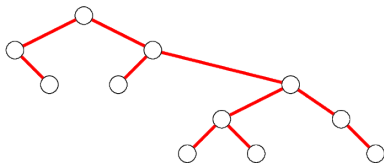Viviane Pons

Maîtresse de conférence, Univ. Paris-Sud Orsay
*viviane.pons@lri.fr – @PyViv*

# Science and Open Source
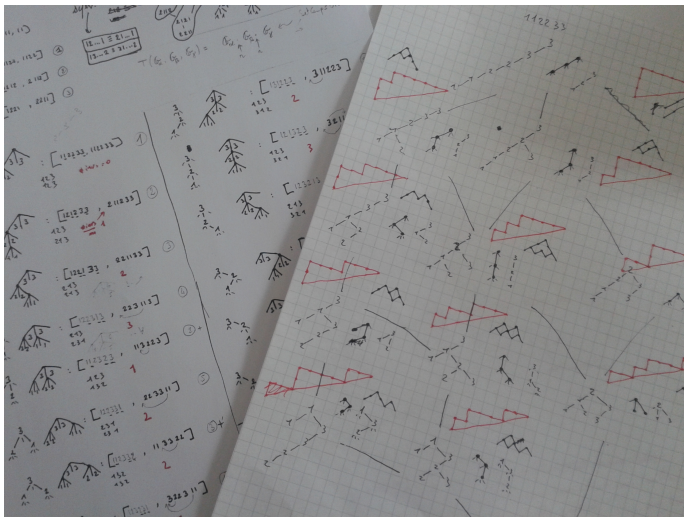
## What do we learn from each other?

# Who am I?

- Both a **mathematician** and a **computer scientist**.



- I do **combinatorics**.

```python
def inf_mperms(p1,p2):
    m = len([i for i in p1 if i==1])
    return perm_to_mperm(inf_perms(mperm_to_perm(p1),mperm_to_perm(p2)),m)

def is_last(perm,i):
    for b in perm[i+1:]:
        if b == perm[i]:
            return False
    return True

def mperm_to_tree(perm):
    values = list(set(perm))
    values.sort()
    values.reverse()
    m = len(perm) / len(values)
    tree = MDecreasingTree(m+1,None)
    for v in values:
        tree = tree.insert_from_mperm(perm,v)
    return tree


def mperm_to_tree2(perm, mfor0 = 1):
    if len(perm)==0:
        return MDecreasingTree(mfor0,None)
    n = max(perm)
    posr = [i for i in xrange(len(perm)) if perm[i]==n]
    m = len(posr)
    children = [[] for i in xrange(m+1)]
    right = {a for a in perm if a!=n}
    for i in xrange(m):
        pos = posr[i]
        for j in xrange(pos-1,-1,-1):
            a = perm[j]
            if a!=n:
                if is_last(perm,j):
                    if a in right:
                        children[i].append(a)
                        right.remove(a)
                elif a in right:
                    right.update([aa for aa in children[i] if aa < a])
                    children[i] = [b for b in children[i] if b >a]
    children[-1] = list(right)
    children_trees = [mperm_to_tree2([a for a in perm if a in c], mfor0 =m) for c in children]
    return MDecreasingTree(m+1,children_trees, label=n)



#tested 2,2 to 2,5
#tested 3,2 to 3,4
def test_sup_max_classes(m,n):
    maxs = list(max_classes(m,n))
```

My research **relies on code**. For every paper I write, there is a program somewhere with experiments and tests.

To know more: see **Experimental pure mathematics using Sage**.

# Science and Software

S.J. Hettrick et al, *UK Research Software Survey 2014*, DOI:10.5281/zenodo.1183562

- ▶ 92% of academics use research software;

# Science and Software

S.J. Hettrick et al, *UK Research Software Survey 2014*, DOI:10.5281/zenodo.1183562

- ▶ 92% of academics use research software;
- ▶ 69% say that their research would not be practical without it;

# Science and Software

S.J. Hettrick et al, *UK Research Software Survey 2014*, DOI:10.5281/zenodo.1183562

- ▶ 92% of academics use research software;
- ▶ 69% say that their research would not be practical without it;
- ▶ 56% develop their own software.

# Science and Software

S.J. Hettrick et al, *UK Research Software Survey 2014*, DOI:10.5281/zenodo.1183562

- ► 92% of academics use research software;
- ► 69% say that their research would not be practical without it;
- ► 56% develop their own software.

What about math?

39 computer algebra systems listed on Wikipedia.

Some **non** open-source

- ▶ Maple: $2275 (Commercial), $2155 (Government), $1245 (Academic), $239 (Personal), $99 (Student)
- ▶ Mathematica: $2495 (Professional), $1095 (Education), $295 (Personal), $140 (Student)
- ▶ Magma: $1440

(numbers from Wikipedia)

# The shared values of science and open-source

# The shared values of science and open-source

- **As a scientist, I want to contribute.**

# The shared values of science and open-source

- **As a scientist, I want to contribute.**
- **As a scientist, I want to create knowledge... and share it.**

# The shared values of science and open-source

- **As a scientist, I want to contribute.**
- **As a scientist, I want to create knowledge... and share it.**
  - All my papers are accesible.
  - All my teaching material is open-source.
  - All my code is open-source.

# The shared values of science and open-source

- **As a scientist, I want to contribute.**
- **As a scientist, I want to create knowledge... and share it.**
  - All my papers are accesible.
  - All my teaching material is open-source.
  - All my code is open-source.
- **As a scientist, I collaborate.** Joining forces to create something better.

# The shared values of science and open-source

- **As a scientist, I want to contribute.**
- **As a scientist, I want to create knowledge... and share it.**
  - All my papers are accesible.
  - All my teaching material is open-source.
  - All my code is open-source.
- **As a scientist, I collaborate.** Joining forces to create something better.
- **As a scientist, I want my results to be re-used and improved.**

**As a scientist, I need to write open-source**
I'm not the only one...

**As a scientist, I need to write open-source**
I'm not the only one...



Richard Stallman
creator of GNU, MIT



Donald Knuth
creator of TeX, Stanford

# Math and open-source

A whole ecosystem of open-source math software.

# Math and open-source

A whole ecosystem of open-source math software.

- Specialized libraries: GAP, Linbox, Pari/GP, MPIR, Singular,...

# Math and open-source

A whole ecosystem of open-source math software.

- ▶ Specialized libraries: GAP, Linbox, Pari/GP, MPIR, Singular,...
- ▶ General purpose systems: SageMath

# Math and open-source

A whole ecosystem of open-source math software.

- ▶ Specialized libraries: GAP, Linbox, Pari/GP, MPIR, Singular,...
- ▶ General purpose systems: SageMath
- ▶ Interactive computing environments: IPython/Jupyter, CoCalc

# Math and open-source

A whole ecosystem of open-source math software.

- ▶ Specialized libraries: GAP, Linbox, Pari/GP, MPIR, Singular,...
- ▶ General purpose systems: SageMath
- ▶ Interactive computing environments: IPython/Jupyter, CoCalc
- ▶ Together with the wider Scientific Python ecosystem

# The SageMath project

# The SageMath project



▶ Started in 2005 by William Stein

# The SageMath project



- Started in 2005 by William Stein
- Built around many pre-existant software.

# The SageMath project



- ▶ Started in 2005 by William Stein
- ▶ Built around many pre-existant software.
- ▶ Grew its own python (and cython) library on top of it.
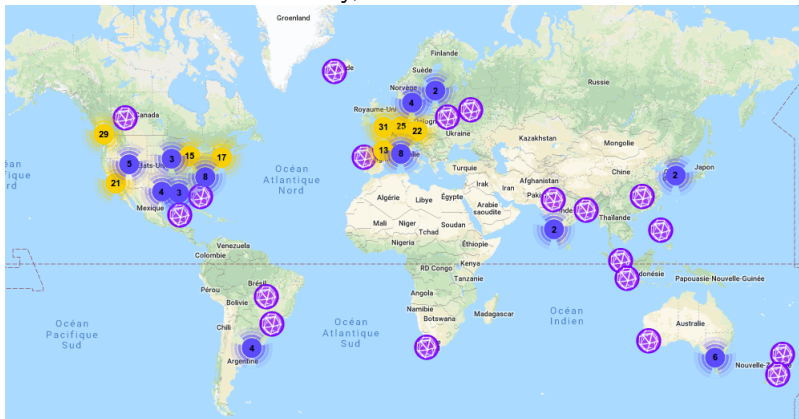
# The SageMath project



- ▶ Started in 2005 by William Stein
- ▶ Built around many pre-existant software.
- ▶ Grew its own python (and cython) library on top of it.
- ▶ And a vibrant community.

# Currently, 271 contributors

# Me and Sage

The combinatorics community "moved" to SageMath shortly before I joined in 2010.

I am a **SageMath native**!

# The challenges...

# The challenges...

**We need funding, recognition, and sustainability**

# The challenges...

**We need funding, recognition, and sustainability**

- ▶ Who pays for the project?

# The challenges…

**We need funding, recognition, and sustainability**

- ▶ Who pays for the project?
- ▶ How is the development work valued in the community?

# The challenges...

**We need funding, recognition, and sustainability**

- ▶ Who pays for the project?
- ▶ How is the development work valued in the community?
- ▶ How is your software going to survive?

# The SageMath development model

The motto is: **for users, by users**. Mostly developed by researchers:

# The SageMath development model

The motto is: **for users, by users**. Mostly developed by researchers: we need **better academic recognition for research software development**.

# The SageMath development model

## Who does the dirty work?

# The SageMath development model

## Who does the dirty work?

Packaging, interfaces, install scripts, low level software interaction, etc.

# The SageMath development model

## Who does the dirty work?

Packaging, interfaces, install scripts, low level software interaction, etc.

## Low research value, high technical value

56% of academics develop their own research software, **only 21%** of those have received training in software development.

# The SageMath development model

### Who does the dirty work?

Packaging, interfaces, install scripts, low level software interaction, etc.

### Low research value, high technical value

56% of academics develop their own research software, **only 21%** of those have received training in software development.

**We need Research Software engineers!**

# The SageMath development model

## Who does the dirty work?
Packaging, interfaces, install scripts, low level software interaction, etc.

## Low research value, high technical value
56% of academics develop their own research software, **only 21%** of those have received training in software development.

## We need Research Software engineers!
This requires

- ▶ Recurrent funding
- ▶ Proper career prospects
- ▶ Flexibility over time and missions

## We're trying hard...

to convince our institutions, our governments, that **open-source, user driven development** is the right way, and **it needs sustainable funding to function.**

## We're trying hard...

to convince our institutions, our governments, that **open-source, user driven development** is the right way, and **it needs sustainable funding to function.**

## We're making progress

## We're trying hard...

to convince our institutions, our governments, that **open-source, user driven development** is the right way, and **it needs sustainable funding to function.**

## We're making progress

Open Digital Research Environment Toolkit for the Advancement of Mathematics

## We're trying hard...

to convince our institutions, our governments, that **open-source, user driven development** is the right way, and **it needs sustainable funding to function.**

## We're making progress

Open Digital Research Environment Toolkit for the Advancement of Mathematics



- ▶ Horizon 2020 European Research Infrastructures Work Programme

## We're trying hard...

to convince our institutions, our governments, that **open-source, user driven development** is the right way, and **it needs sustainable funding to function.**

## We're making progress

Open Digital Research Environment Toolkit for the Advancement of Mathematics



- ▶ Horizon 2020 European Research Infrastructures Work Programme
- ▶ A budget of 7.6 millions run by 18 partner institutions in 7 countries

## We're trying hard...

to convince our institutions, our governments, that **open-source, user driven development** is the right way, and **it needs sustainable funding to function.**

## We're making progress

Open Digital Research Environment Toolkit for the Advancement of Mathematics



- ▶ Horizon 2020 European Research Infrastructures Work Programme
- ▶ A budget of 7.6 millions run by 18 partner institutions in 7 countries
- ▶ started in 2015...

## We're trying hard...

to convince our institutions, our governments, that **open-source, user driven development** is the right way, and **it needs sustainable funding to function.**

## We're making progress

Open Digital Research Environment Toolkit for the Advancement of Mathematics



- ▶ Horizon 2020 European Research Infrastructures Work Programme
- ▶ A budget of 7.6 millions run by 18 partner institutions in 7 countries
- ▶ started in 2015...... ends in 2019.

## We're trying hard...

to convince our institutions, our governments, that **open-source, user driven development** is the right way, and **it needs sustainable funding to function.**

## We're making progress

Open Digital Research Environment Toolkit for the Advancement of Mathematics



- ▶ Horizon 2020 European Research Infrastructures Work Programme
- ▶ A budget of 7.6 millions run by 18 partner institutions in 7 countries
- ▶ started in 2015...... ends in 2019.

**What's next?**

# The challenges...

# The challenges...

**We need inclusivity and diversity**

# The challenges...

**We need inclusivity and diversity**
We need to develop software **for** everyone, **with** everyone.

# Some numbers...

# Some numbers...

In academia,

88% of Windows users and  93% of Linux users  **use** research software
41% of Windows users and  90% of Linux users  **develop** research software

# Some numbers...

In academia,

88% of Windows users and   93% of Linux users   **use** research software
41% of Windows users and   90% of Linux users   **develop** research software

Native Windows SageMath install: 2017 (thank you OpenDreamKit!)

# Some numbers...

In academia,

| | | |
|---|---|---|
| 88% of Windows users and | 93% of Linux users | **use** research software |
| 41% of Windows users and | 90% of Linux users | **develop** research software |

Native Windows SageMath install: 2017 (thank you OpenDreamKit!)

## Open-source is free to use but can be difficult to access

- ▶ Never forget the technical cost.
- ▶ Never forget the cultural aspects.

# Some other numbers...

# Some other numbers...

- 56% of researchers develop their own software,
- 70% of male researchers develop their own software, and only 30% of female researchers do so.
- 92% of researchers use research software (same number for men and women!)

# Some other numbers...

- 56% of researchers develop their own software,
- 70% of male researchers develop their own software, and only 30% of female researchers do so.
- 92% of researchers use research software (same number for men and women!)

Lack of training? Lack of confidence? **What can we do?**

# Support women coders and women initiative

# Check your privileges

- How did you get where you are?

# Check your privileges

- ▶ How did you get where you are?
- ▶ When did you have your first computer?

# Check your privileges

- How did you get where you are?
- When did you have your first computer?
- When did you write your first line of code?

# Check your privileges

- How did you get where you are?
- When did you have your first computer?
- When did you write your first line of code?
- What field did you study?

# Check your privileges

- ▶ How did you get where you are?
- ▶ When did you have your first computer?
- ▶ When did you write your first line of code?
- ▶ What field did you study?
- ▶ When did you first hear about open-source?

# Check your privileges

- ▶ How did you get where you are?
- ▶ When did you have your first computer?
- ▶ When did you write your first line of code?
- ▶ What field did you study?
- ▶ When did you first hear about open-source?
- ▶ What country are you from?

# Check your privileges

- ▶ How did you get where you are?
- ▶ When did you have your first computer?
- ▶ When did you write your first line of code?
- ▶ What field did you study?
- ▶ When did you first hear about open-source?
- ▶ What country are you from?
- ▶ Did your parents go to university?

# Check your privileges

- ▶ How did you get where you are?
- ▶ When did you have your first computer?
- ▶ When did you write your first line of code?
- ▶ What field did you study?
- ▶ When did you first hear about open-source?
- ▶ What country are you from?
- ▶ Did your parents go to university?
- ▶ What is the color of your skin?

# Check your privileges

- How did you get where you are?
- When did you have your first computer?
- When did you write your first line of code?
- What field did you study?
- When did you first hear about open-source?
- What country are you from?
- Did your parents go to university?
- What is the color of your skin?
- Are you straight? Non disabled? Cis-gendered?

Remember…
We want software **for everybody, by everybody**.
**You want people in, who are not like you.**

Remember…

We want software **for everybody, by everybody**.
**You want people in, who are not like you.**

Thank you!