significant divergences with respect to the **mean spatial agenda** of all media on a particular time period

(*Le Monde*, week of December 15th, 2014)



MTA Model: Mean Temporal Agenda



Knowing that **Le Monde** made 64 citations about **Cuba** within the whole corpus, and that 18.8% of all citations about **Cuba** were concentrated on **that particular week?**

→ One would then expect **12 citations**, so **16** is a little bit more.

Given a country $s \in S$, detect tempo-media irregularities $(m_i, t_k) \in M \times T$.

s	t_1	t_2	t_3
m_1	3	4	3
m_2	3	5	4
m_3	2	7	1

→

s	t_1	t_2	t_3
m_1			
m_2			
m_3			

8 16 8 32

→

s	t_1	t_2	t_3
m_1	2.5	5	2.5
m_2	3	6	3
m_3	2.5	5	2.5

Raw values:

$$v(m_i, s, t_k)$$

Marginal values:

$$v(m_i, s, \cdot) = \sum_{t_k \in T} v(m_i, s, t_k)$$

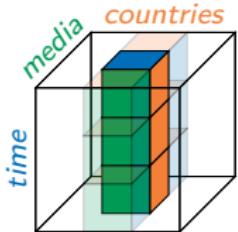
$$v(\cdot, s, t_k) = \sum_{m_i \in M} v(m_i, s, t_k)$$

$$v(\cdot, s, \cdot) = \sum_{m_i \in M} \sum_{t_k \in T} v(m_i, s, t_k)$$

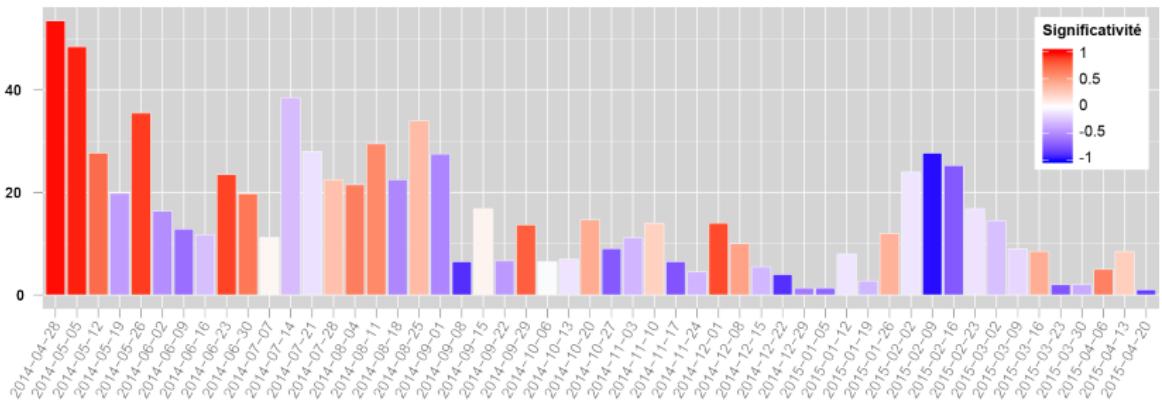
Expected values:

$$v^*(m_i, s, t_k) = \frac{v(m_i, s, \cdot) v(\cdot, s, t_k)}{v(\cdot, s, \cdot)}$$

MTA Model: Mean Temporal Agenda



This data model measures significant divergences with respect to the **mean temporal agenda** of all media regarding a particular country
(Ukraine, Le Monde)

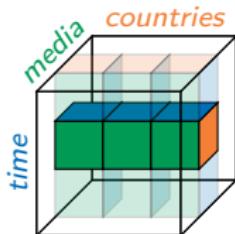


Part II

The Multiple Scales of Media Flows

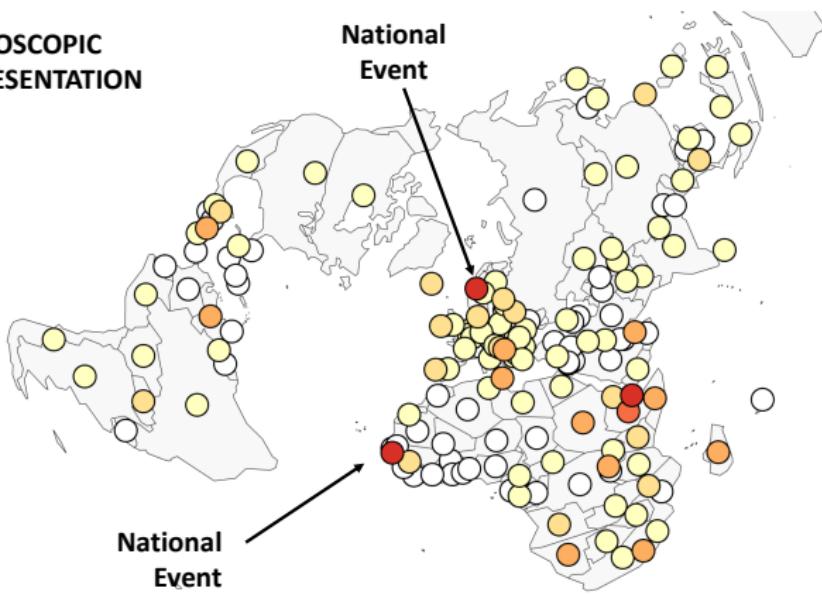
R. Lamarche-Perrin, Y. Demazeau, and J.-M. Vincent. "Building Optimal Macroscopic Representations of Complex Multi-agent Systems". In *Transactions on Computational Collective Intelligence*, vol. XV, LNCS 8670, p. 1-27. Springer-Verlag Berlin, Heidelberg, 2014.

Looking for Geographical Scales

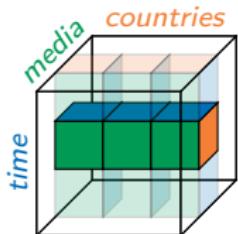


ISTA Model
Le Monde
July 2011

MICROSCOPIC
REPRESENTATION

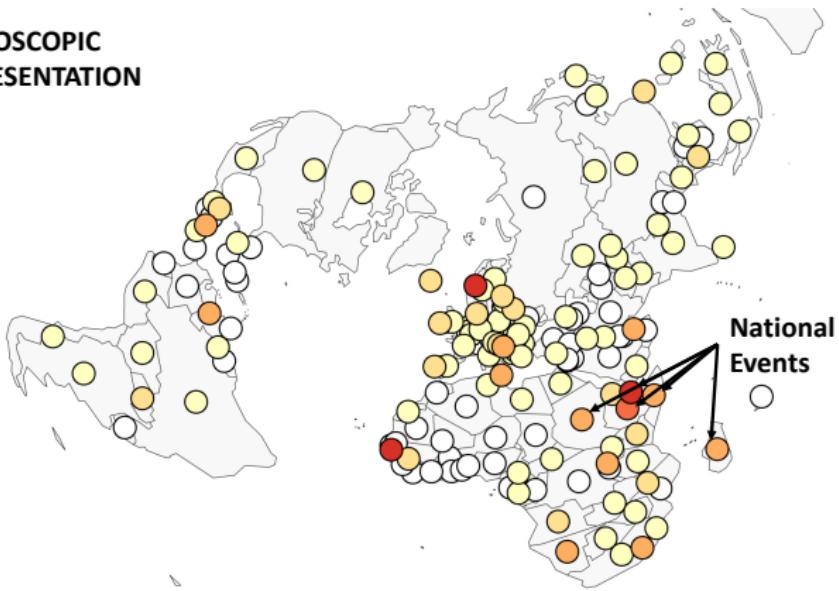


Looking for Geographical Scales



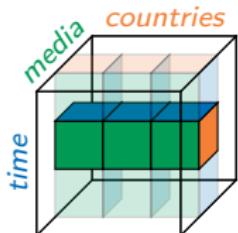
MICROSCOPIC
REPRESENTATION

ISTA Model
Le Monde
July 2011



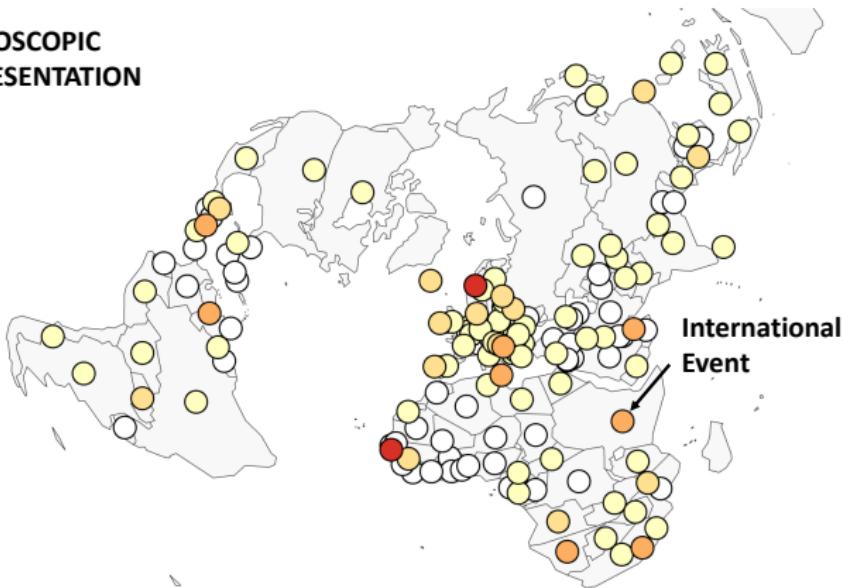
m	s_1	s_2	s_3
t_1	3	3	4
t_2	3	6	3
t_3	2	7	1

Looking for Geographical Scales



MICROSCOPIC
REPRESENTATION

ISTA Model
Le Monde
July 2011

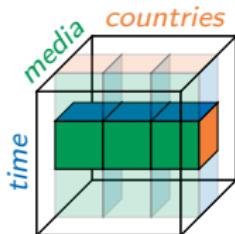


m	s_1	s_2	s_3
t_1	3	3	4
t_2	3	6	3
t_3	2	7	1

Aggregation
→

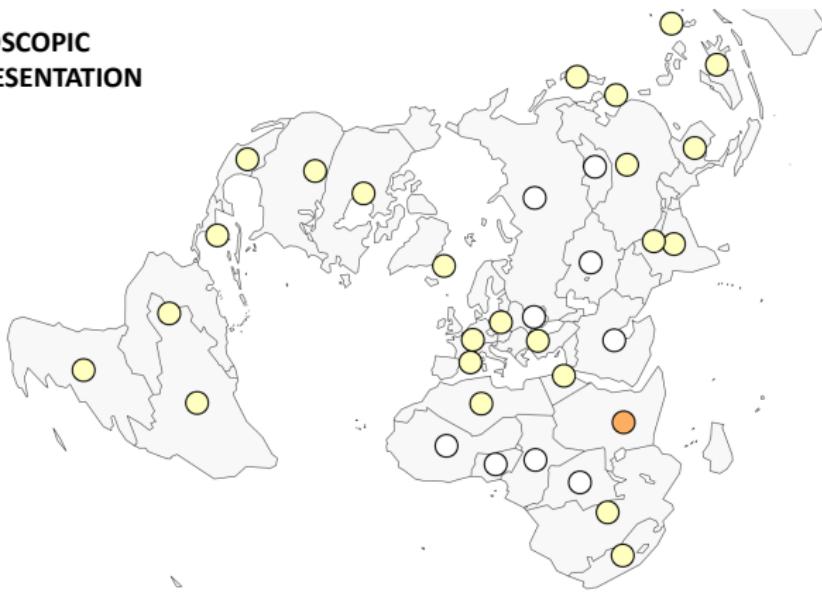
m	s_1	s_2	s_3	
t_1	3	7		10
t_2	3	9		12
t_3	2	8		10
	8	16	8	32

Looking for Geographical Scales



MESOSCOPIC
REPRESENTATION

ISTA Model
Le Monde
July 2011

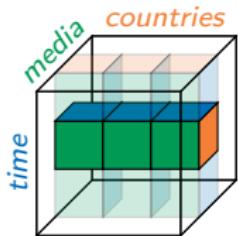


m	s_1	s_2	s_3
t_1	3	3	4
t_2	3	6	3
t_3	2	7	1

Aggregation
→

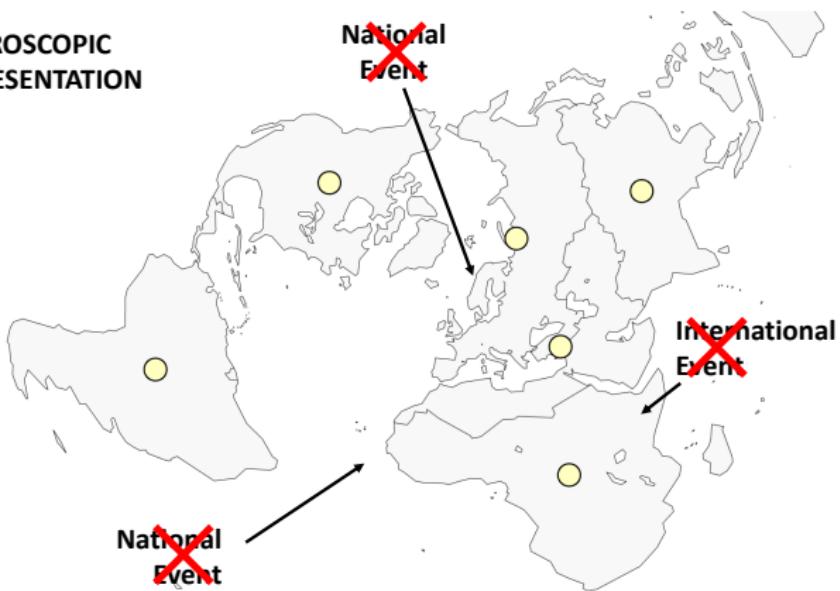
m	s_1	s_2	s_3	
t_1	3	7		10
t_2	3	9		12
t_3	2	8		10
	8	16	8	32

Looking for Geographical Scales



ISTA Model
Le Monde
July 2011

MACROSCOPIC
REPRESENTATION

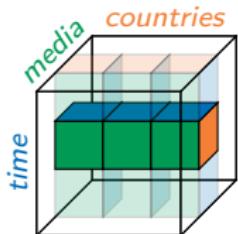


m	s_1	s_2	s_3
t_1	3	3	4
t_2	3	6	3
t_3	2	7	1

Aggregation
→

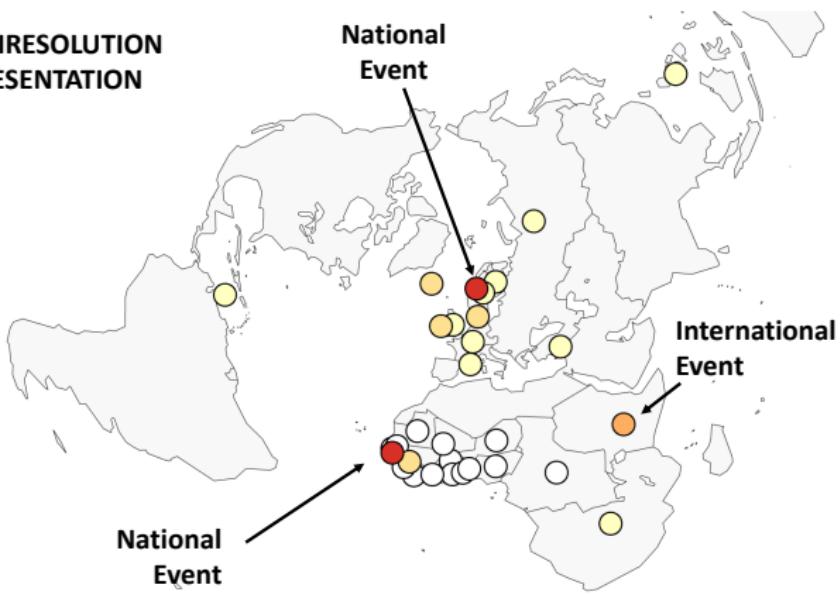
m	s_1	s_2	s_3	
t_1	3	7		10
t_2	3	9		12
t_3	2	8		10
	8	16	8	32

Looking for Geographical Scales



MULTIRESOLUTION
REPRESENTATION

ISTA Model
Le Monde
July 2011



m	s_1	s_2	s_3
t_1	3	3	4
t_2	3	6	3
t_3	2	7	1

Aggregation
→

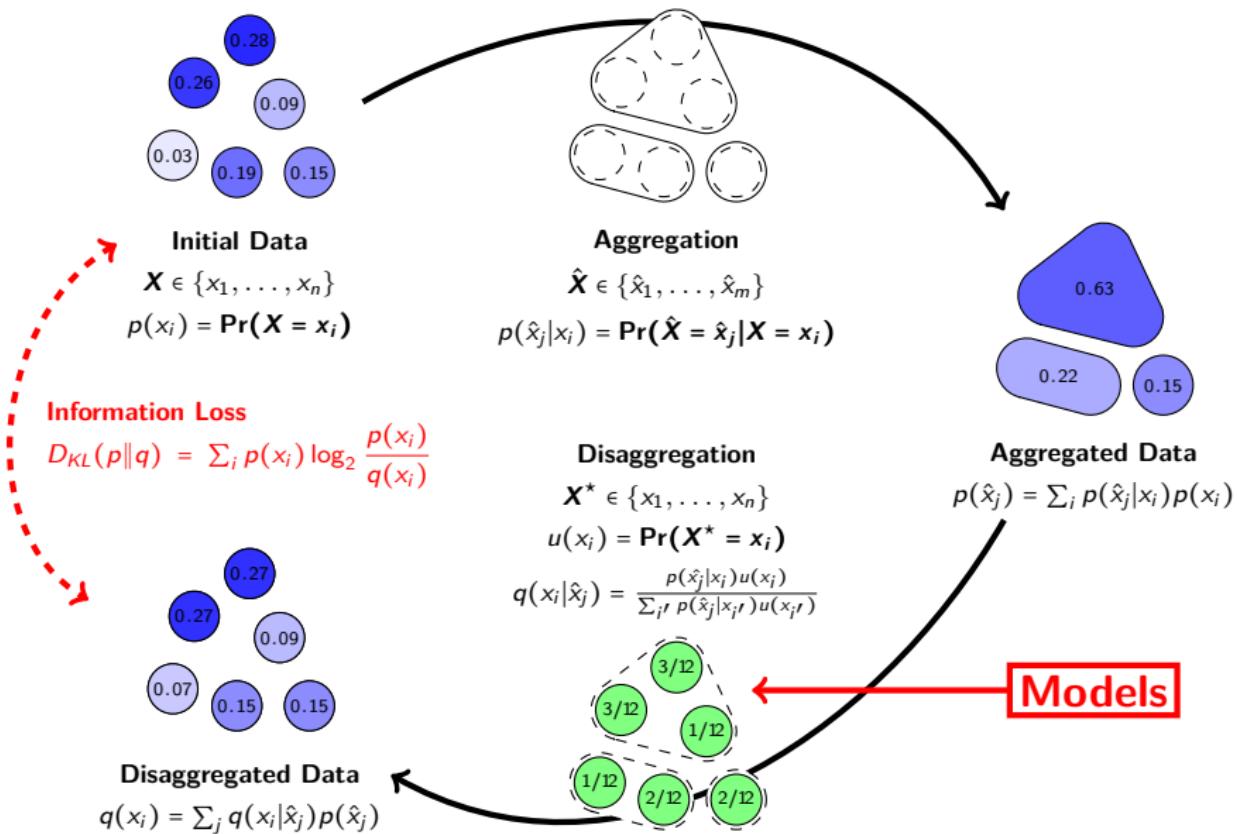
m	s_1	s_2	s_3
t_1	3	7	
t_2	3	9	
t_3	2	8	

8 16 8 32

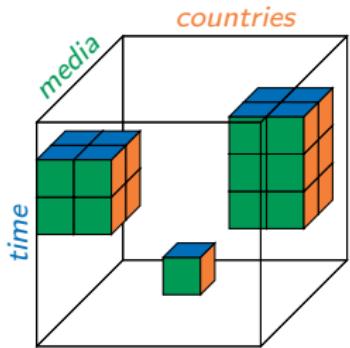
Disaggregation
→

m	s_1	s_2	s_3
t_1	3	4.7	2.3
t_2	3	6	3
t_3	2	5.3	2.7

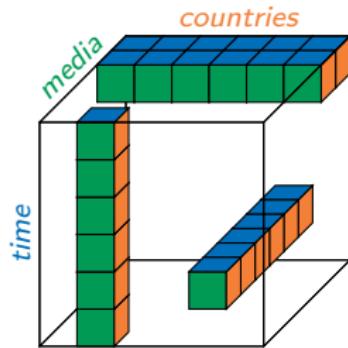
Data Aggregation and Information Loss



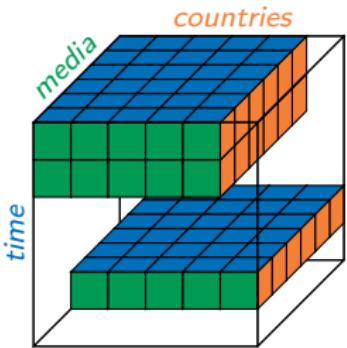
Multidimensional Aggregation in the Cube



No privileged dimension

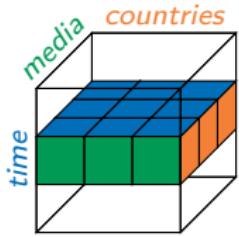


One privileged dimension

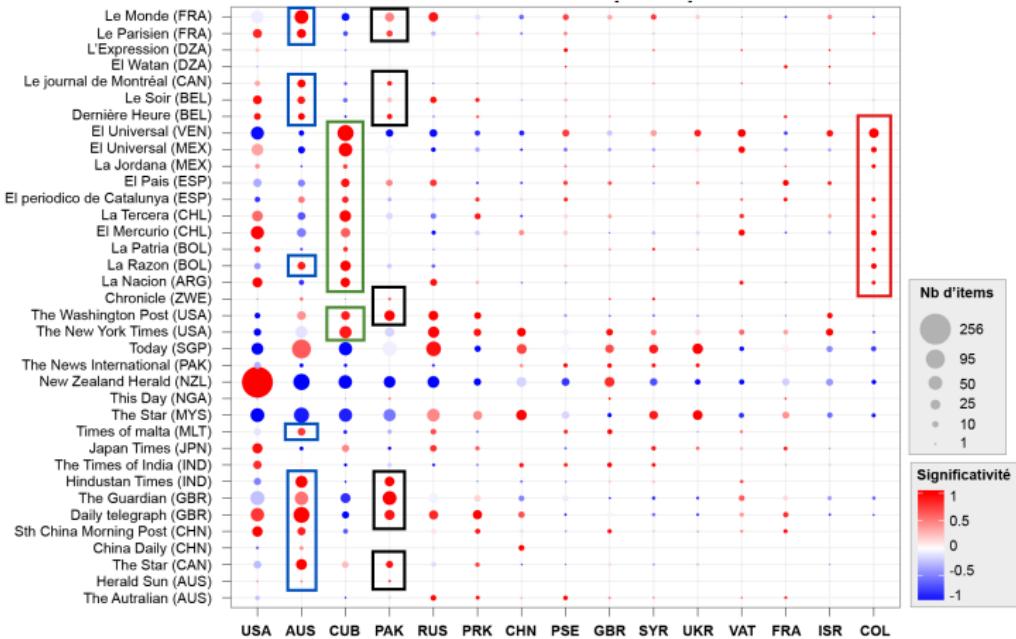


Two privileged dimensions

Media Aggregation

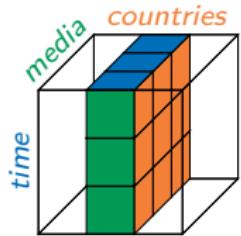


MSA Model
Week of
15/12/2014

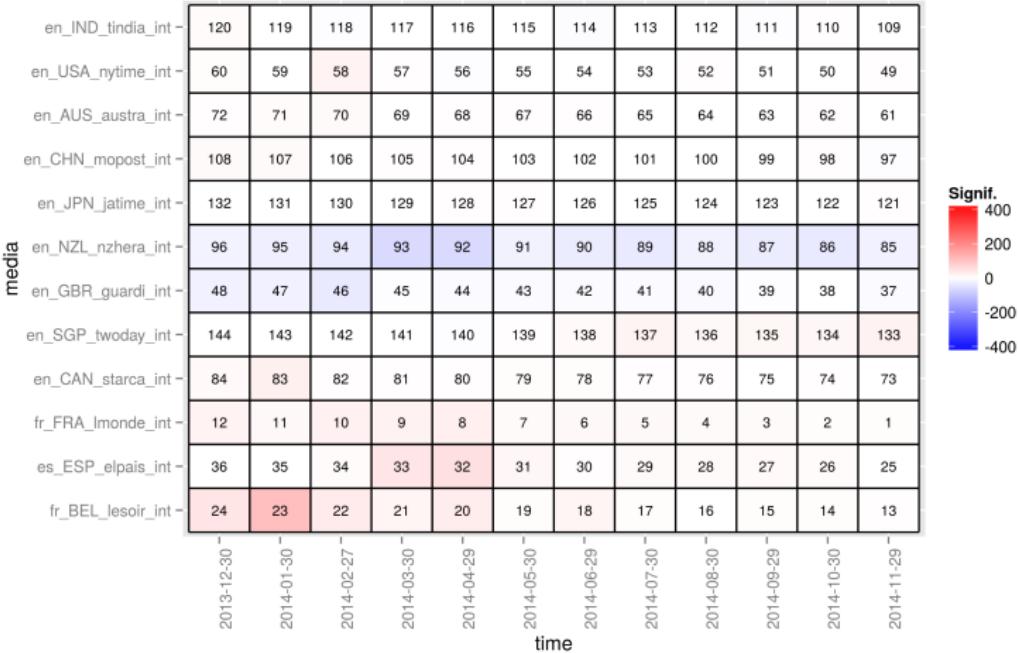


Media × Time Aggregation

No information loss → 144 aggregates

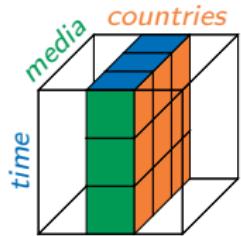


MSA Model
Ukraine

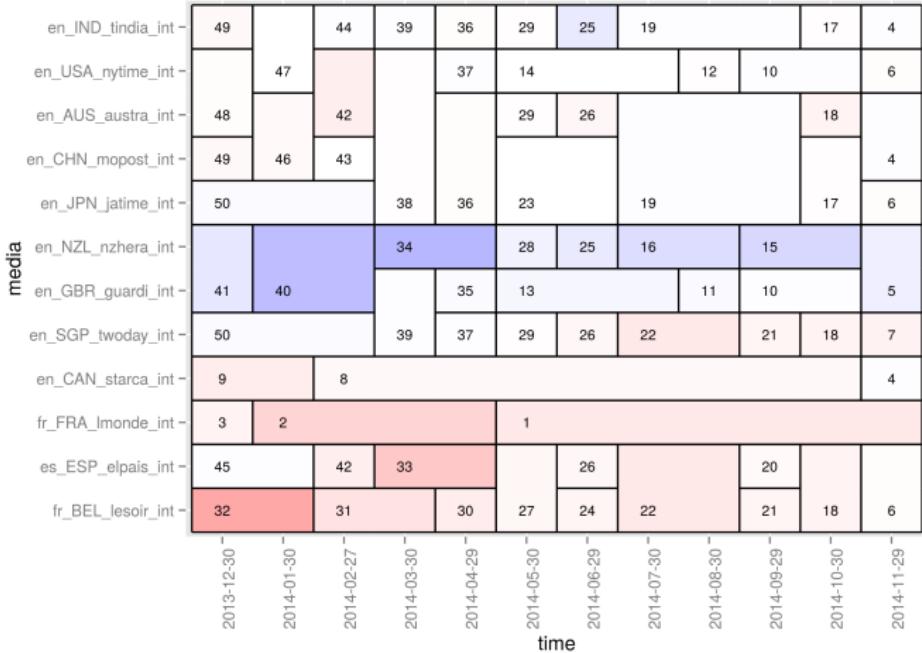


Media × Time Aggregation

0.7% of information loss → 50 aggregates



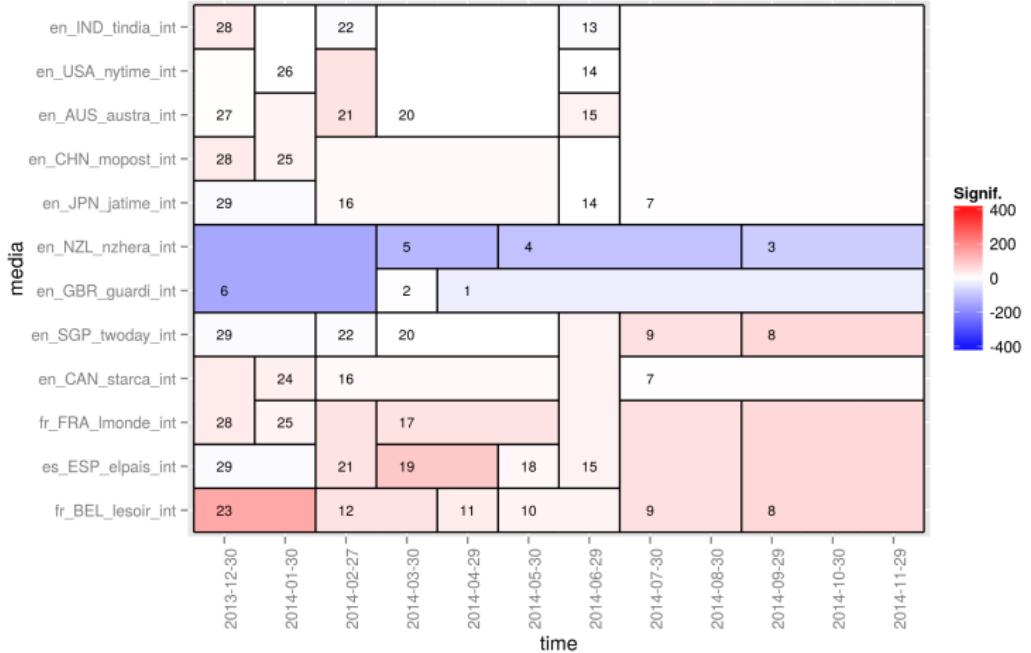
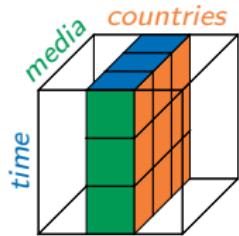
MSA Model
Ukraine



Media × Time Aggregation

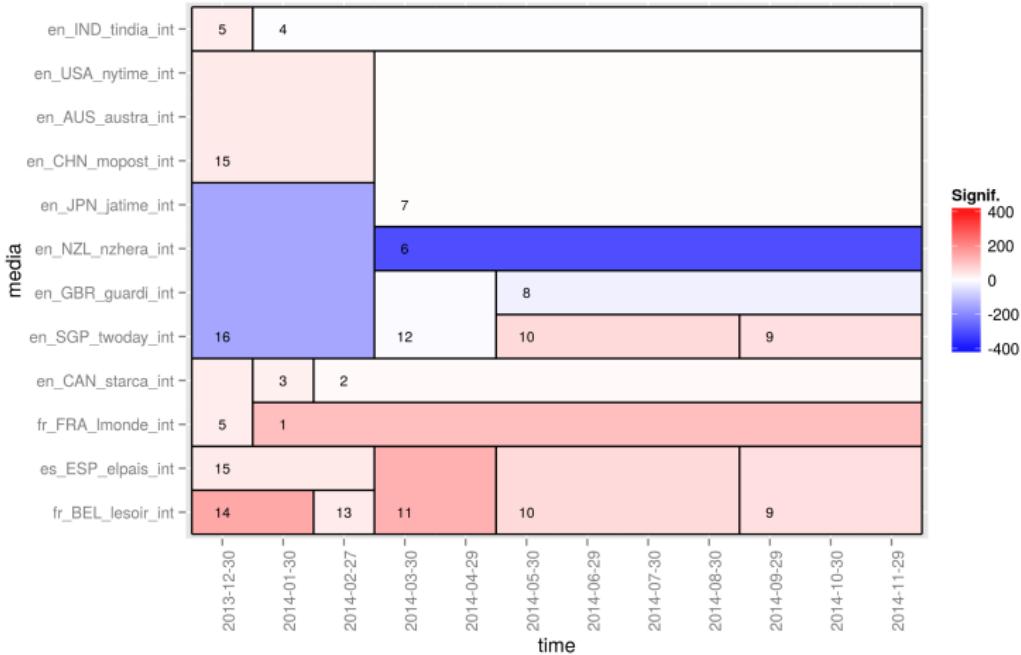
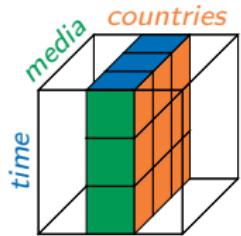
2.9% of information loss

→ 29 aggregates



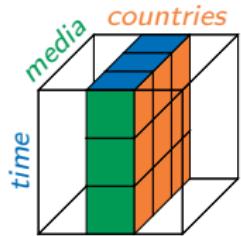
Media × Time Aggregation

5.9% of information loss → 16 aggregates

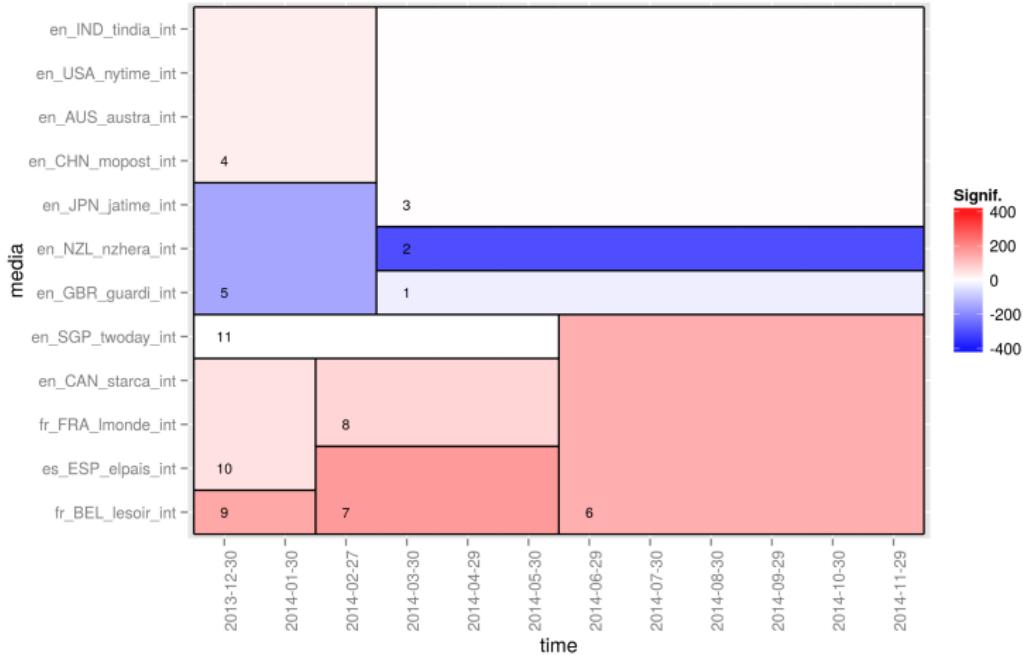


Media × Time Aggregation

8.5% of information loss → 11 aggregates

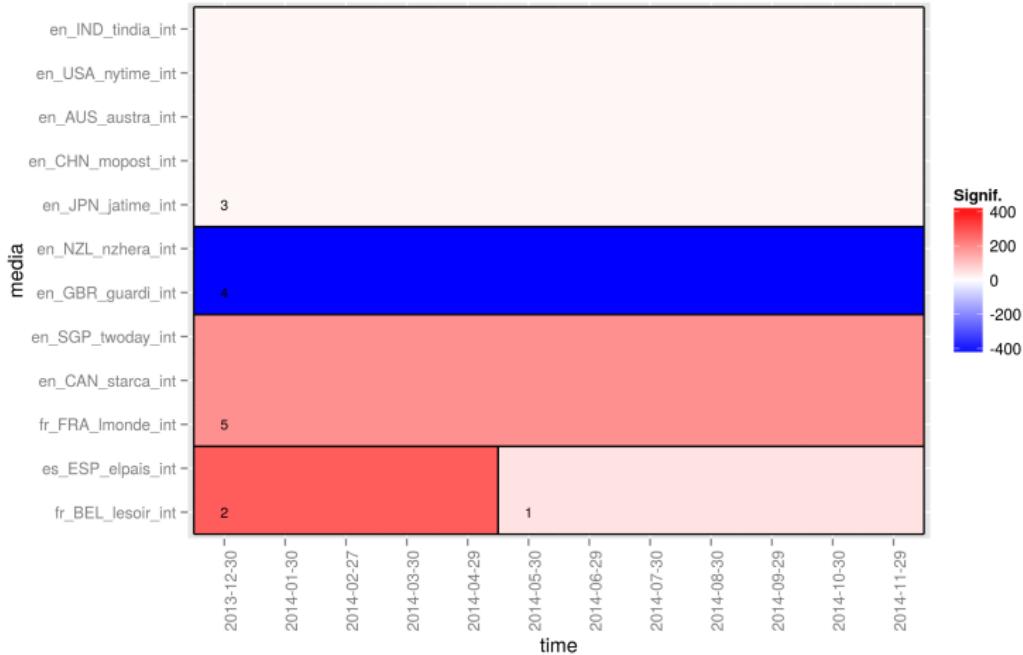
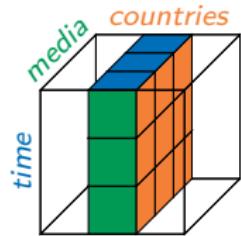


MSA Model
Ukraine



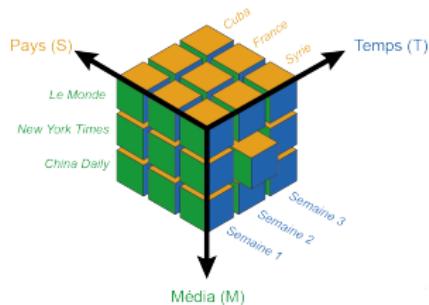
Media × Time Aggregation

18.3% of information loss → 5 aggregates



Conclusion

Geomedia Cube



Limitations and Perspectives

- Events are defined as outliers **with respect to the selected corpus**
- Aggregation should also take into account the semantics of the three dimensions (**rdv tomorrow at 9:50 in Session 5**)
- Measure of statistical significance could be improved (now using Poisson counting processes)
- We need heuristics (sub-optimal algorithms) to work on the whole cube

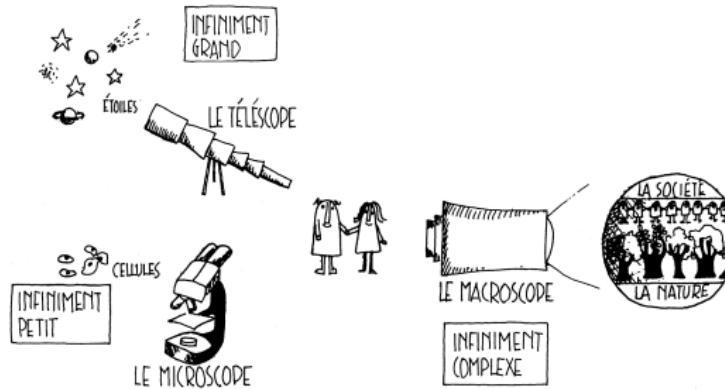
Ressources

- Aggregation algorithms: github.com/Lamarche-Perrin/optimal_partition
- GEOMEDIA Project: geomedia.hypotheses.org
- GEOMEDIA Data: [www.gis-cist.fr/en/
mise-a-disposition-dun-echantillon-de-la-base-de-donnees-geomedia](http://www.gis-cist.fr/en/mise-a-disposition-dun-echantillon-de-la-base-de-donnees-geomedia)
- GEOMEDIA Workshop: [www.gis-cist.fr/en/events/
english-geomedia-structure-and-dynamics-of-media-flows](http://www.gis-cist.fr/en/events/english-geomedia-structure-and-dynamics-of-media-flows)

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*.