Publications, bibliometry and cooperation indicators

daily experiences, issues and traps

www.cnrs.fr

Tashkent seminar : 23/24 april 2013

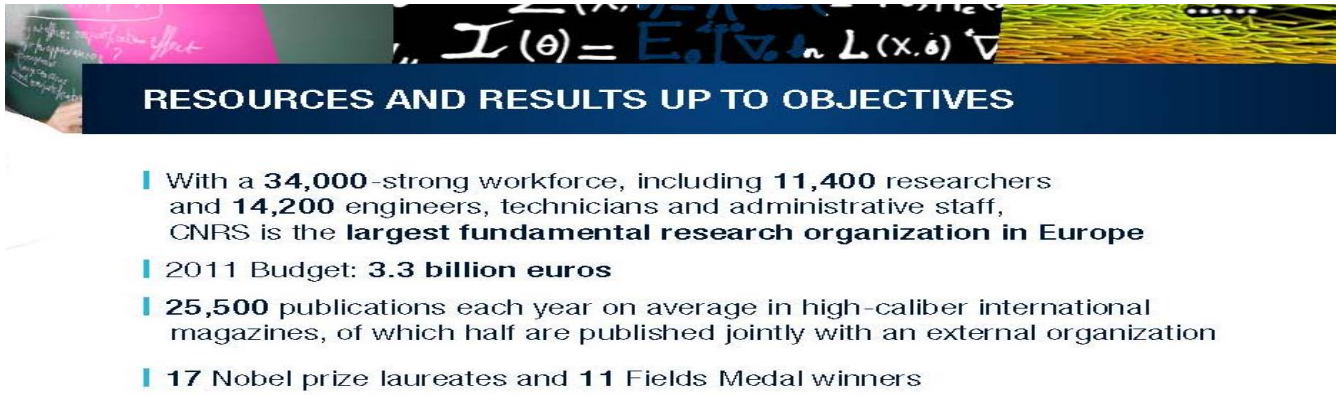
A brief presentation

What is CNRS ?

-CNRS is a scientific and technological public organization

-All fields of knowledge are covered through 10 specialized institute

-Nearly 1,100 research units, of which 94% are joint research units (with universities, (inter)national research institutions as well as private companies



- **Who am I ?**

- 15 years in CNRS (12 y. in the international cooperation office as an analyst and since 2010, in a sub-unit of the « territorial research organization' department

- Our sub-unit, namely the SAPPS, for « Supporting Service for Scientific Prospective & Policy » where I'm in charge of R&D's indicators for policy research issues and more precisely focused on international scientific cooperation (ISC) indicators

Our daily job, our daily issues :
To wonder very simple questions, very ancient questionings.

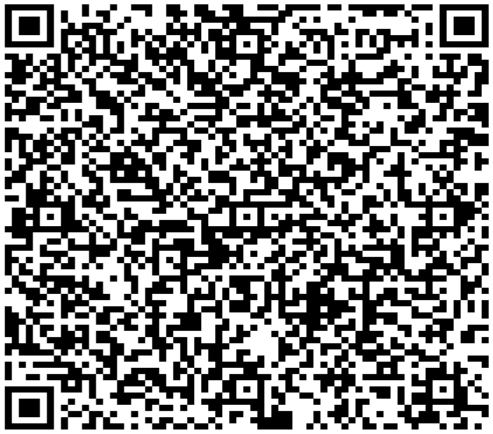
If one can answer the ancient «greek Pythian questioning»

Who, Why, How,
When, With whom… ?

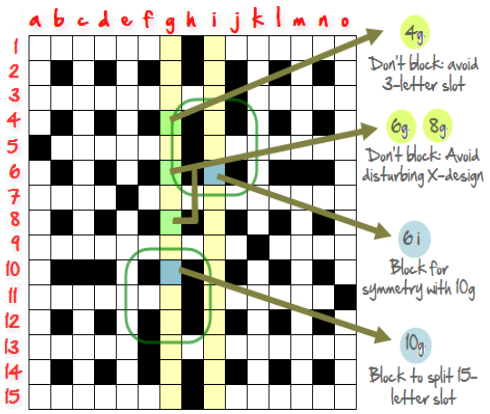
Then a big step
forward is done

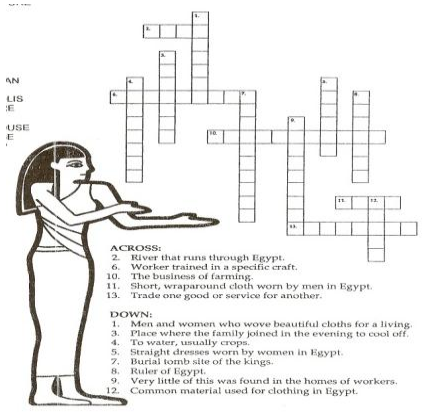
It could have been the end of my presentation, but sadly, the answer is sometimes …

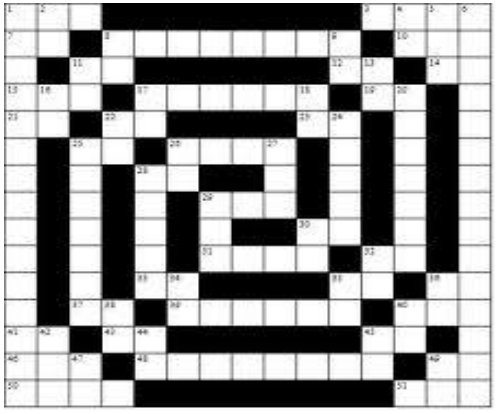
Confusing…



The main part of the job (its difficulties and its intersting side) is to put in order and organize this chaos







PLAN

- Why and How ?
- Who for and Whom with ?
- Traps and pitfalls
- Next steps
- Conclusion

Why and How ?

- 1- 1st step

- We (CNRS) assume that before all, we need to characterize ourself

- 2- 2nd step

- We (CNRS) suppose that once 1st step is done we Have to be **COMPARABLE**

Why and How ? …

1 - Characterization

We use worldwide databases for bibliometry (SCI and/or WoS)
because it's a **RELIABLE** output of research

But, it has to be **SUITABLE** for us and our environment (i.e. ERA)

These 3 words **Comparability – Reliability – Suitability**
are very important to assert we can use and spread our figures

Why and How ? …

To be reliable :

•We use bibliometric databasis for we have the capability since 15 years to produce studies through a long period of time

•We have to clean the data, especially the affiliation of authors

As CNRS' labs are mostly joint units with universities, 25/30% of them don't mention their CNRS' affiliation.

It's then a big added-value to clean and improve the information on the data

Why and How ? …

To be suitable :

One have to distinguish

- bibliometric databasisn (SCI-Wos …)
- production databasis (arXiv, Hal for instance)

- The former allow us to be comparable, reliable and suitable on a long period
- The latter sketches a production at a given time for an entity (i.e. Ribac in CNRS)
- No SHS (Social & Humanities Sciences) in our bibliometric report (different idiosynchrasy in dissemination of knowledge –books, books-chapter, conference…)

Why and How ? …

To be comparable :

•We operate in a world context, with partners (not competitors)

•A usual issue is to make worlwide comparable indicators and to keep national/institutionnal specificity

•More than this, one of the main worry (for all research institutions and/or universities) is to be **VISIBLE**

•**To be quantified (counted) and qualified we must be visible**

•This visibility is not that easy for an institution (i.e. for CNRS or MPG or CNR…)

The « publish or perish » motto, beyond its kinda funny-dictatorial sense, make us grow conscience that :

to *Publish* = shows a capability of doing, a « know-how »

or *Perish* = even if I did publish, what's the use if nobody knows (if I'm not recognised)

To « make know » where/who you are, be visible

Why and How ? …

This matrix shows our institute's (except ShS) commitment in scientific fields worldwide comparable (the matrix could be reversible)

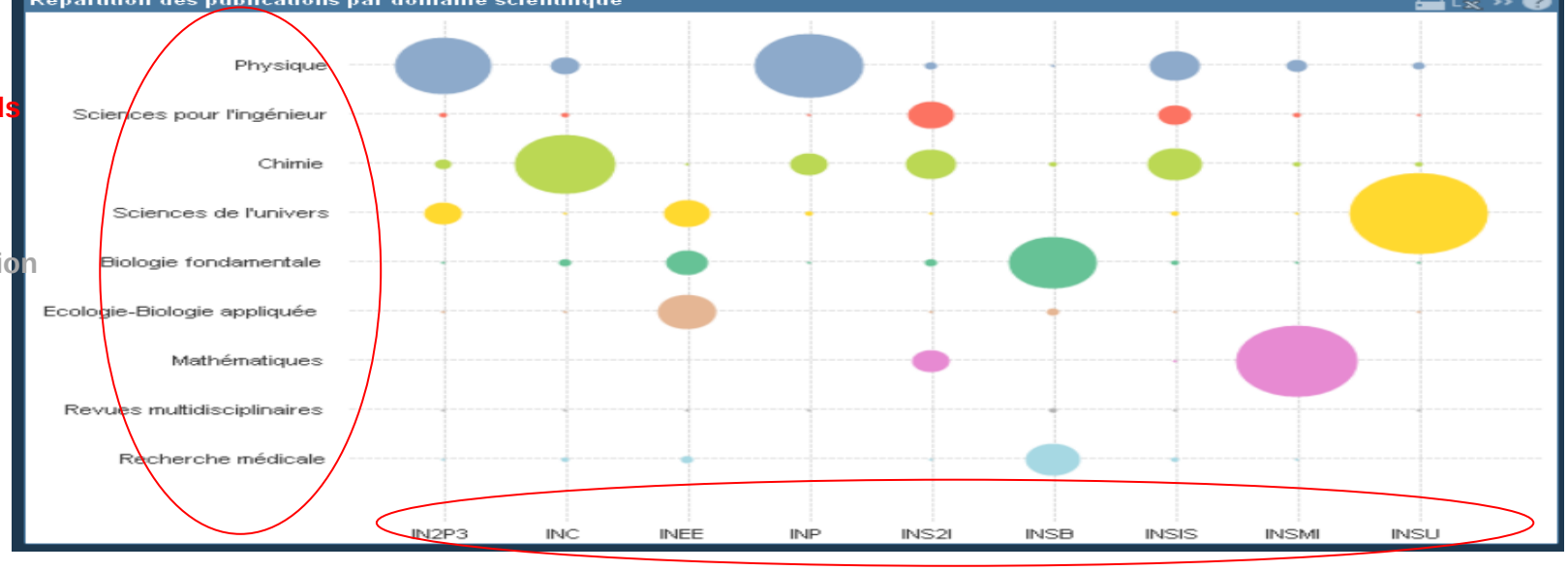
Nombre moyen annuel de publications du CNRS

29 019

Nombre moyen annuel de publications par institut

Institut	IN2P3	INC	INEE	INP	INS2I	INSB	INSIS	INSMI	INSU
Nombre moyen annuel de publications	1 218	6 993	2 206	4 643	721	5 130	4 685	1 788	4 467

Répartition des publications par domaine scientifique



Scient. fields for external mapping

Standart classification

CNRS' Institute for internal mapping

Internal classification

Why and How ? …

Other sources of identification

Beyond bibliometry, CNRS use other sources

2 examples

1 - The innovation and business relations' department (http://www.cnrs.fr/en/workingwith/innovation-business.htm)

deals with patents' issues.

It does have the same approach and questionning we do have in our unit.

A way to solve a part of the issue (*visiblity through comparability, reliability and suitability*)

has been to define some strategic innovation axis (*ASI in french*) in order to :

- generate partnership
- by presenting a thematic landscape of the CNRS

Why and How ? …

2 examples

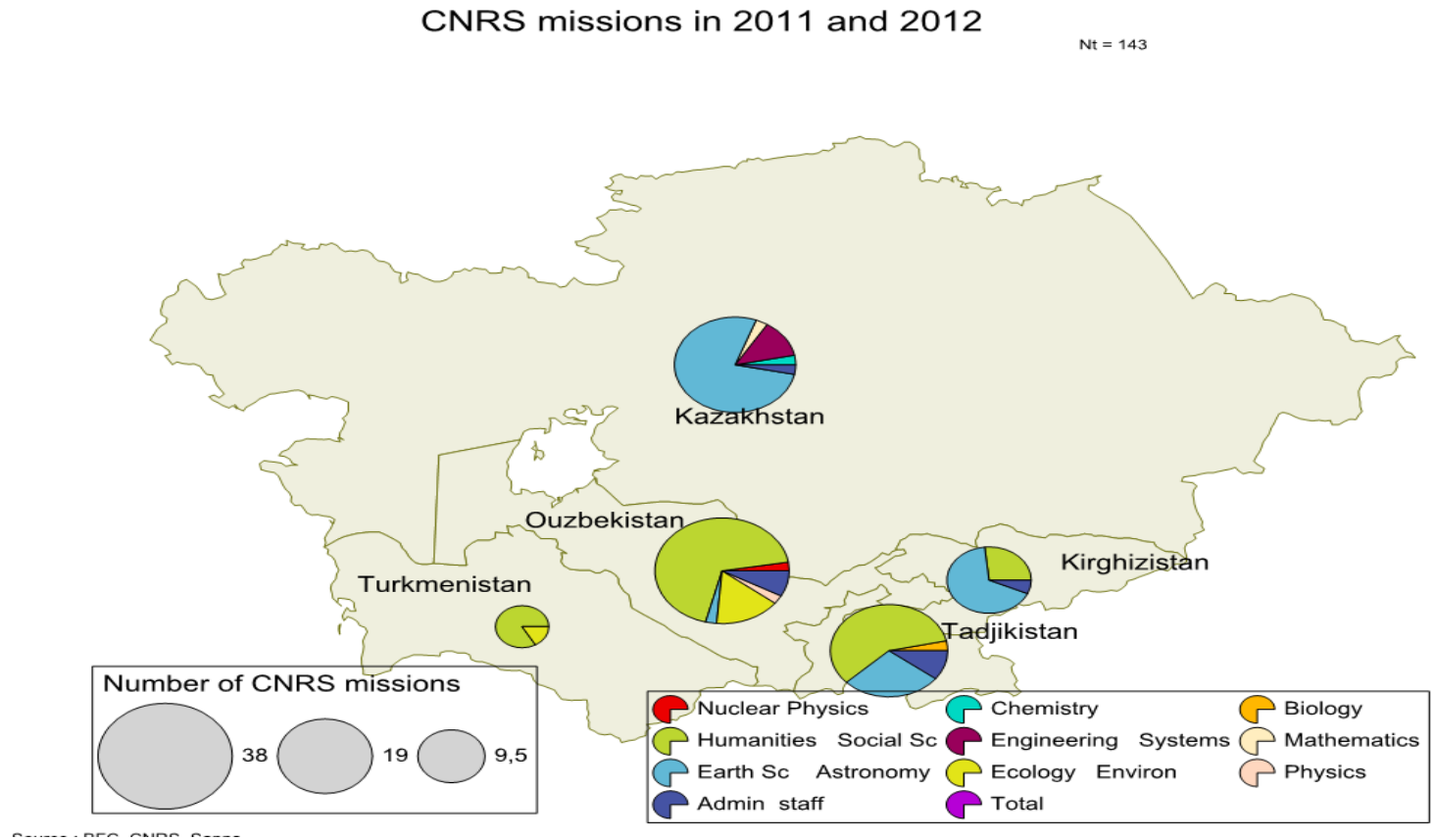
2 - The missions abroad (scientific trips / project abroad)

We collect every year the figures of missions made by researchers/engineers abroad to spread or collect scientific information or to work on common project.

More than 55,000 missions abroad (2/3 in Europe) made by CNRS

This amount of missions is big enough to be a suitable indicator of the CNRS activity abroad

Why and How ? …



Source: CNRS-BFC, SAPPS

- Who for and Whom with ?

We firstly exploit our figures for our direction's expectations, for policy makers
The french research landscape is deeply moving; universities and research organizations are gathering on bigger scientific sites (*cf*. Shangaï classification of universities)

We use to present indicators in the CNRS'annual report
(http://www.cnrs.fr/fr/organisme/docs/espacedoc/IndicateursChiffres-2011.pdf)

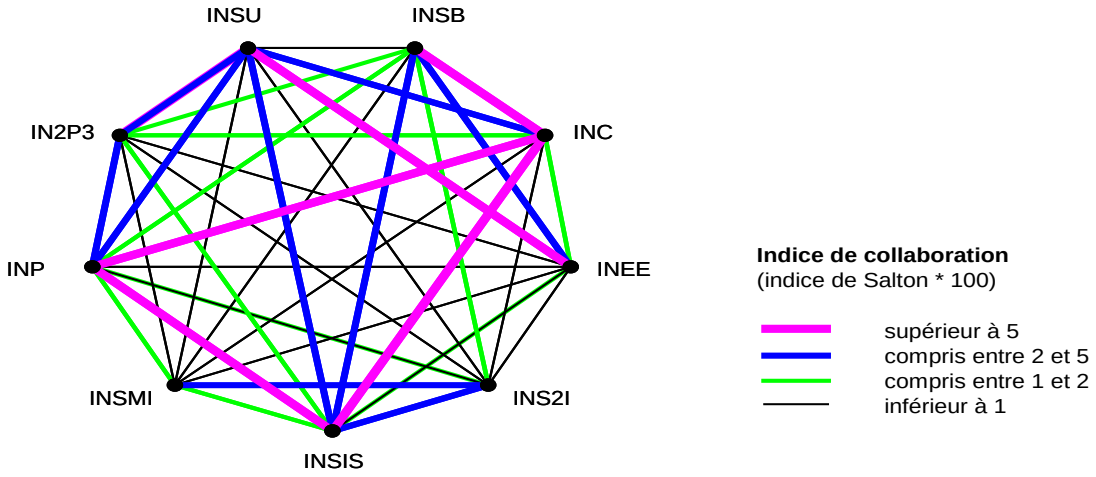
A reporting tool has been made, through Business Intelligent process(B.I.)
A good way to understand is to show the homepage of our tool, which shows our permanent worry of mapping our activity at different scale. This tool is a major gate for direction to respond to 99% of their daily questioning.

Traps and pitfall - 1

An overview of the CNRS'production : the interdisciplinarity in between CNRS'Institute

This graph presents another approach, another axis (point of view) to characterize the CNRS' scientific production

BUT, the data are indeed reliable, suitable BUT NOT comparable outside CNRS.



Source : données SCI (DVD Edition / Thomson Reuters) ; traitement CNRS / SAP2S

Traps and pitfall - 2

Due to specificities in the habits (*cf.* idiosyncrasy) of different fields, one can not compare 2 fields together.

For instance, high energy (or nuclear physicist) scientists use to share big facilities in common for their work. The publication from this common work will be counted as a copublication which actually don't turn out to a be bilaterral cooperation but a com-mitment using same facilities and a common participation to community of nuclear scientists.

In other words : a copublication doesn't always imply a bilateral cooperation

In other words : a copublication in field A « not-eq » a copublication in field B

Traps and pitfall - 3

We use to estimate the effort of international copublications by regarding the level of internatinal copublication / total number of publication

Country	% of international commintments
KZ-Kazakhstan	60,2%
KG-Kirghizistan	74,4%
UZ-Ouzbekistan	53,5%
TJ-Tadjikistan	61,1%
TM-Turkmenistan	100,0%
CNRS	57,0%
Vietnam	79,9%
Japan	**26,4%**
USA	30,3%

Sometimes figures are misleading and their interpretation can be equivocal

USA or Japan can certailly find in their domestic research area the potentiality needed.

Is Vietnam a very much open-minded country or is its scientific domestic potential not rich enough?

Traps and pitfall – (last and not least)

- We do not make any indicator of performance, we do not classify nor evaluate!

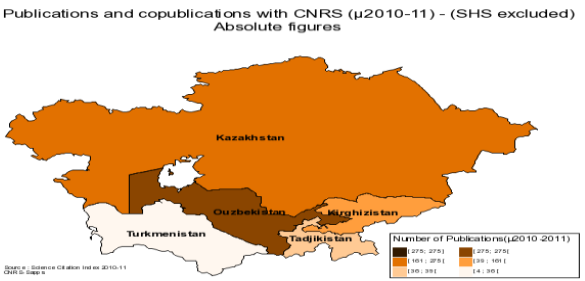
- We do not make « raw counts » as « performance ratio » such as nber of publis / reseacher

Please avoid this!

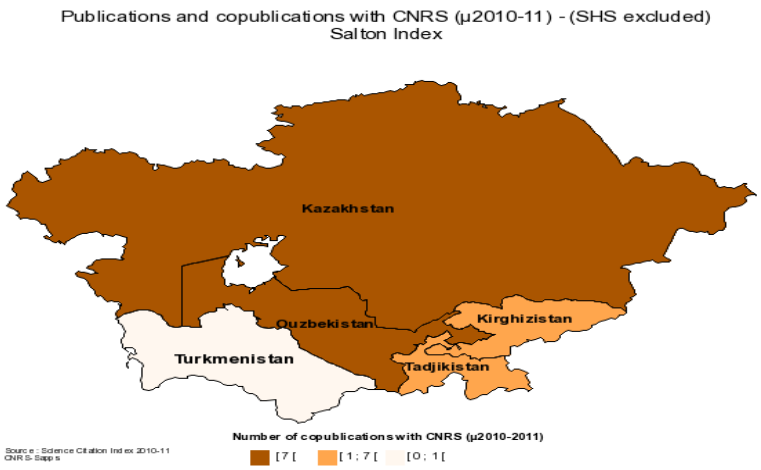
- We ‘d better use (when possible) relative parts rather than absolute number

Two different approaches of figures

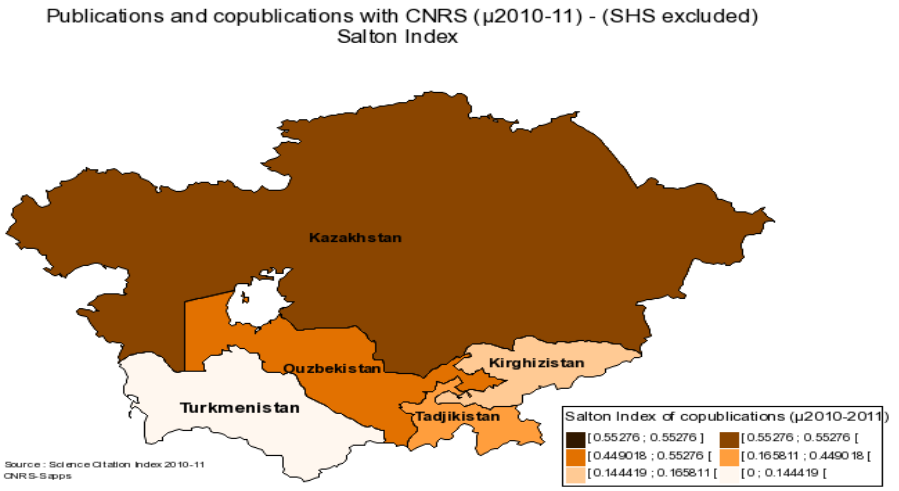
Number of publications



1-Absolute number of copub. with CNRS



2–Ratio qualifying the intensity of copub. with CNRS



Salton Index = Cxy/(Cx*Cy)^1/2 *100
0>S>1

Next Steps (Challenges and milestones)

I need your feedback to build a synthetic index of cooperatic

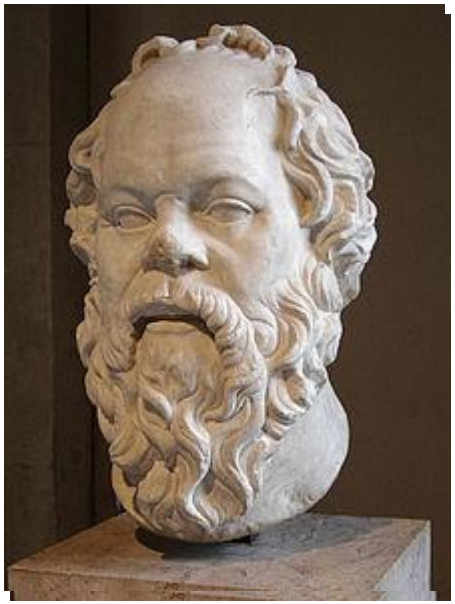
As the PNUD (United NationsDevelopt Programme) does for its development index, I want to try to build a cooperation index in science and technology mixing informal and formal data (copublications, location of conference-proceedings, missions or projects abroad, EU projects participation, international structures of cooperation

Futhermore, we are to include the proceedings in our analysis which will allow us to cover some more fields which are usually not very well covered by the SCI (let's say the computing sciences, mathematics …

Conclusion

To complete the circle of my presentation

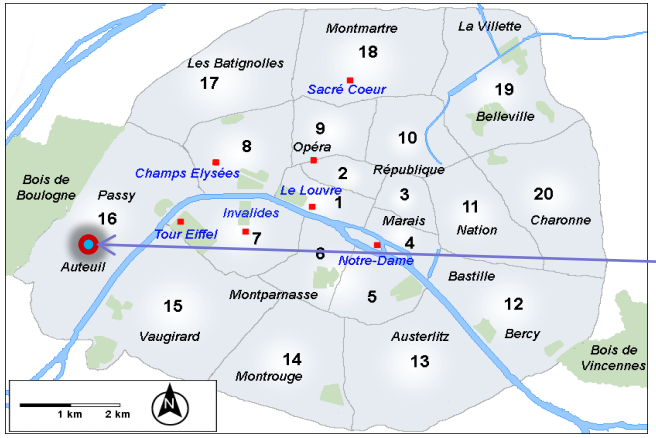
I would like to introduce an new greek contributor



Portrait of Socrate - Musée du Louvre
Roman bust (1st A.C.)

Know thyself (connais-toi toi-même)
(γνῶθι σεαυτόν - *gnōthi seauton)*

THANK YOU



CNRS-Headquarter - 3, rue Michel-Ange, 75016 Paris

Didier Journo

didier.journo@cnrs-dir.fr

+33-144 96 45 49

Annual Report : facts and figures

http://www.cnrs.fr/fr/organisme/docs/espacedoc/IndicateursChiffres-2011.pdf