

# Le calcul intensif en géographie : une tradition bien ancrée

*Le cas de la simulation*

Sébastien Rey-Coyrehourcq (**IGR UMR IDEES**) \*

Arnaud Banos (**DR UMR IDEES**)

Juste Raimbault (**Post Doc CASA**)



@reyman64



<https://slides.com/sebastienreycoyrehourcq/jcad2019>

ideas

9 oct. 2019 - Toulouse  
JCAD 2019

# Aperçu des pratiques historiques ...

*simulation & centre de calculs*



chapitre de thèse [Rey-Coyrehourcq 2015]

stage 4 mois 2019 : "*Pérénnisation & valorisation code sources des pionniers de la quantitative en géographie*"

# Simulation en géographie

[Varenne2017, Rey-Coyrehourcq 2015]

## Les pionniers Américains et Suédois

1950's - 1960's

Univ. Washington  
(Marble, Pitts, Tobler, etc.)

transfert (1957)

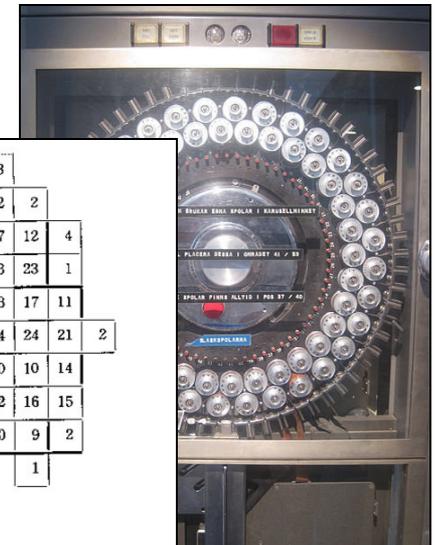
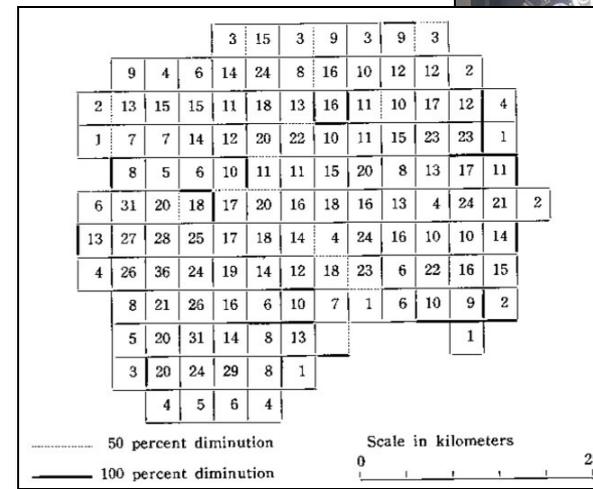


T. Hagerstrand

1956 / SMIL at Lund



Washington University **IBM 650**  
& Northwestern's  
Vogelback **CDC 3400/ 6400**



*Simulation de diffusion par Monte-Carlo*

[ex Hager III and Hager IV Fortran program / 1967]

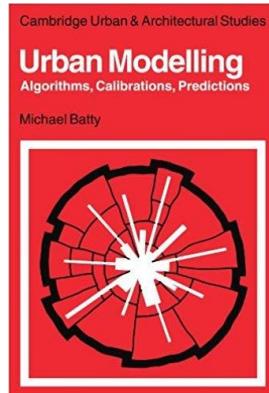
# Simulation en géographie

[Rey-Coyrehourcq 2015]

Les pionniers au Royaume-Uni  
1970's & 1980's

*HPC and the Art of Parallel  
Programming - 1999*

## Urban Modelling



M. Batty 1976

H. Couclelis  
M. Batty  
A. Wilson

...

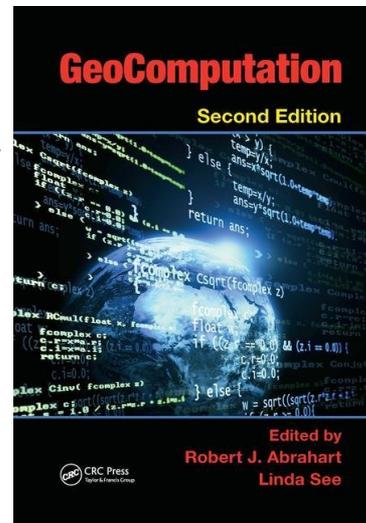


*pratiques  
peu documentées*

## Equipe de Leeds

S. Openshaw  
R. Abraham  
I. Turton

...



*HPC & Geocomputation - 2000*

High-performance  
Computing and the Art of  
Parallel Programming  
An introduction for geographers, social  
scientists and engineers

Stan Openshaw and Ian Turton

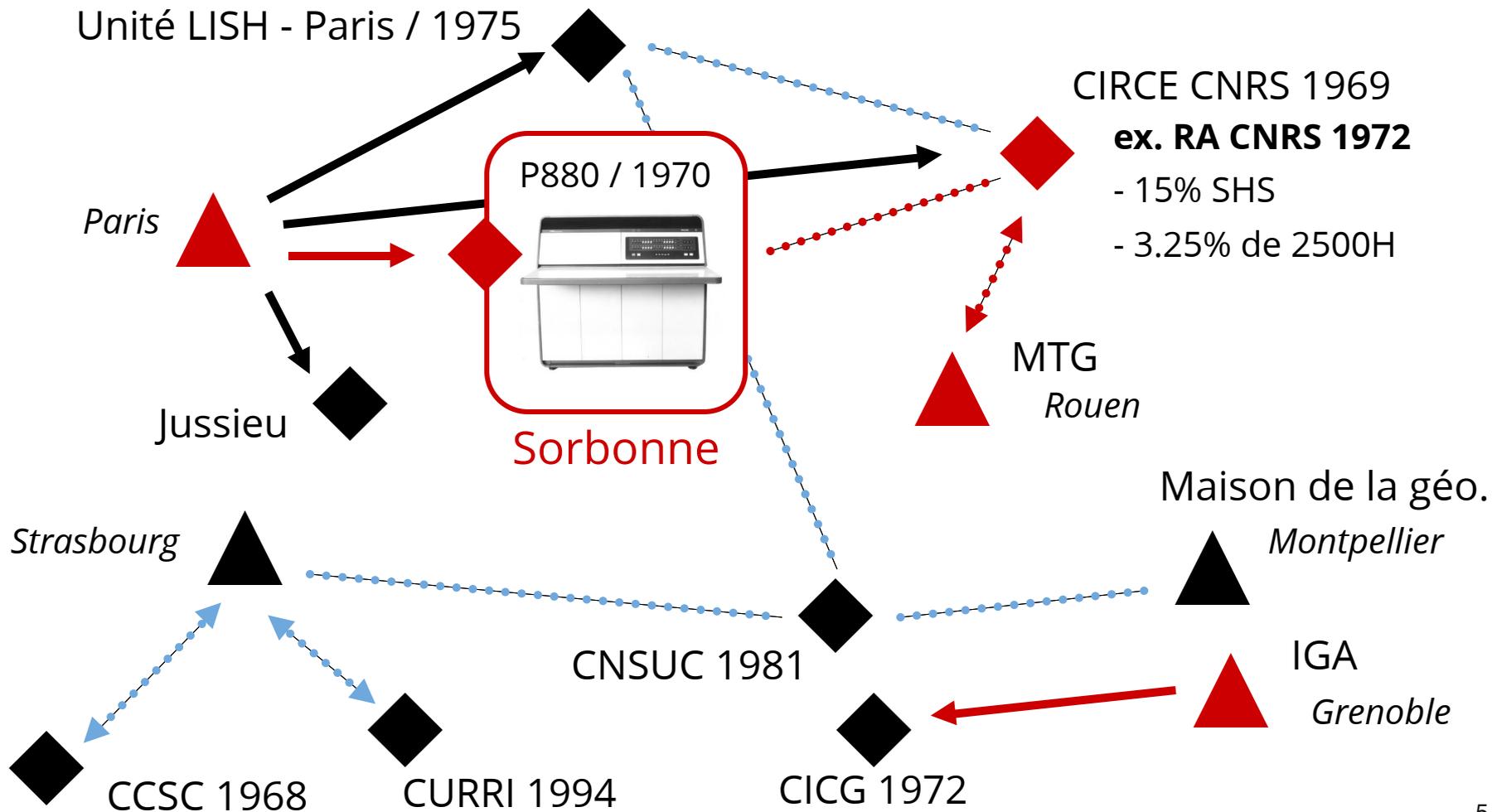
*// **GeoComputation** is [...] meant to imply the adoption of a large-scale computationally intensive scientific paradigm as a tool for doing all manner of geographical research."*

# Simulation en géographie

[Rey-Coyrehourcq 2015]

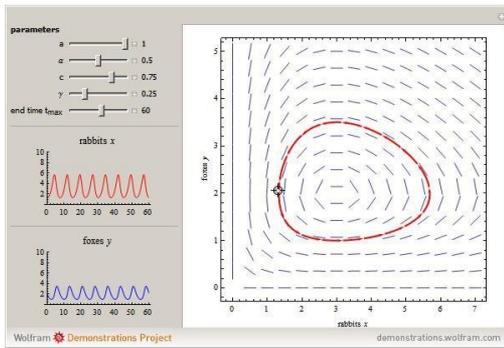
Les pionniers en France  
1970's - 1990's

- ◆ Calculateurs
- ▲ Géographes (calculs & **simulation**)
- Connections
- ➔ Déplacements



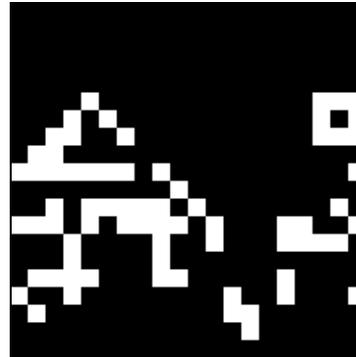


# Modèles dynamiques



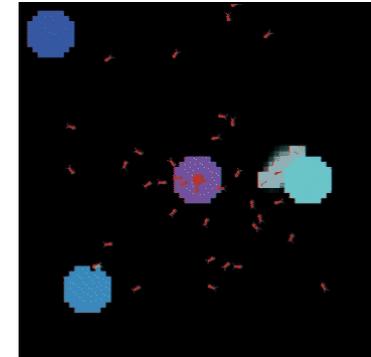
proie prédateur

Systemes dynamiques



jeu de la vie

Automates  
Cellulaires



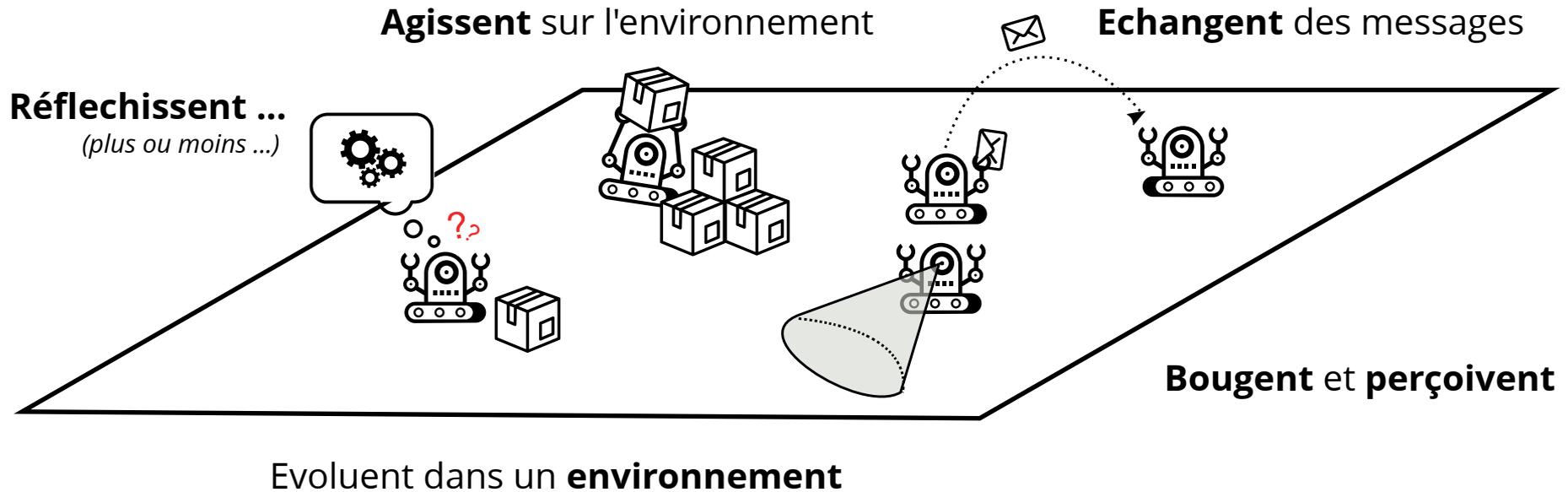
fourmis

Modèles multi-agent

***Couplage possible de ces trois approches***

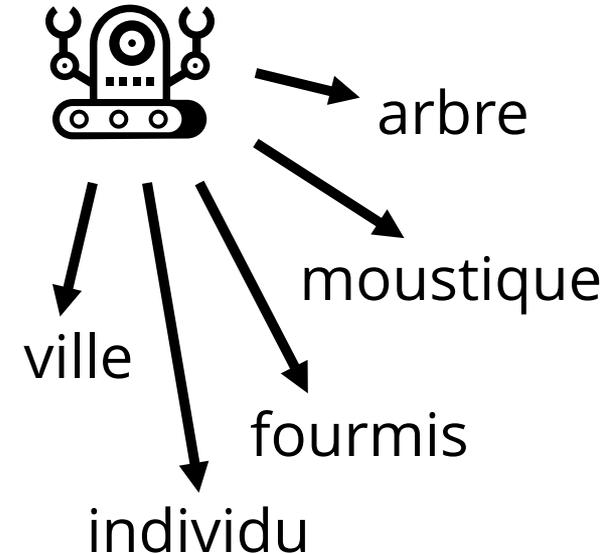
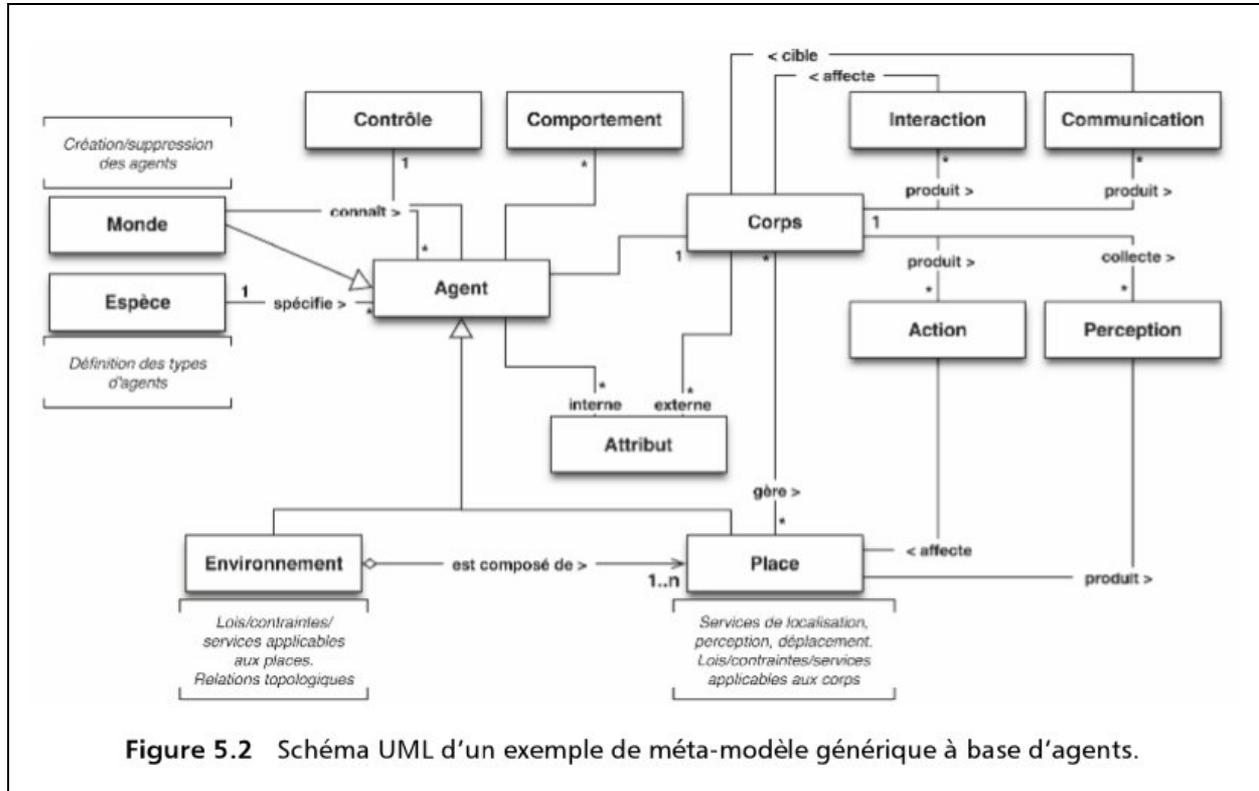
# Modèle multi-agents

[Ferber 1999, Drogoul 2009]



# Modèle multi-agents

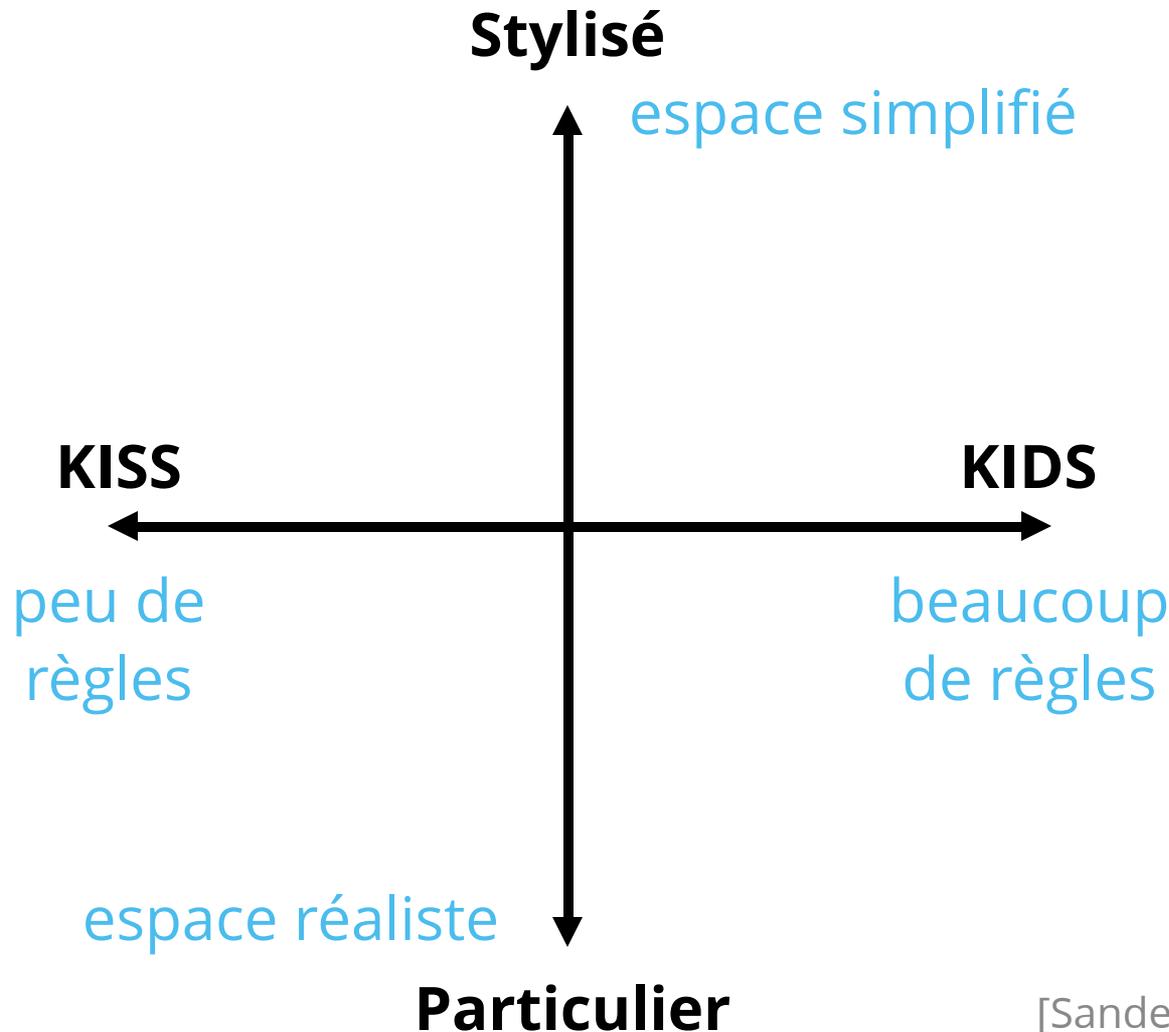
Un **méta-modèle** plus qu'un modèle



[Drogoul Treuil 2008]

flexibilité de représentation &  
emboîtement d'**échelle** !

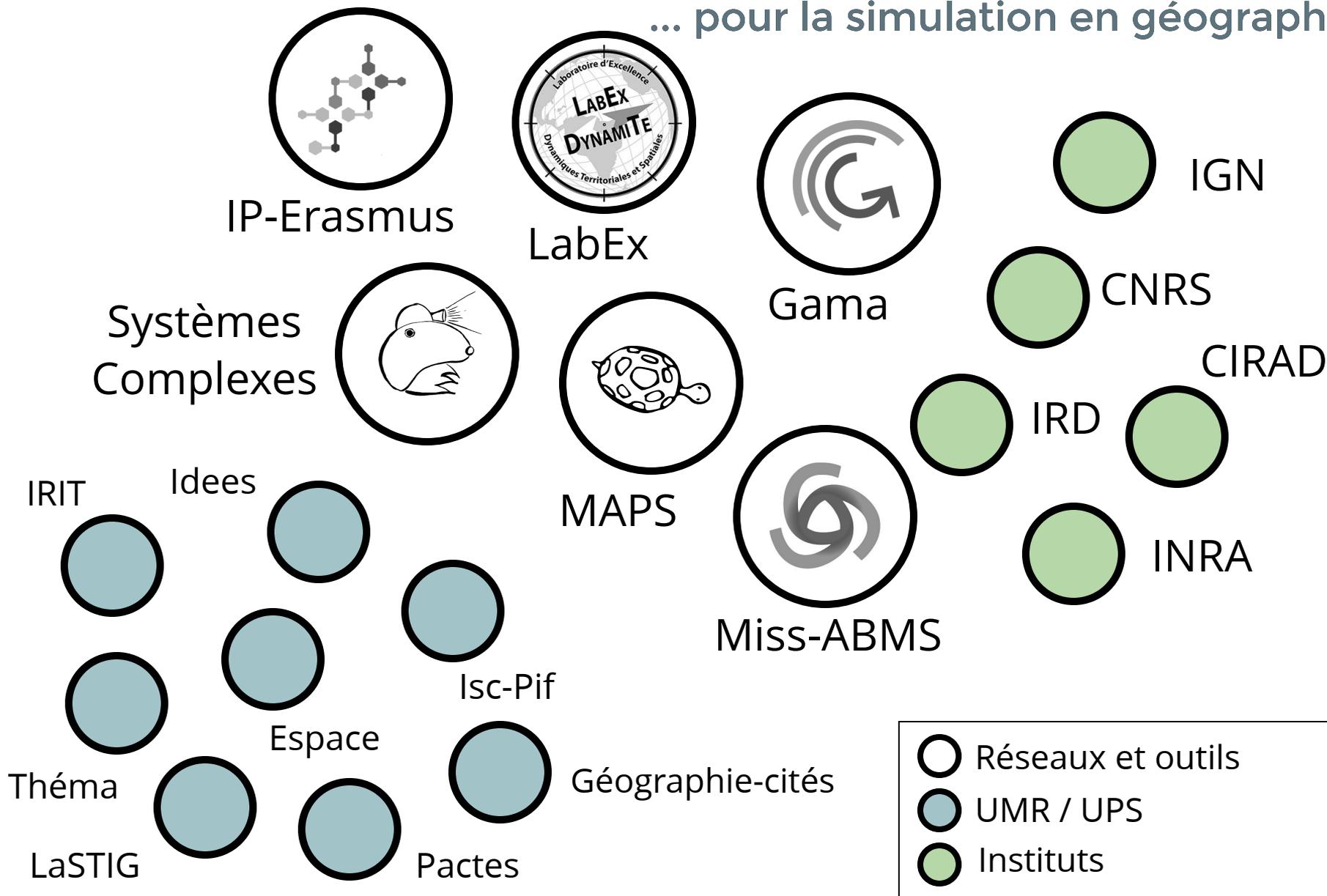
# Une variété de pratiques



[Sanders & Banos 2011,  
2013, Banos 2013]

# Une diversité d'acteurs ...

... pour la simulation en géographie



**Besoins en "calcul" ?  
Quelles solutions pour la simulation**

# Les *racines* du mal

Les quatre cavaliers de l'apocalypse

... **curse dimensionality** ...

Une exploration directe est impossible

... **non linéarité** ...

diversité de formes et d'évolutions

... **stochasticité** ...

un résultat n'est jamais sûr

... **équifinalité** ...

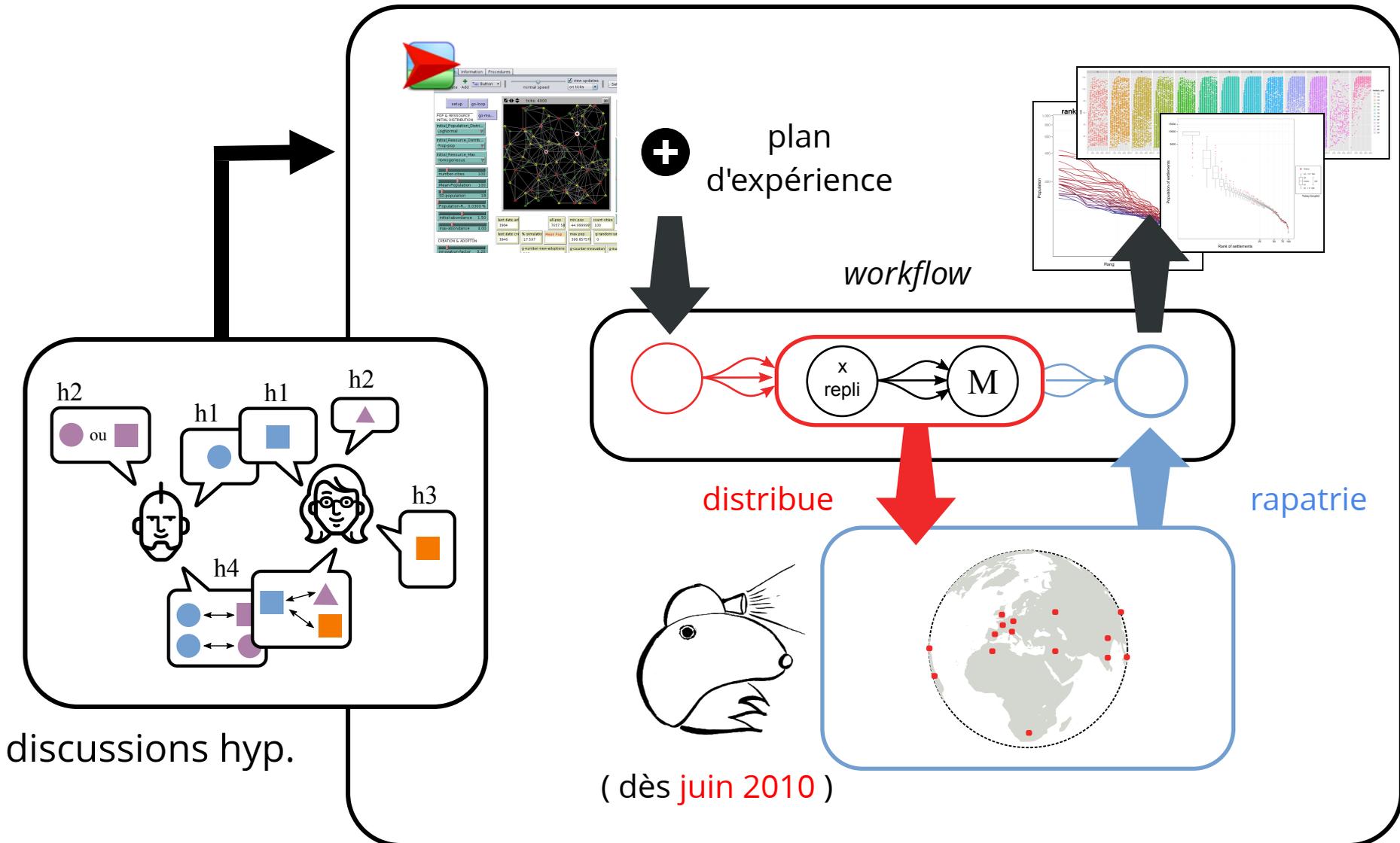
un état final est atteignable de  
1000 façon différentes  
par 1000 personnes différentes

→ **Important** en shs, [pumain 2005 ]  
"cumulativité des connaissances"



# Accès aux ressources

... de nombreux aller retours entre modèles et résultats ...





# OpenMOLE

[next.openmole.org](http://next.openmole.org)

[iscipif.fr](http://iscipif.fr)

Une plateforme et une communauté **en expansion**

Créer des "chaînes de traitements"  
**parallélisables, reproductibles, simples** (ou complexes)  
sans **aucune connaissance** des technologies **HPC**

```
val i = Val[Double]  
val res = Val[Double]
```

```
val exploration =  
ExplorationTask(i in (0.0 to 100.0 by 1.0))
```

```
val model =  
  ScalaTask("val res = i * 2") set (  
    inputs += i,  
    outputs += (i, res)  
  )
```

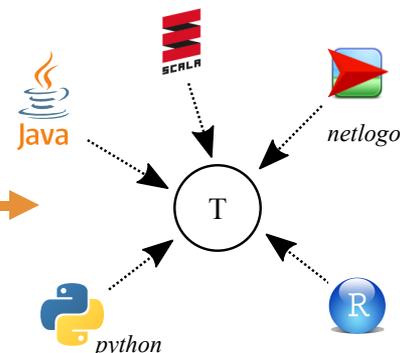
```
val env = LocalEnvironment(4)
```

```
exploration -< (model on env hook ToStringHook())
```

Langage de scripts (DSL) **simple**  
**et flexible** pour composer les  
workflow

Plusieurs types de plan  
d'expériences : **optimisation &**  
**exploration.**

Ecosystème riche:  
**simulation, etc.**



**Support de multiples**  
**technologies HPC & Grilles**

# Exemples SMA

plateformes  
SMA  
ou java/scala



GAMA (gaml)

<https://gama-platform.github.io/>

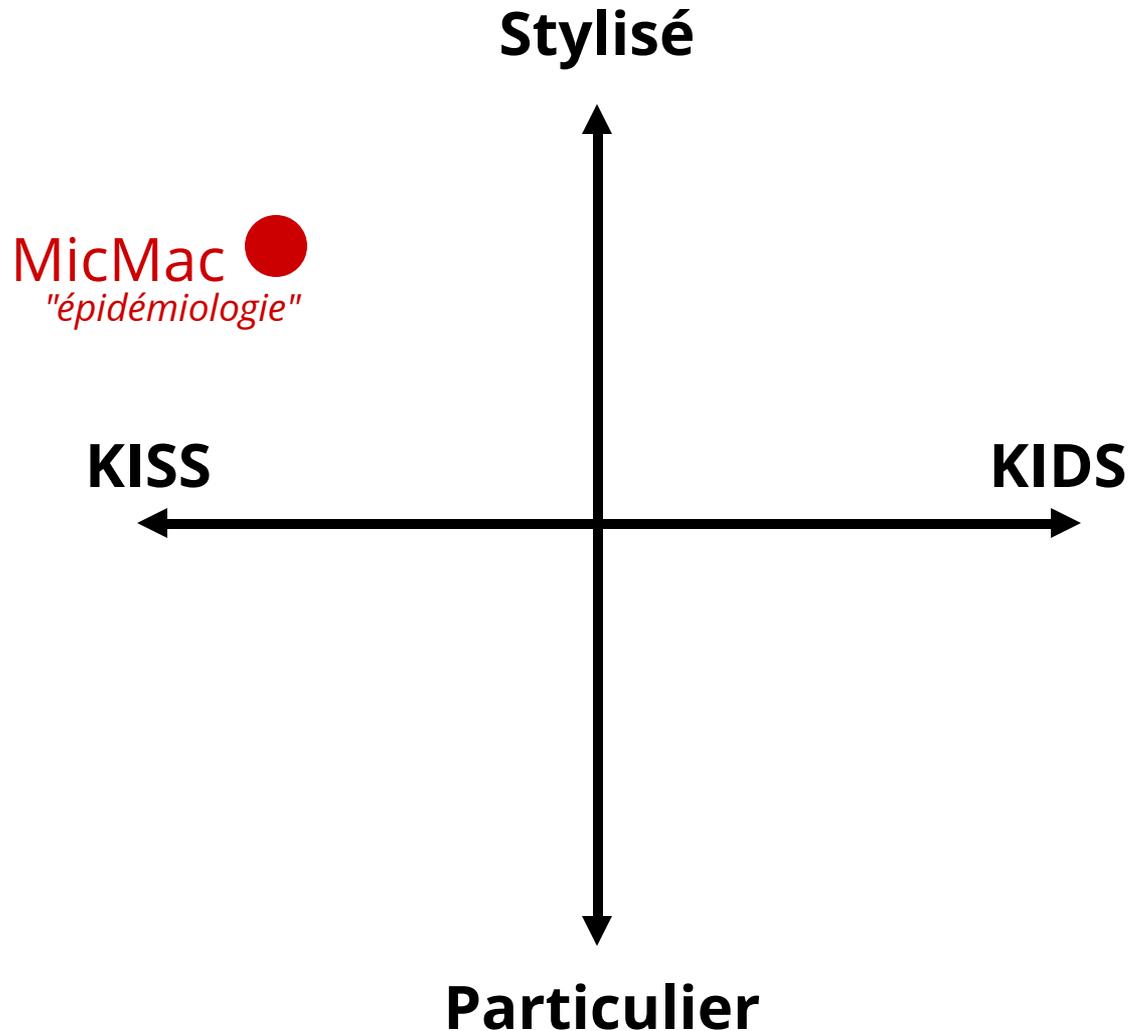


Netlogo (logo)

<https://ccl.northwestern.edu/netlogo/>

# 6 exemples

& une grande diversité de besoins en calculs : **Grille & HPC !**



# MicMac

[Banos et al. 2015]

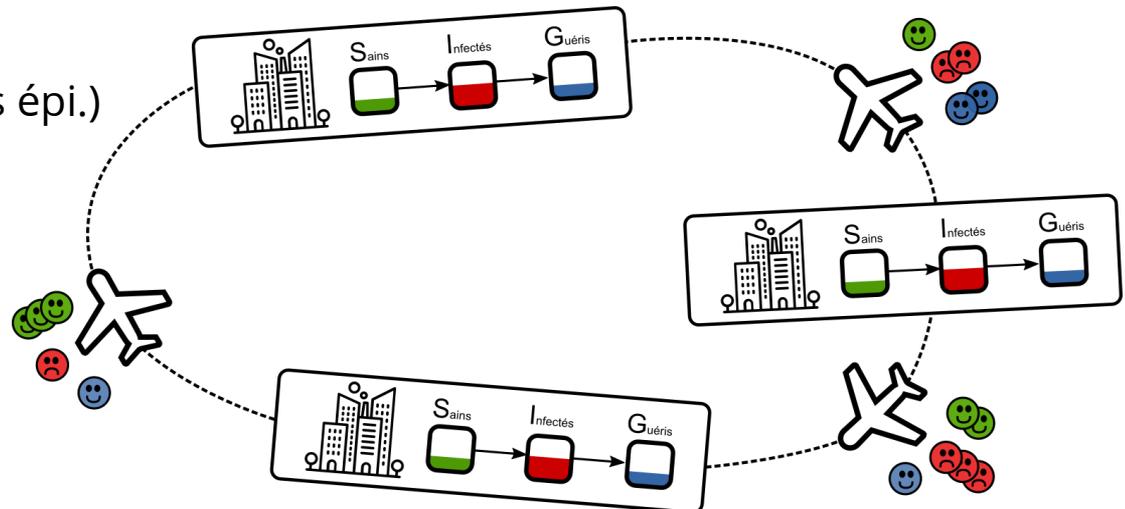
**Question :** Quel est la combinaison de **stratégies (quarantaine, évitement, responsabilisation ind., protection ind.)** qui permet de limiter à la fois les pertes économiques, et la diffusion de l'épidémie dans un tel système ?

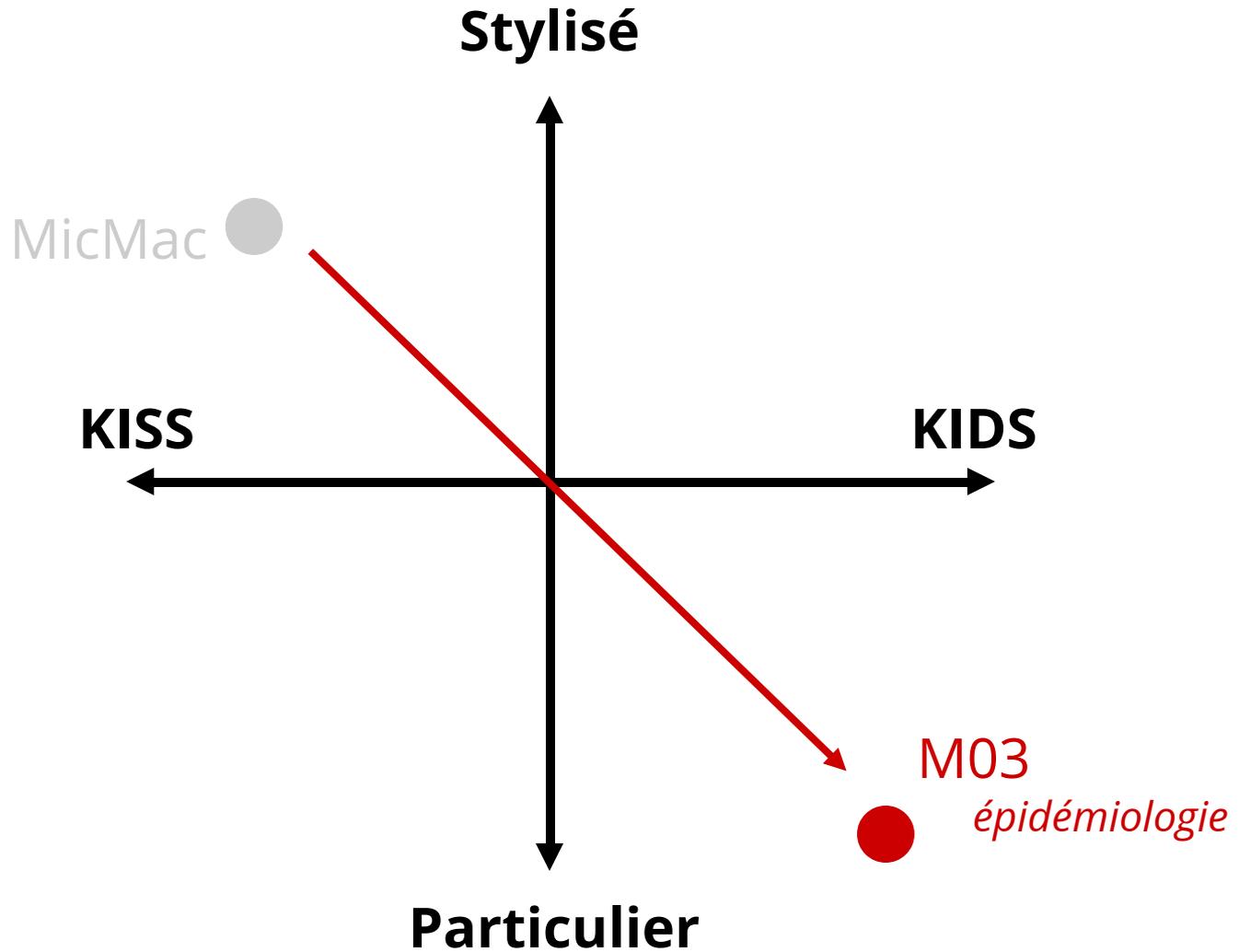


~ **5/10** minutes d'exécution  
1 simulation par CPU

**agents** = personnes & avion & villes ;  
couplage avec ODE  
**4** stratégies à explorer  
**2** objectifs à minimiser (écon. vs épi.)  
**pas** de données

$$\left\{ \begin{array}{l} \frac{dS}{dt} = -\beta IS \\ \frac{dI}{dt} = \beta IS - \alpha I \\ \frac{dR}{dt} = \alpha I \end{array} \right.$$



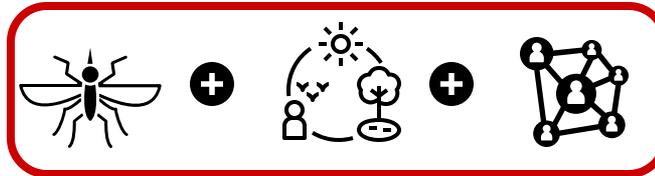




[Daudé et al. 2017]

2011 & 2019

# Objectif : Comprendre et **lutter** contre l'émergence et la diffusion de la **Dengue** dans les milieux urbains



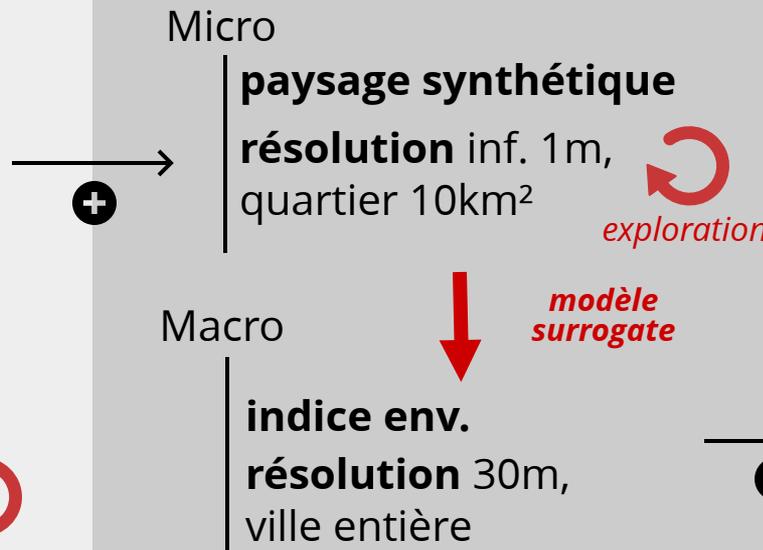
*ou ?  
comment ?  
quand ?*

 **Moma**  
[Maneerat et al. 2017]

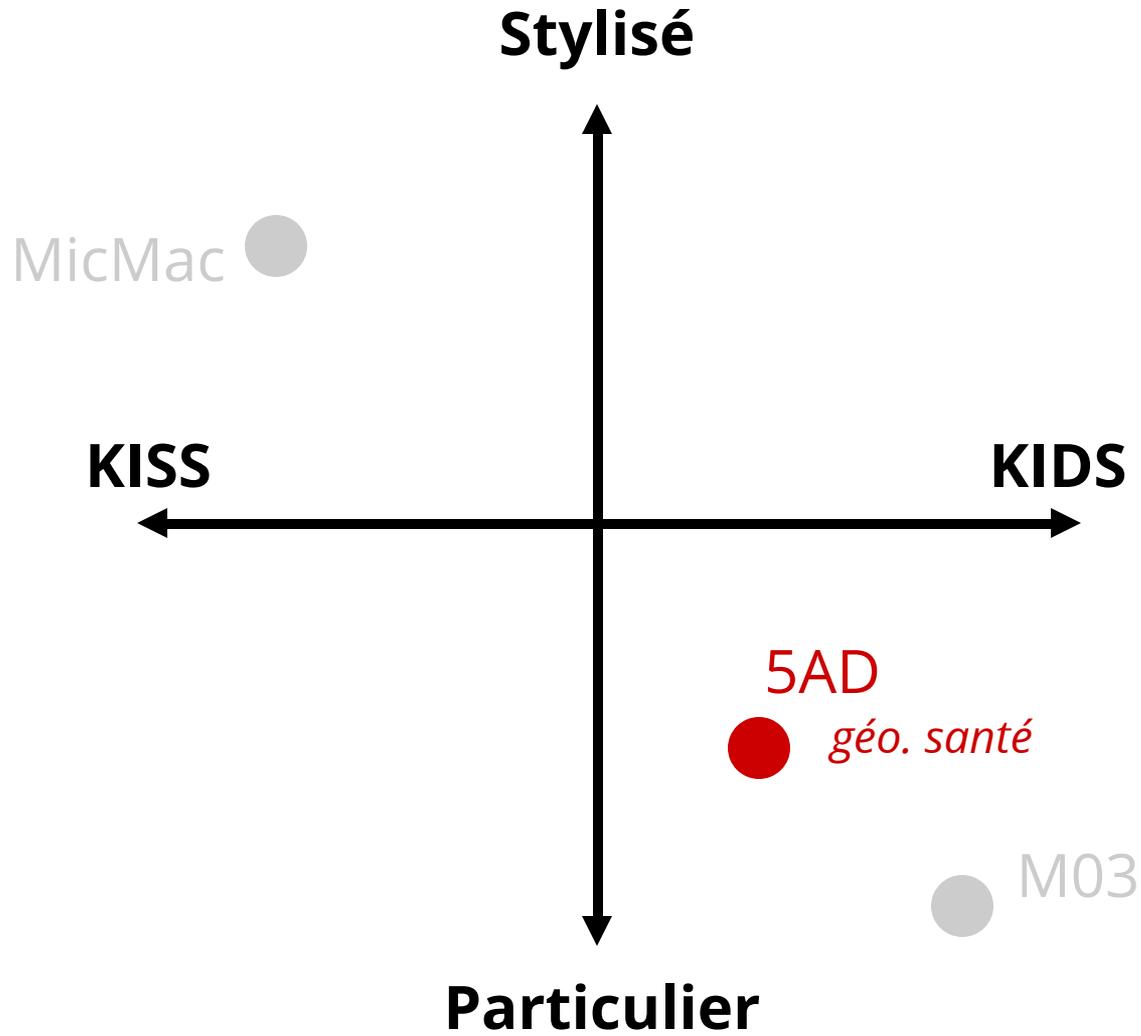
 **Mode**  
[Misslin 2017]

 **Momos**  
[Cebeillac et al. 2017]

**agents** = moustiques & nids  
**résolution** inf. 1m  
**exécution** : + 12 heures  
+ 50 paramètres  
**1 simulation = 1 cpu**  
**16 Go Ram**  
**DOE** : LHS sur grille EGI  
500000 sim, ~20 ans calculs



**agent**: individu mobile avec agenda  
**résolution** 30m, ville entière  
**données**: twitter



# 5Ad



Etudier l'effet de la **segregation sociale** dans l'évolution des **comportements de santé** sur une journée type de 24h.

[Cottineau et al. 2020, à paraître]



**Question :** Arrive-t-on à **reproduire** de façon cohérente les données et l'inégalité de comportement de santé en fonction des groupes sociaux ?

**agents:** personnes / 8.16 M agents mobiles, en interaction

**espace:** IDF à 1km x 1km

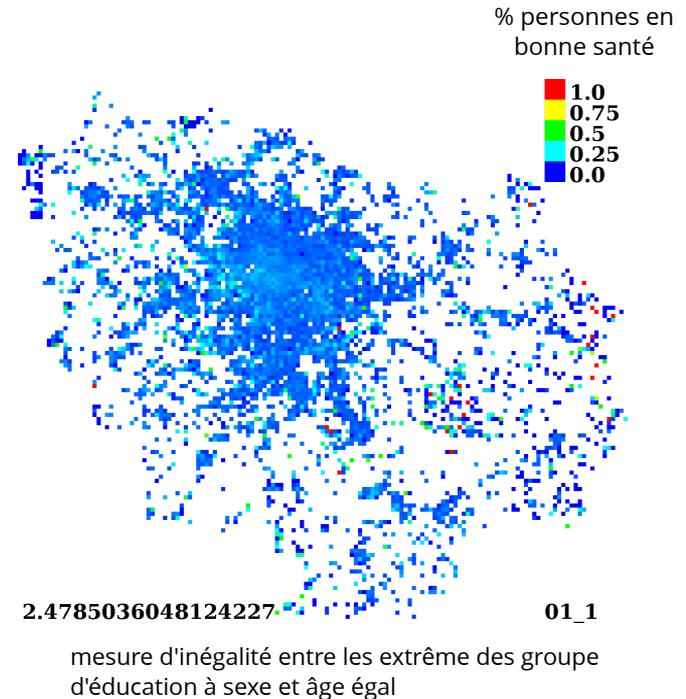
**données:** nombreuses, et à différentes dates (NPPS, CSP 18 catégories, OD, baromètre santé, etc.)

1 modèle changement d'opinion

5 paramètres

3 objectifs à minimiser

4 scénarii de mobilité (random, données, etc.)



~ 25 minutes d'exécution

~ 3 Go ram par simulation

1 core, mais **multithreading possible**

**test :** 100.000 execution

**calibrage:** 200.000 executions



# h24

2 scénarii à comparer :

	Attribution résidence aléatoire	Attribution résidence & mobilité via les données
pas de mobilité (résidence)	<b>Scenario 1</b>	Scenario 2
mobilité journalière (nuit,matin,soir)	Scenario 3	<b>Scenario 4</b>

via 3 objectifs à minimiser :

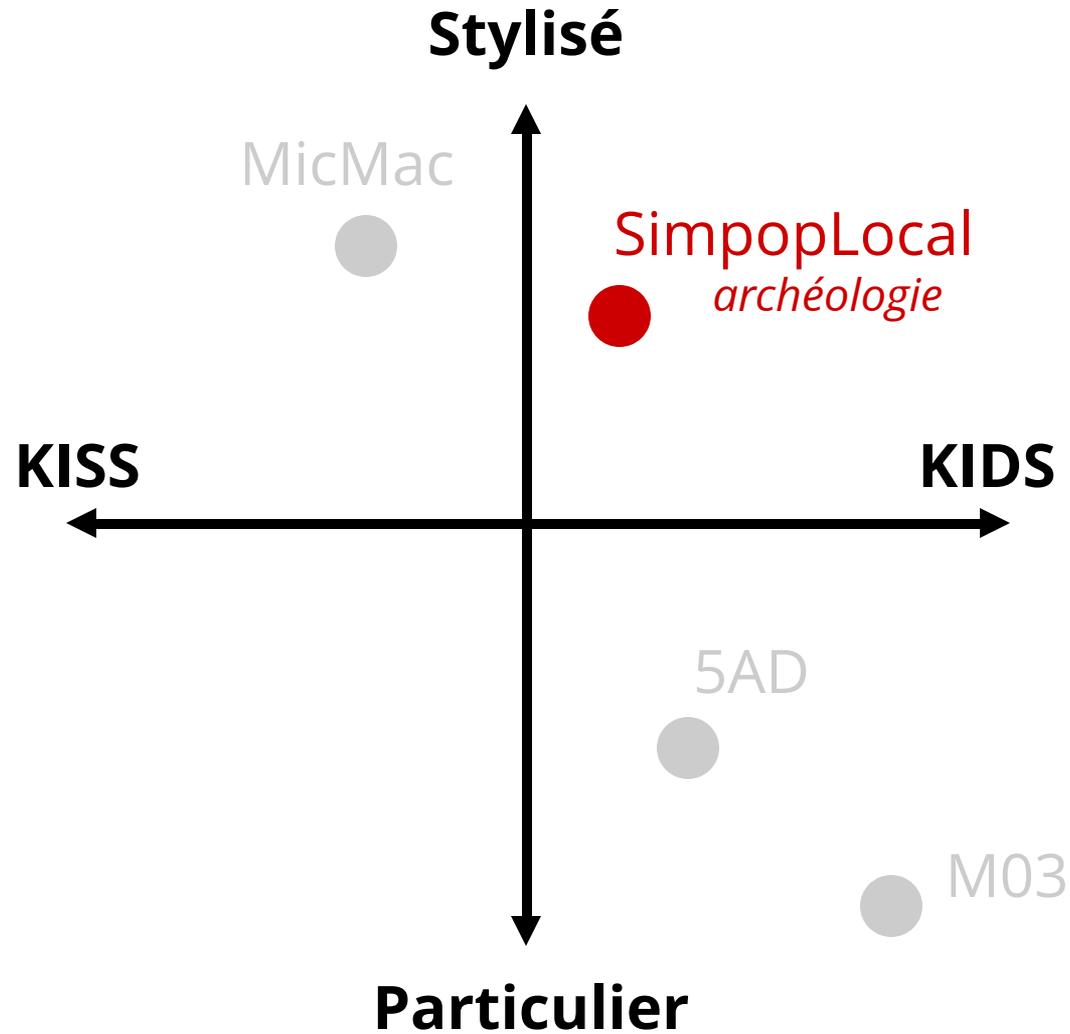
**Obj 1. :** MSE sum **Scenario 1**

$$\sum_{i \in (2002, 2008)} MSE_i$$

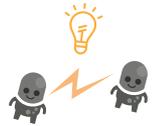
**Obj. 2 :** MSE sum **Scenario 4**

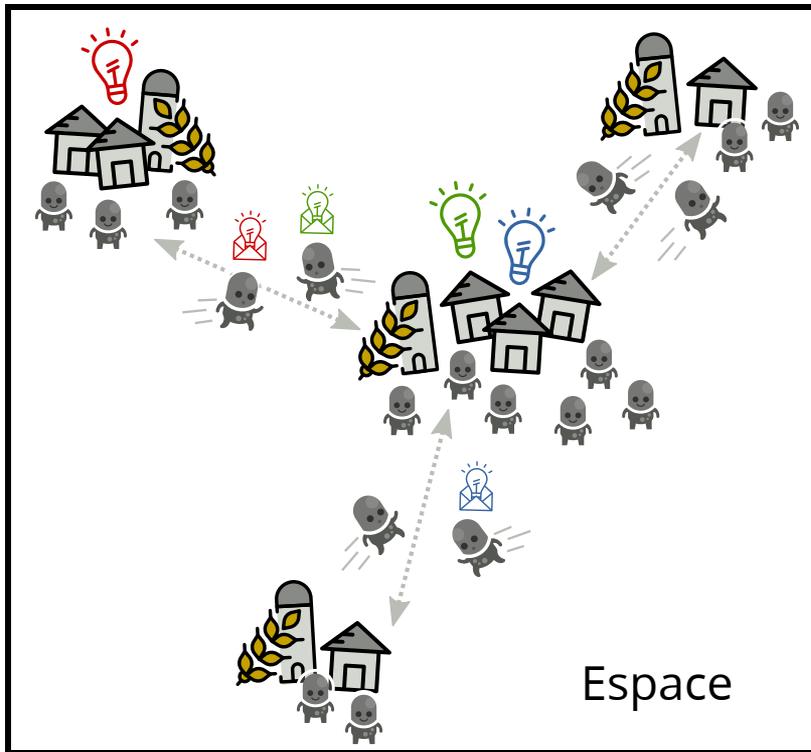
**Obj. 3 :** Le delta sur la mesure d'**inégalité (SI)** entre S1 et S4

$$\Delta(SI_{random}, SI_{mobility})$$



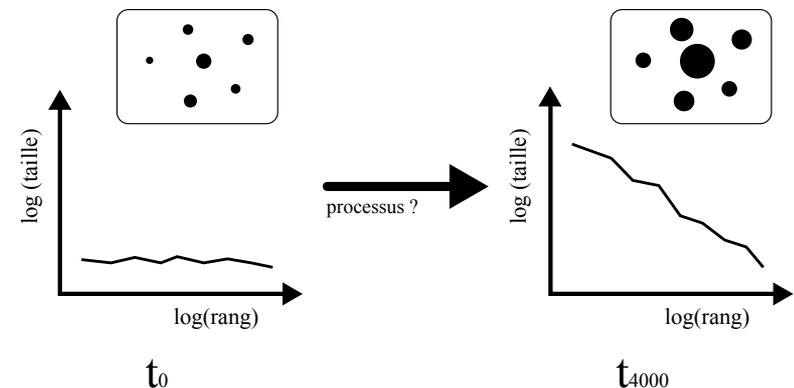
# SimpopLocal

**Question : La production et la diffusion de l'innovation**   
est il un mécanisme nécessaire pour produire une **hiérarchisation**  
**crédible** du système de peuplement en 4000 ans au **néolithique** ?



système de peuplements

**agents:** villes & innovations   
**7** paramètres libres  
**4** mécanismes  
**très peu** de données en entrée/sorties  
**1 loi empirique** à reproduire



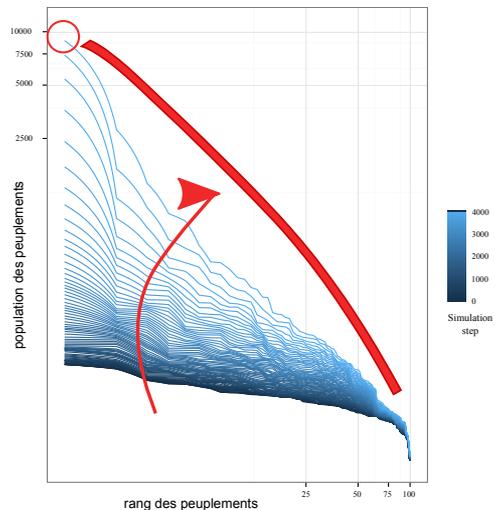
# SimpopLocal



## Calibration

NSGA 2 - Ilôts

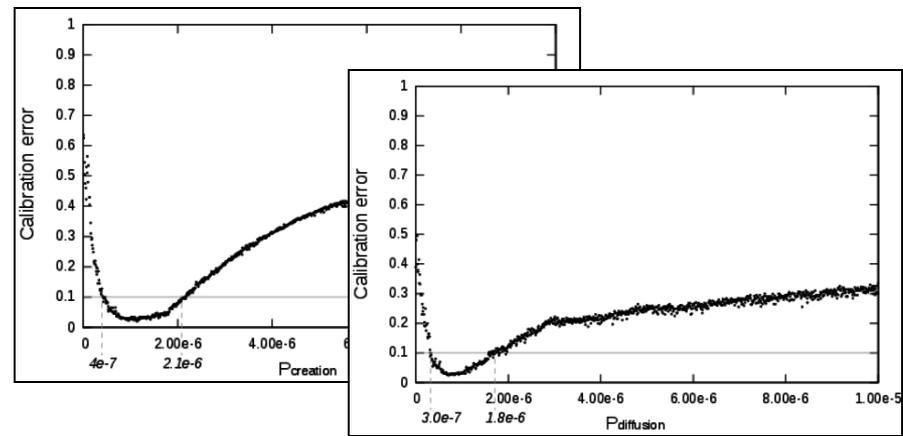
3 objectifs : (temps, forme, taille)



## Analyse de sensibilité

CP-Profile

1 objectif :  $f = \max(01,02,03)$



obj. f est **impossible** si 0 innovations

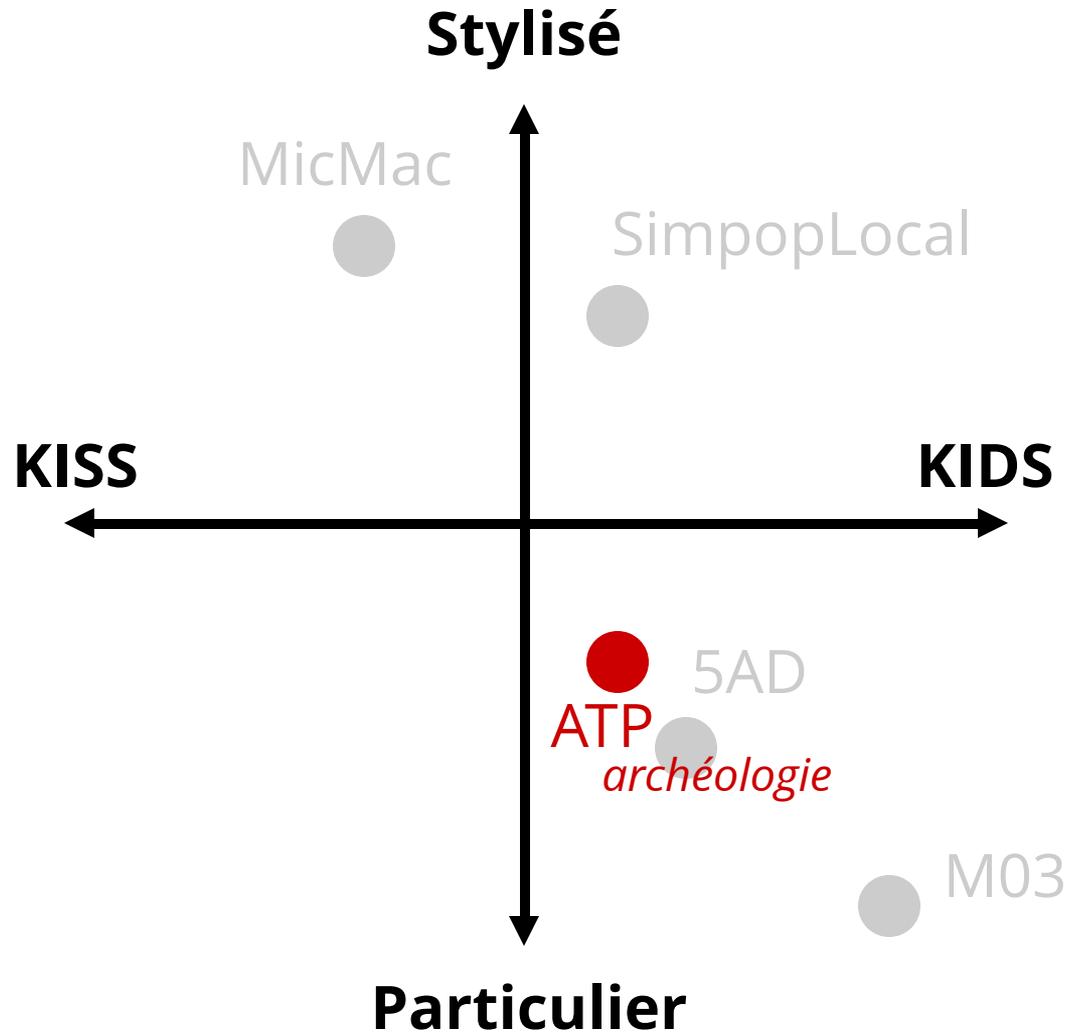


**1 simulation = 1 coeur**  
**1 Go Ram**  
~ **5/10** minutes d'exécution  
**500** millions de simulation  
**eq. 20** ans de calculs  
**4000 cpu EGI Biomed**



→  
réécriture modèle

**1 simulation = 1 coeur**  
**1 Go Ram**  
~ **1/2** minutes d'exécution  
**eq. 12** ans calcul **par profil**  
**15 jours** de calculs (profil x 6)  
**4000 cpu EGI Biomed**



interdisciplinarité

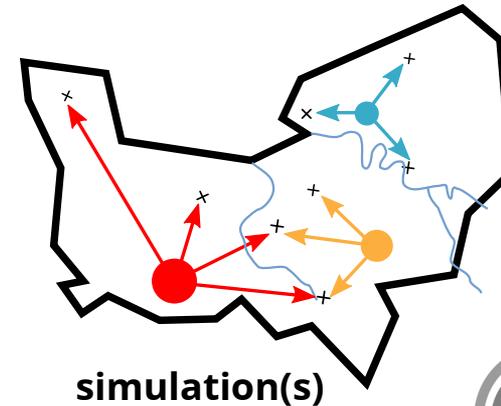
# Rin ATP



CRIANN



modèle(s)  
→  
acteurs / objets  
processus  
ordre de grandeurs



simulation(s)



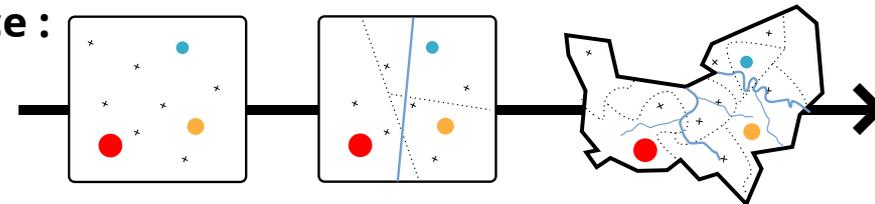
**Objectif** : Simuler pour mieux comprendre les règles de circulation des archéomatériaux (craies) depuis les carrières jusqu'aux sites de construction, en Normandie, entre le 11ème - 15ème siècle



**1 simulation = 1 coeur**  
**1 Go Ram**  
~ **20** minutes d'exécution  
**... campagne en cours ...**

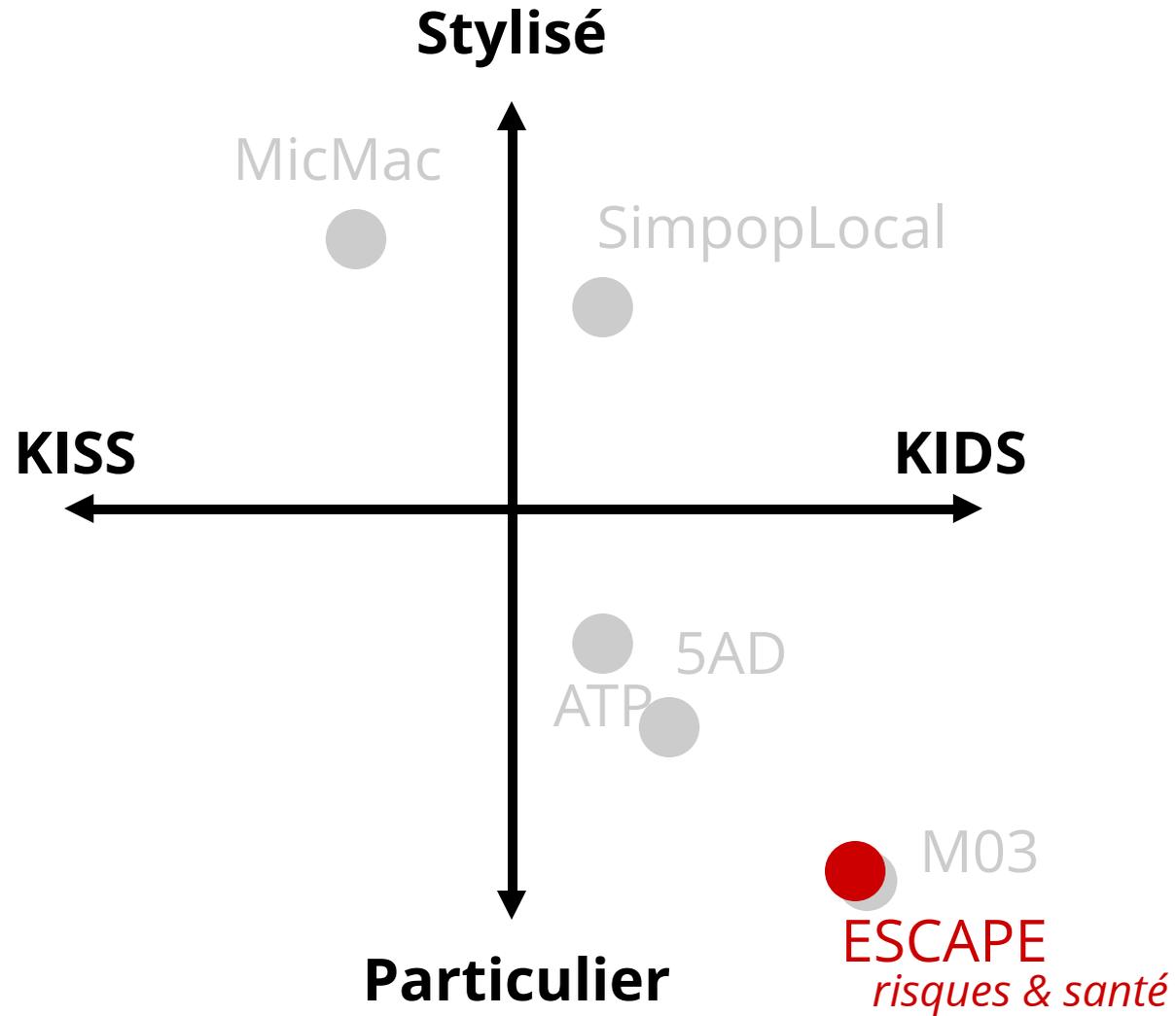
**agents:** carrières et sites

**espace :**



**données** : très peu, fait stylisés & ordre de grandeurs

**observable** : volume par type de matériaux  
par type de batiments



**Objectif :** Le projet ESCAPE a pour objectif de développer un système d'aide à la modélisation et à la simulation d'**évacuation** de villes afin d'**évaluer** des scénarios de **gestion de crise**.

**agents:** automobiles, bus, piétons, organisations, etc.

**mécanismes:** pop synthétique, agenda, multimodalité, arbres décisions, etc.

**données :** trafic (voiture, transports en communs), Insee, EMD, etc.



Trois démonstrateurs :

**Rouen (risque tech),**

Hanoï et val d'Authion (risque nat)



**1 simulation = 1 coeur**  
**~ 8/16 Go Ram**  
**+30 minutes**



# Questions ?



<https://slides.com/sebastienreycyorehourcq/jcad2019>



[Lien vers la bibliographie](#)

# Table ronde

Quelques chiffres pour les autres laboratoires  
de géographie quanti. (hors GPU)

4 laboratoires de quantitatives (avec simulation) en géographie

**UMR LaSTIG** : Laboratoire en Sciences et technologies de  
l'information géographique

*discussions avec Julien Perret / DR LaSTIG STRUDEL*

**UMR Géographie-cités / Equipe P.a.r.i.s**

*discussions avec Romain Reuillon / CR Géographie-cités*

**UMR IDEES** : Identité & Différenciation des Territoires

*Sébastien Rey-Coyrehourcq / IGR Université*

**UMR Théma** : Théoriser et modéliser pour aménager

*Gilles Vuidel / IGR CNRS*

# Table ronde

## ex des projets aux LaSTIG

	<b>Cluster</b>	<b>Grid</b>	<b>Cloud</b>	<b>Autres</b>	<b>Outil</b>
<b>Simurbapy</b>	via cloud	EGI VO ISCPIF	LAL SLURM		OpenMole
<b>ArtiScales</b>	via cloud	EGI VO ISCPIF	LAL SLURM		OpenMole
<b>SoDuCo</b>				HUMANUM (vm)	(OpenMole)
<b>HIATUS</b>				Zebulon (SSH)	MicMac IGN
<b>IQmulus</b>	via cloud		LAL SPARK		OpenMole
<b>Alegoria</b>			LAC DC/OS	LAL (vm)	spark/mesos
<b>Anatole G.</b>				CNES	
<b>5ad</b>		EGI VO ISCPIF		LAL (vm)	OpenMole
...					

# Table ronde

## Référents / Interface moyen de calculs

	<b>GRID</b>	<b>Cluster</b>	<b>Cloud</b>	<b>Openmole</b>
Quoc-Dinh NGUYEN			<b>X</b>	
Laurent Caraffa & Mathieu Brédif		via cloud	X (SPARK)	
Paul Chapron & Julien Perret	X (EGI VO ISCPHF)	via cloud	X (SLURM LAL)	<b>X</b>
Didier Richard*		X	X	

\* *gestion et/ou achats des machines*

# Table ronde

## ex des projets HPC à Géographie-cités ( via ISCPIF )

	Cluster	Grid	Cloud	Autres	Outil
<b>5Ad</b>		EGI VO ISCPIF	LAL (vm)		OpenMole
<b>World3</b>		EGI VO ISCPIF	LAL (vm)		OpenMole
<b>Escape</b>	SLURM méso CRIANN	EGI VO ISCPIF	DSI UR (vm)		OpenMole
<b>SimFéodal</b>				WS (12cpu)	BASH / SSH
<b>Mo3</b>	SLURM méso CRIANN	EGI VO ISCPIF	DSI UR (vm)		OpenMole
...					

## Référents / Interface moyen de calculs

	GRID	Cluster	Cloud	Autre	Openmole
Mathieu Leclaire*	<b>X</b> ( EGI / VO ISCPIF)			<b>X</b>	X
Romain Reuillon*	<b>X</b> (EGI / VO ISCPIF)		<b>X</b> (LAL)	<b>X</b> (Zebulon)	X
Mazyiar PANAHI*			<b>X</b> (LAL)	<b>X</b> (Zebulon)	X
Robin Cura				<b>X</b>	

\* *gestion et/ou achats des machines*

# Table ronde

## ex des projets HPC à IDEES

	Cluster	Grid	Cloud	Autres	Outil
<b>Mo3</b>	SLURM CRIANN	EGI VO ISCPHF	DSI UR (vm)	WS (12 cpu)	OpenMole
<b>Escape</b>	SLURM CRIANN		DSI UR (vm)	WS (12 cpu)	OpenMole
<b>ATP</b>	SLURM CRIANN	EGI VO ISCPHF	DSI UR (vm)		OpenMole
<b>5Ad</b>		EGI VO ISCPHF	LAL (vm)		OpenMole
...					

## Référents / Interface moyen de calculs

	GRID	Cluster	Cloud	Autre	Openmole
Sébastien R.C. *	X ( EGI / VO ISCPHF)	X (SLURM CRIANN)	X (DSI)	X (WS)	X

\* *gestion et/ou achats des machines*

# Table ronde

## ex des projets HPC à Théma

	Cluster	Grid	Cloud	Autres	Outil
<b>Pixscape</b>				(48 cpu en MSH)	MPI / Cuda
<b>MobiSim</b>	méso UFC			(48 cpu en MSH)	MPI
<b>Graphab</b>	méso UFC				MPI
<b>ArtiScales</b>	via cloud	EGI VO ISCIPIF	LAL SLURM		OpenMOLE
...					

## Référents / Interface moyen de calculs

	GRID	Cluster	Cloud	Autre	Openmole
Gilles Vuidel	?	méso UFC	?	?	

\* *gestion et/ou achats des machines*

# Table ronde

	<b>IDEES</b>	<b>THEMA</b>	<b>Géographie-Cité</b>	<b>LASTIG</b>
Référents	1	1	3	6
Projets*	4 (3)	4 (1)	5 (3)	8 (2)

\* *non exhaustif*      (*) communs entre les laboratoires*

- OpenMOLE vs MPI ...      • WS locale => prototypage ?
- Diversité des approches, mais sources similaires
  - Grille ( EGI ) & Clusters ( UFC, CRIANN )
  - Cloud VM ou instancié cluster ( LAL, DSI UR )
- Référents +/- nombreux & spécialisés !
- Projet redondants => petite communauté (ou biais ?)

# Table ronde

Complémentarité des approches / IMHO géographe

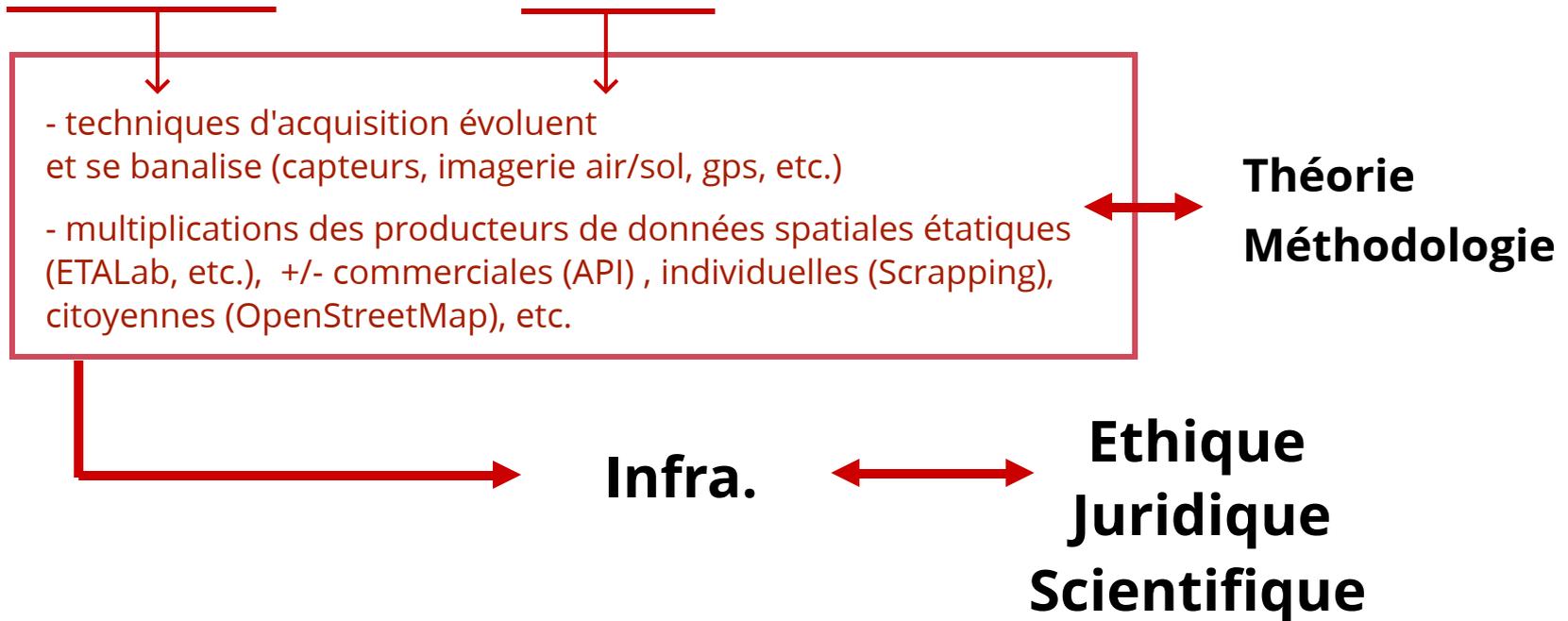
	<b>Grille</b>	<b>Cluster</b>
<b>Accès</b>	certificats	comptes
<b>Attribution</b>	accès au fil de l'eau	dépôt de projet via campagne : évaluation shs ? quantification ?
<b>Tarification</b>	gratuit	?? variable selon méso ??
<b>Quota</b>	pas de limite nb/h	plafond nb/h
<b>Performance</b>	+/++/+++ variable	+++
<b>Résilience</b>	++ (sauf webdav...)	++
<b>Stockage</b>	+	+++
<b>Environnement</b>	open bar	assez stricte
<b>Accompagnement</b>	??	utilisateurs & équipe régionale (formation, support)
<b>Participation financière ?</b>	Noeuds via DIRAC	?? variable selon méso ??
<b>Diversité CPU/GPU</b>	??	CPU/GPU
<b>Mémoire</b>	3Go	presque open bar

# Table ronde

Réflexions sur le HPC en Géo (et en SHS?)

a) Enjeux **multiples** autour des données, nbx acteurs en jeu

Volumétrie, Vitesse, Variété



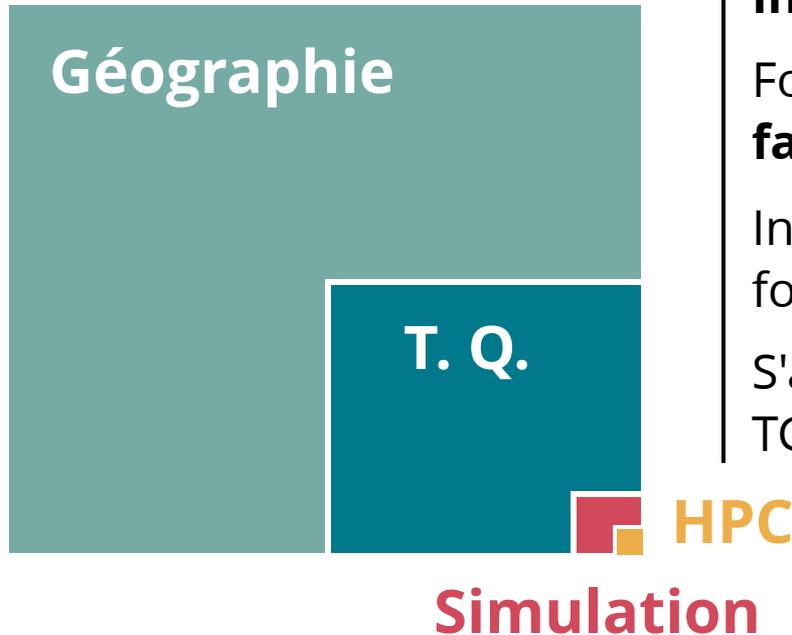
# Table ronde

Réflexions sur le HPC en Géo (et en SHS?)

## b) Enjeux **formations & valorisations (géo / shs ?)**

*Issue d'une discussion par mail avec  
Stéphane Pouyllau / IGR Directeur  
Technique TGIR HumaNum*

➔ *"demande marginale"  
"accompagnement encore limité"  
"chercheurs pas assez solide (méthodes, calcul)  
mais progression ... "*



**Inventaire** des projets/formations HPC SHS

Formations internes SHS **adaptés pour favoriser autonomie (outils, méthodes)**

Intégrer le HPC et l'accès aux HPC dans les formations & écoles d'été => **dédramatiser**

S'appuyer sur les réflexions **existantes** :  
TGIR Humanum, RMSH, GDR Magis

➔ Groupe de travail ?