

# RÉPLICABILITÉ ET REPRODUCTIBILITÉ DE LA RECHERCHE: IMPACT SUR LES PRATIQUES DES CHERCHEUSES ET CHERCHEURS

---

Arnaud Legrand



Colloque MITI  
Septembre 2023



# NO TRANSPARENCY NO CONSENSUS



*Computer science is not more related to computers than Astronomy to telescopes*  
– E. Dijkstra

L'informatique est une discipline "jeune" héritant des  
Mathématiques, Ingénierie, Sciences Naturelles, Linguistique, ...

# L'INFORMATIQUE EN TANT QUE SCIENCE

L'informatique est une discipline "jeune" héritant des Mathématiques, Ingénierie, Sciences Naturelles, Linguistique, ...

À quoi ressemble une contribution?

**Maths** définition, théorème/lemme, démonstration

**Physique** proposition hypothèse, expériences, mesure, (in)validation

**Botanique** observations, description, catégorisation, modélisation

**Ingénierie** voici les calculs permettant de construire ce pont, les matériaux, etc. On l'a construit et il tient le coup!

...

**Informatique** Un joyeux mélange 😊

- Avec des échelles de temps assez différentes des autres disciplines...

## Grid Computing as Science?

- You can tell you are a scientific discipline if
    - You can read a paper, **easily** reproduce (at least a subset of) its results, and improve
    - You can tell to a grad student "Here are the standard tools, go learn how to use them and come back in one month"
    - You can give a 1-hour seminar on **widely accepted** tools that are the basis for doing research in the area
  - We are not there today
    - But perhaps I can give a 1-hour seminar on **emerging** tools that could be the basis for doing research in the area, provided a few open questions are addressed
- Need for standard ways to run "Grid experiments"

Pratiques courantes il y a 20 ans (section "Experiments and Methods")

- On a fait des simulations.

Pratiques courantes il y a 20 ans (section "Experiments and Methods")

- On a fait des simulations.
- On a écrit un simulateur en Java.

Pratiques courantes il y a 20 ans (section "Experiments and Methods")

- On a fait des simulations.
- On a écrit un simulateur en Java.
- On a fait des *expériences* sur un Pentium III à 300MHz avec 500Mo de RAM, Linux 2.1 ...



Pratiques courantes il y a 20 ans (section "Experiments and Methods")

- On a fait des simulations.
- On a écrit un simulateur en Java.
- On a fait des *expériences* sur un Pentium III à 300MHz avec 500Mo de RAM, Linux 2.1 ... et ~~Java~~ Python

Pratiques courantes il y a 20 ans (section "Experiments and Methods")

- On a fait des simulations.
- On a écrit un simulateur en Java.
- On a fait des *expériences* sur un Pentium III à 300MHz avec 500Mo de RAM, Linux 2.1 ... et ~~Java~~ Python
- On a développé un nouveau simulateur, en open source,

Pratiques courantes il y a 20 ans (section "Experiments and Methods")

- On a fait des simulations.
- On a écrit un simulateur en Java.
- On a fait des *expériences* sur un Pentium III à 300MHz avec 500Mo de RAM, Linux 2.1 ... et ~~Java~~ Python
- On a développé un nouveau simulateur, en open source, sur demande,

Pratiques courantes il y a 20 ans (section "Experiments and Methods")

- On a fait des simulations.
- On a écrit un simulateur en Java.
- On a fait des *expériences* sur un Pentium III à 300MHz avec 500Mo de RAM, Linux 2.1 ... et ~~Java~~ Python
- On a développé un nouveau simulateur, en open source, sur demande, sur ma page web (dead within a year)

Pratiques courantes il y a 20 ans (section "Experiments and Methods")

- On a fait des simulations.
- On a écrit un simulateur en Java.
- On a fait des *expériences* sur un Pentium III à 300MHz avec 500Mo de RAM, Linux 2.1 ... et ~~Java~~ Python
- On a développé un nouveau simulateur, en open source, sur demande, sur ma page web (dead within a year)
- On s'est reposé sur NS2/GTNetS/GridSim/OptorSim/...

## Pratiques courantes il y a 20 ans (section "Experiments and Methods")

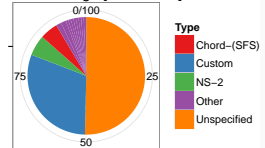
- On a fait des simulations.
- On a écrit un simulateur en Java.
- On a fait des *expériences* sur un Pentium III à 300MHz avec 500Mo de RAM, Linux 2.1 ... et ~~Java~~ Python
- On a développé un nouveau simulateur, en open source, sur demande, sur ma page web (dead within a year)
- On s'est reposé sur NS2/GTNetS/GridSim/OptorSim/...
  - Mais... quelle version, avec quels paramètres?

# UNE "CRISE" DE LA REPRODUCTIBILITÉ EN INFORMATIQUE ? (2/2)

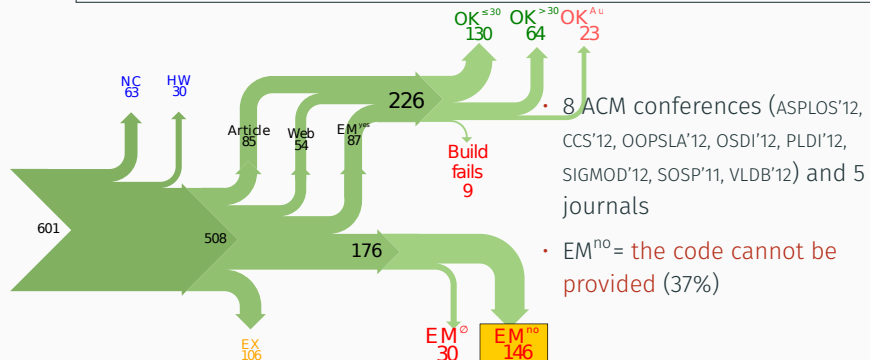
Naicken, Stephen et Al., *Towards Yet Another Peer-to-Peer Simulator*, HET-NETs'06.

From 141 P2P sim.papers, 30% use a custom tool,  
50% don't report used tool

Simulator usage [Naicken06]



Collberg, Christian et Al., *Measuring Reproducibility in Computer Systems Research*,  
<http://reproducibility.cs.arizona.edu/>



**Transparence** (compréhension) Comment est-ce que ça a été fait ?

**Confiance** (vérification): Est-ce juste ? Sous quelles conditions ?

Cherry-Picking ?

**Réutilisation** *Se hisser sur les épaules de géants*

Conditions nécessaires:

1. Accès pérenne au code et aux données
2. Description précise et compréhensible des procédures
3. Plus de "zones grises" (open source, open data, open hardware, open laboratory notebook, etc.)



## 2005-2015 nouveaux outils et standards émergents

- Plates-formes: Forge, GitHub, GitLab, ArXiv, Zenodo, SWH
- Notebooks: Jupyter, Rstudio, Org-Mode, ...
- Une myriade de "micro"-outils communautaires

# ÉVOLUTION DES PRATIQUES 1/2

**2005-2015** nouveaux outils et standards émergents

- Plates-formes: Forge, GitHub, GitLab, ArXiv, Zenodo, SWH
- Notebooks: Jupyter, Rstudio, Org-Mode, ...
- Une myriade de "micro"-outils communautaires

**2010 - ...** changement des pratiques de publications

- *Manifeste* dans des équipes pionnières:
  - "I solemnly pledge" (WSSSPE, Lorena Barba, FAIR)



# ÉVOLUTION DES PRATIQUES 1/2

## 2005-2015 nouveaux outils et standards émergents

- Plates-formes: Forge, GitHub, GitLab, ArXiv, Zenodo, SWH
- Notebooks: Jupyter, Rstudio, Org-Mode, ...
- Une myriade de "micro"-outils communautaires

## 2010 - ... changement des pratiques de publications

- *Manifeste* dans des équipes pionnières:
  - "I solemnly pledge" (WSSSPE, Lorena Barba, FAIR)
- Expérimentation petite conférences/workshops  
~> conférences phare
  - Artifact evaluation: SC
  - Reproducibility challenges: NeurIPS
  - Open reviews: ICLR



# ÉVOLUTION DES PRATIQUES 1/2

## 2005-2015 nouveaux outils et standards émergents

- Plates-formes: Forge, GitHub, GitLab, ArXiv, Zenodo, SWH
- Notebooks: Jupyter, Rstudio, Org-Mode, ...
- Une myriade de "micro"-outils communautaires

## 2010 - ... changement des pratiques de publications

- *Manifeste* dans des équipes pionnières:
  - "I solemnly pledge" (WSSSPE, Lorena Barba, FAIR)
- Expérimentation petite conférences/workshops  
↪ conférences phare
  - Artifact evaluation: SC
  - Reproducibility challenges: NeurIPS
  - Open reviews: ICLR
- Valorisation internationale: ACM Reproducibility badges



- Engagement national/international: DORA, PNSO, COSO, COARA, etc.



Reproductibilité = "nouveau" *critère* de qualité!

Perceptible dans les publications, et au comité national

- Code source sur GitHub 😊
- Jupyter Notebook 😊

Reproductibilité = "nouveau" critère de qualité!

Perceptible dans les publications, et au comité national

- Code source sur GitHub 😊
- Jupyter Notebook 😊
- *"Dans un soucis de reproductibilité, le code et les données sont disponibles à l'URL suivante" (Zenodo ?)* 😊

Reproductibilité = "nouveau" critère de qualité!

Perceptible dans les publications, et au comité national

- Code source sur GitHub 😊
- En un seul commit! Boom! 😞
- Jupyter Notebook 😊
- *"Dans un soucis de reproductibilité, le code et les données sont disponibles à l'URL suivante" (Zenodo ?)* 😊

## ÉVOLUTION DES PRATIQUES 2/2

Reproductibilité = "nouveau" critère de qualité!

Perceptible dans les publications, et au comité national

- Code source sur GitHub 😊
- En un seul commit! Boom! 😞
- Jupyter Notebook 😊
- Des codes assez trivaux qui ne précisent pas ou mal leur environnement logiciel 😞
- *"Dans un soucis de reproductibilité, le code et les données sont disponibles à l'URL suivante" (Zenodo ?)* 😊



# ÉVOLUTION DES PRATIQUES 2/2

Reproductibilité = "nouveau" critère de qualité!

Perceptible dans les publications, et au comité national

- Code source sur GitHub 😊
- Jupyter Notebook 😊
- *"Dans un soucis de reproductibilité, le code et les données sont disponibles à l'URL suivante" (Zenodo ?) 😊*
- En un seul commit! Boom! 😞
- Des codes assez trivaux qui ne précisent pas ou mal leur environnement logiciel 😞
- Zenodo, SWH 😞
  - dépôt GitHub vide
  - Google-drive, un `.zip` de 300Mo avec un `.exe`
  - Une archive avec un binaire dont certaines parties sont spécifiques au supercalculateur

# ÉVOLUTION DES PRATIQUES 2/2

Reproductibilité = "nouveau" critère de qualité!

Perceptible dans les publications, et au comité national

- Code source sur GitHub 😊
- Jupyter Notebook 😊
- *"Dans un soucis de reproductibilité, le code et les données sont disponibles à l'URL suivante" (Zenodo ?)* 😊
- En un seul commit! Boom! 😞
- Des codes assez trivaux qui ne précisent pas ou mal leur environnement logiciel 😞
- Zenodo, SWH 😞
  - dépôt GitHub vide
  - Google-drive, un `.zip` de 300Mo avec un `.exe`
  - Une archive avec un binaire dont certaines parties sont spécifiques au supercalculateur
- Les outils et les ressources existent maintenant et on peut faire un excellent travail quand on s'en saisit.
- Mais tant que les "vieux" (mal formés) ne s'assureront pas que les "jeunes" sont bien formées, ça prendra du temps. 😊