

# Indices et classements : à qui profitent les chiffres ?

Michel Beaudouin-Lafon  
Université Paris-Sud

# Classements

- Classement de Shanghai
- Classement du Times
- Classement de Mines ParisTech
- Classement Webometrics

# Classement de Shanghai

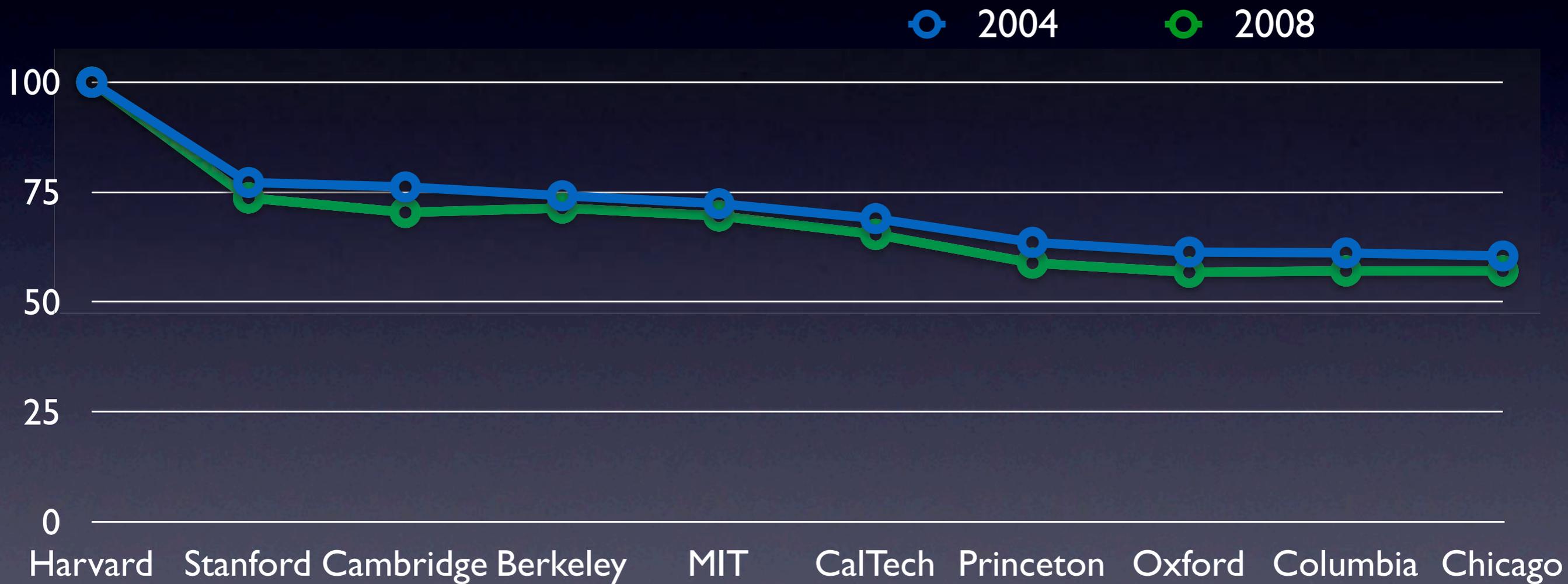
Criteria	Indicator	Code	Weight
Quality of Education	Alumni of an institution winning Nobel Prizes and Fields Medals	<i>Alumni</i>	10%
Quality of Faculty	Staff of an institution winning Nobel Prizes and Fields Medals	<i>Award</i>	20%
	Highly cited researchers in 21 broad subject categories	<i>HiCi</i>	20%
Research Output	Articles published in <i>Nature</i> and <i>Science</i>	<i>N&amp;S*</i>	20%
	Articles Indexed in Science Citation Index-Expanded and Social Science Citation Index	<i>SCI</i>	20%
Size of Institution	Academic performance with respect to the size of an institution	<i>Size</i>	10%
Total			100%

# Shanghai : 2004 vs 2008

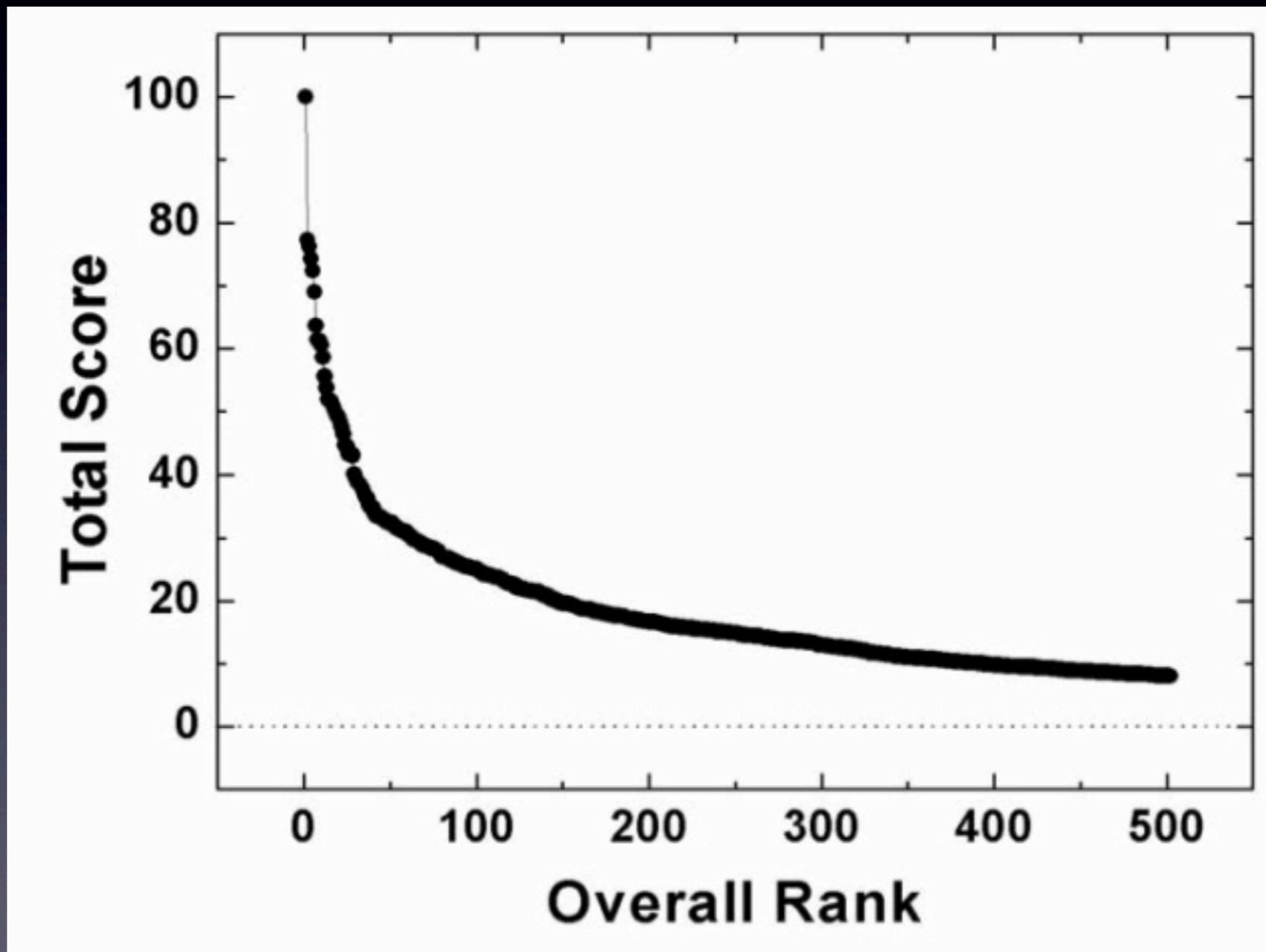
World Rank	Institution	Country
1	<a href="#">Harvard Univ</a>	USA
2	<a href="#">Stanford Univ</a>	USA
3	<a href="#">Univ Cambridge</a>	UK
4	<a href="#">Univ California - Berkeley</a>	USA
5	<a href="#">Massachusetts Inst Tech (MIT)</a>	USA
6	<a href="#">California Inst Tech</a>	USA
7	<a href="#">Princeton Univ</a>	USA
8	<a href="#">Univ Oxford</a>	UK
9	<a href="#">Columbia Univ</a>	USA
10	<a href="#">Univ Chicago</a>	USA

World Rank	Institution*	Region
1	<a href="#">Harvard Univ</a>	Americas
2	<a href="#">Stanford Univ</a>	Americas
3	<a href="#">Univ California - Berkeley</a>	Americas
4	<a href="#">Univ Cambridge</a>	Europe
5	<a href="#">Massachusetts Inst Tech (MIT)</a>	Americas
6	<a href="#">California Inst Tech</a>	Americas
7	<a href="#">Columbia Univ</a>	Americas
8	<a href="#">Princeton Univ</a>	Americas
9	<a href="#">Univ Chicago</a>	Americas
10	<a href="#">Univ Oxford</a>	Europe

# Shanghai : 2004 vs 2008



# Distribution des scores



# La France dans le classement de Shanghai

<b>Institution</b>	<b>Rang 2004 - 2008</b>	<b>Score 2004 - 2008</b>
Paris-6	41 - 42	32.6 - 33.1
Paris-11	48 - 49	31.2 - 32.1
Strasbourg	82 - ~125	25.7 - 22.3
ENS Ulm	85 - 73	25.4 - 27.7

# Classement du Times

World University Rankings 2008



## THE THE TOP 200 WORLD UNIVERSITIES

2008 RANK	2007 RANK	INSTITUTION	COUNTRY	PEER REVIEW SCORE	EMPLOYER REVIEW SCORE	STAFF/STUDENT SCORE	CITATIONS/STAFF SCORE	INTERNATIONAL STAFF SCORE	INTERNATIONAL STUDENTS SCORE	OVERALL SCORE
1	1	Harvard University	US	100	100	96	100	87	81	100
2	2=	Yale University	US	100	100	100	98	89	71	99.8
3	2=	University of Cambridge	UK	100	100	99	89	98	95	99.5
4	2=	University of Oxford	UK	100	100	100	85	96	96	98.9
5	7=	California Institute of Technology	US	100	74	98	100	100	93	98.6
6	5	Imperial College London	UK	99	100	100	83	98	100	98.4
7	9	University College London	UK	96	99	100	89	96	100	98.1
8	7=	University of Chicago	US	100	99	98	91	78	83	98.0
9	10	Massachusetts Institute of Technology	US	100	100	90	100	33	94	96.7
10	11	Columbia University	US	100	99	98	94	29	89	96.3

## INDICATORS TO EVALUATE THE OVERALL POSITION OF A UNIVERSITY

### **Academic Peer Review**

A global survey of academics asking respondents to identify universities they consider excellent in their own broad field of knowledge.

### **Employer Review**

A global survey of employers, with experience of recruiting from universities, asking respondents to identify universities they consider to be best at preparing their graduates for the workplace.

### **Citations per Faculty Member**

Using exported information from Scopus, this measure combines research productivity and quality, taking into account the scale of an institution.

### **Student Faculty Ratio**

In lieu of a globally available evaluation of teaching quality, this indicator is designed to serve as a widely available proxy for commitment to teaching.

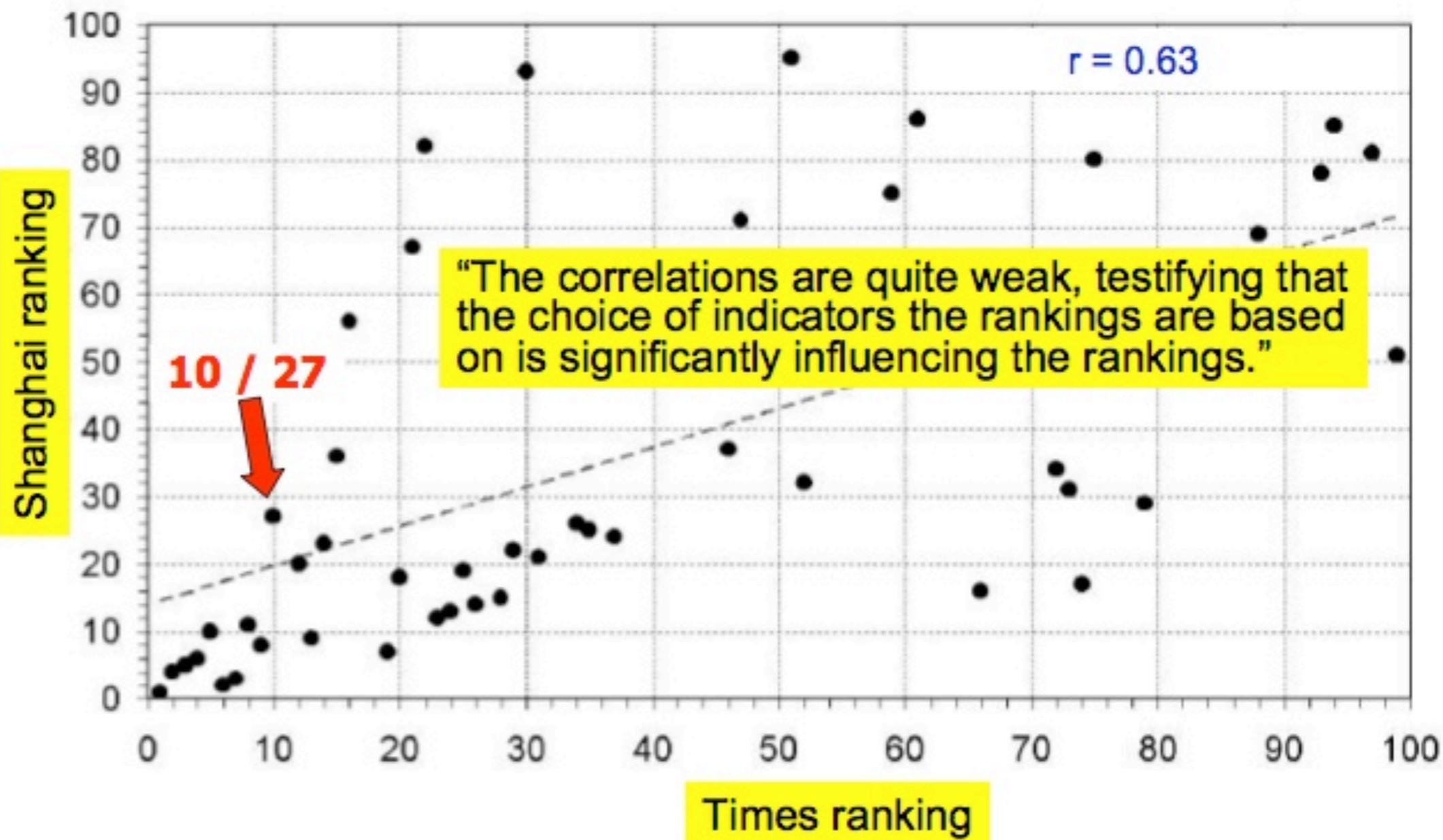
### **International Students**

The proportion of international students is used to evaluate not only a university's broad approach to internationalization but also to give prospective students feeling for an institution's commitment and facilities for students from overseas.

### **International Faculty**

This indicator supports the previous one in assessing a university's commitment to providing a global education for an increasingly global post-education environment.

# Corrélation Shanghai-Times



Tibor Braun, Ildikó Dióspalonyi, Erika Zádor, Sándor Zsindely:  
Journal gatekeepers indicator-based top universities of  
the world, of Europe and of 29 countries - A pilot study  
Scientometrics, Vol. 71, No. 2 (2007) 155-178

# La France dans le classement du Times

- 28 : ENS Ulm
- 34 : Ecole Polytechnique
- 140 : ENS Lyon
- 149 : Paris 6

# Comment améliorer le classement d'un pays ?

- Trouver le “bon” critère

Rang	Institution	Score
1	Tokyo Univ	100,0
2	Harvard Univ	89,0
3	Stanford Univ	57,0
4	Keio Univ	48,0
5	Univ Pennsylvania	39,0
6	Waseda Univ	36,0
7	HEC	35,0
8	Kyoto Univ	30,0
8	Univ Oxford	30,0
10	ENA	29,5
11	Inst for Study of Politics - Paris	26,5
12	Yale Univ	26,0
13	Massachusetts Inst Tech (MIT)	24,0
13	Osaka Univ	24,0
15	Ecole Polytechnique	23,5
16	Seoul Natl Univ	23,0
17	Univ Muenster	22,0
18	Fordham Univ	21,0
18	Pennsylvania State Univ - Univ Park	21,0
20	Duke Univ	20,0
20	Ecole Natl Super Mines - Paris	20,0
20	Columbia Univ	20,0
23	Chalmers Univ Tech	18,0
23	Chuo Univ	18,0
23	Univ Chicago	18,0
23	Univ Iowa	18,0
27	INSEAD	15,5

# Classement de Mines ParisTech

Nous avons donc choisi de retenir un critère simple, non déclaratif et vérifiable : **le nombre d'anciens élèves occupant le poste de n°1 exécutif (Chief Executive Officer ou équivalent) dans une des 500 plus grandes entreprises internationales**, à la date et avec les éléments (entreprises et dirigeants) du classement « **Fortune Global 500** » établi par le magazine Fortune en 2007 à partir du chiffre d'affaires publié par les entreprises mondiales.

Ce critère se veut être l'équivalent au niveau des entreprises du critère « anciens élèves ayant obtenu le Prix Nobel ou la médaille Fields » utilisé par le classement de l'Université Jiaotong à Shanghai, les nombres d'anciens concernés étant du même ordre de grandeur. Mais à la différence du classement de Shanghai, ce critère permet de mettre l'accent sur les performances des formations délivrées dans les établissements d'enseignement supérieur, plutôt que sur les performances de recherche de ces établissements.

# Classement Webometrics

90	ROYAL NETHERLANDS METEOROLOGICAL INSTITUTE		175	125	45	240
91	URBAN INSTITUTE		205	62	451	41
92	THOMAS JEFFERSON LAB NATIONAL ACCELERATOR FACILITY		98	198	28	170
93	ROYAL NETHERLANDS ACADEMY OF ARTS AND SCIENCES		117	133	243	183
94	JAPAN AEROSPACE EXPLORATION AGENCY		67	157	116	315
95	RUSSIAN ACADEMY OF SCIENCES		103	171	94	218
96	NYU MEDICAL CENTER NYU SCHOOL OF MEDICINE		104	119	169	364
97	UMR8623 LABORATOIRE DE RECHERCHE EN INFORMATIQUE		61	221	77	147
98	ENTE PER LE NUOVE TECNOLOGIE L'ENERGIA EL'AMBIENTE		96	201	60	191
99	ISTITUTO NAZIONALE DI STATISTICA		105	96	215	445
100	INSTITUT PASTEUR		126	132	213	262

1 | 101 | 201 | 301 | 401 | Institutes 1 to 100 of 500

2007

# Webometrics

## WEBOMETRICS RANK

VISIBILITY (external inlinks) <b>50%</b>	SIZE (web pages) <b>20%</b>
	RICH FILES <b>15%</b>
	SCHOLAR <b>15%</b>

2007

**Webometrics Ranking of World Universities**  
January '07

home world countries world rank european rank latin american rank spanish rank

> home > top 500 R&D

**Data**

- Top 3000 Universities
- Premier League
- Top USA & Canada
- Top Latin America
- Top Europe
- Top Asia
- Top Middle East
- Top Oceania
- Top Africa
- Top 500 R&D Institutes
- Research Councils
- Distribution by Country
- Specials
- Best Practices
- PDF Library

**Comparative Analysis**

- Productivity
- Visibility
- Impact
- Methodology

**Catalogue**

- Universities by country
- R&D Centres by country

**Information**

- Methodology

**Top 500 R&D**  
First | Previous | Next | Last | Institutes 1 to 100 of 500

WORLD RANK	INSTITUTE	COUNTRY	POSITION			
			SIZE	VISIBILITY	RICH FILES	SCHOLAR
1	NATIONAL INSTITUTES OF HEALTH		2	2	21	1
2	NATIONAL AERONAUTICS AND SPACE ADMINISTRATION		3	4	3	10
3	NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION		5	5	8	23
4	US GEOLOGICAL SURVEY		8	9	14	21
5	US ENVIRONMENTAL PROTECTION AGENCY		7	8	36	36
5	MAX PLANCK GESELLSCHAFT		11	16	12	20
7	CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS		15	18	17	8
8	EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH		6	24	1	19
9	WORLD WIDE WEB CONSORTIUM		1	1	53	73
10	CENTERS FOR DISEASE CONTROL AND PREVENTION		31	7	29	16
11	NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY		21	13	18	29
12	JET PROPULSION LABORATORY		24	14	26	15
13	INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE		19	23	24	11
14	COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANIZATION		23	37	16	6
15	NATIONAL LIBRARY OF MEDICINE		4	3	199	2

# Webometrics 2009

Top 4000 Universities						
First   Previous   Next   Last   Universities 1 to 50 of 4000						
<u>WORLD RANK</u>	<u>UNIVERSITY</u>	COUNTRY	POSITION			
			SIZE	VISIBILITY	RICH FILES	SCHOLAR
1	Massachusetts Institute of Technology		1	3	2	6
2	Stanford University		2	2	3	12
3	Harvard University ***		3	1	17	1
4	University of California Berkeley		6	4	5	24
5	Cornell University		4	5	8	37
6	University of Michigan		10	6	15	22
7	California Institute of Technology ***		8	8	21	17
8	University of Minnesota		9	16	4	19
9	University of Illinois Urbana Champaign *		14	10	6	38
10	University of Texas Austin		11	9	10	45

- Que signifient ces classements ?
  - A quoi servent ces classements ?
  - A qui servent ces classements ?
- => aux media ... et parfois aux politiques

- Problèmes de fiabilité
  - Qualité des sources
  - Précision des mesures
- Problèmes de validité
  - Pertinence des mesures
  - La mesure perturbe le système

# Indices bibliométriques

- h-index
- Facteur d'impact



J. Hirsch

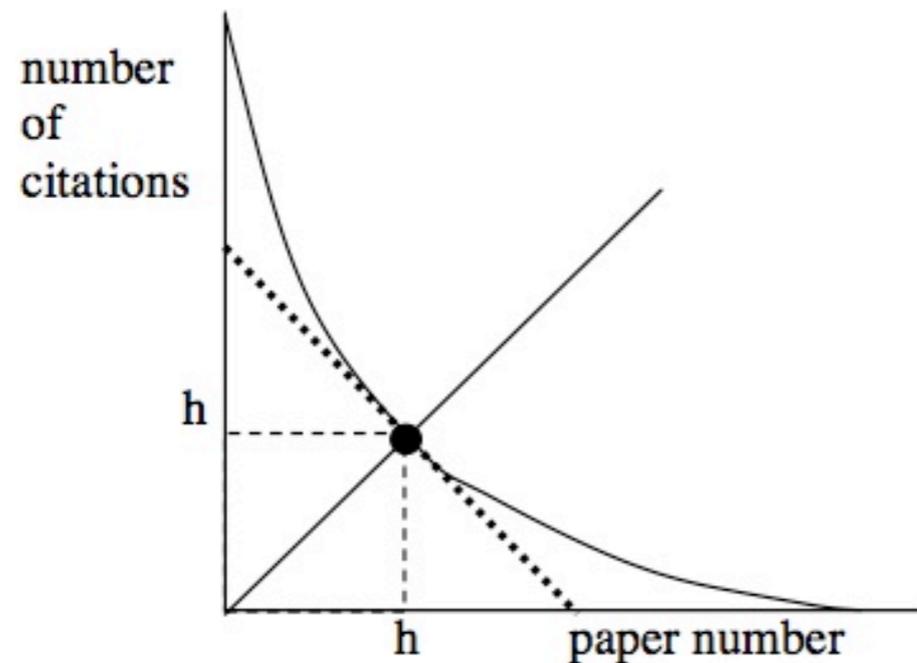


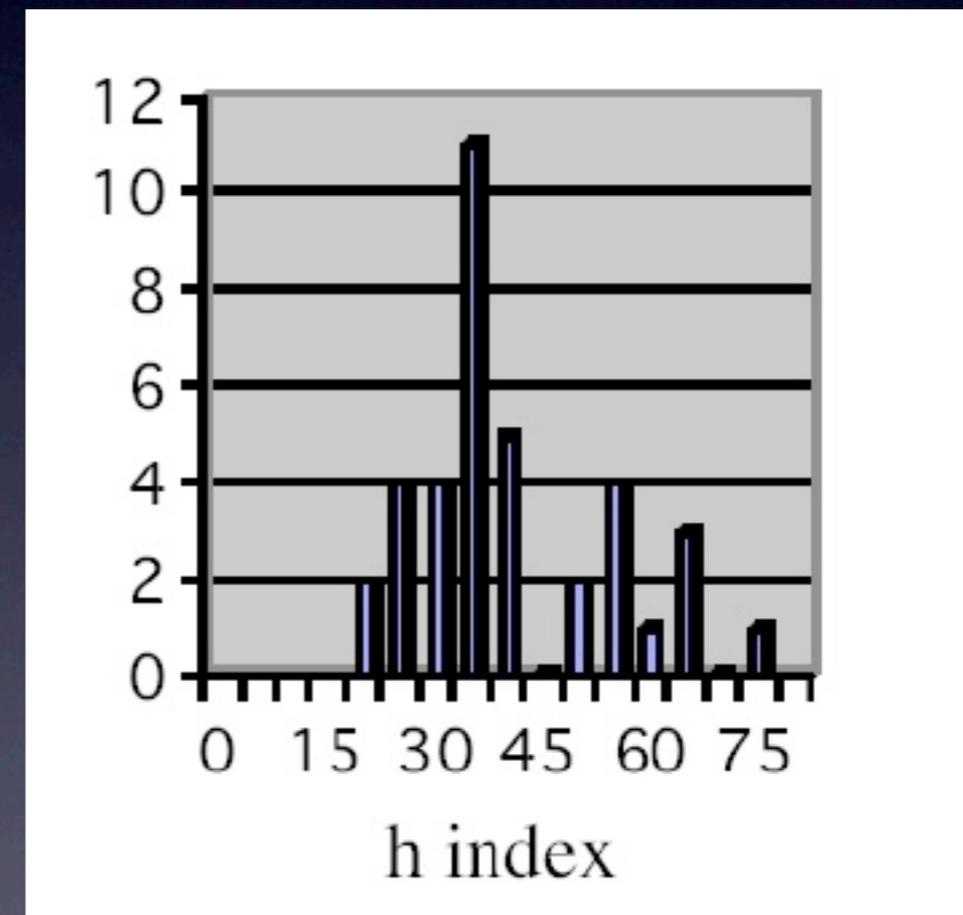
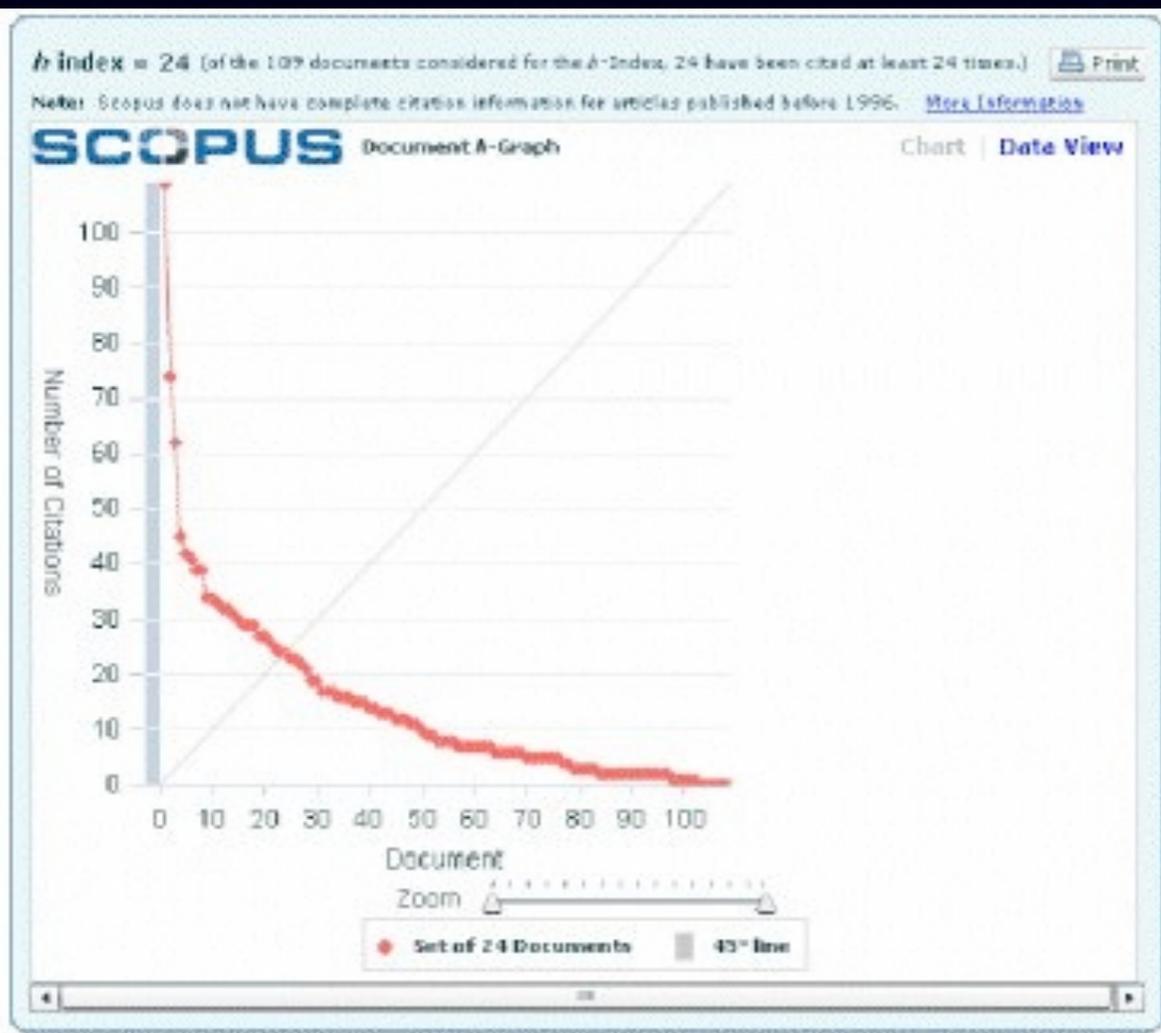
FIG. 1: The intersection of the 45 degree line with the curve giving the number of citations versus the paper number gives  $h$ . The total number of citations is the area under the curve. Assuming the second derivative is non-negative everywhere, the minimum area is given by the distribution indicated by the dotted line, yielding  $a=2$  in Eq. 1.

From inspection of the citation records of many physicists I conclude:

- (1) A value  $m \sim 1$ , i.e. an  $h$  index of 20 after 20 years of scientific activity, characterizes a successful scientist.
- (2) A value  $m \sim 2$ , i.e. an  $h$ -index of 40 after 20 years of scientific activity, characterizes outstanding scientists, likely to be found only at the top universities or major research laboratories.
- (3) A value  $m \sim 3$  or higher, i.e. an  $h$ -index of 60 after 20 years, or 90 after 30 years, characterizes truly unique individuals.

In summary, I have proposed an easily computable index,  $h$ , which gives an estimate of the importance, significance and broad impact of a scientist's cumulative research contributions. I suggest that this index may provide a useful yardstick to compare different individuals competing for the same resource when an important evaluation criterion is scientific achievement, in an unbiased way.

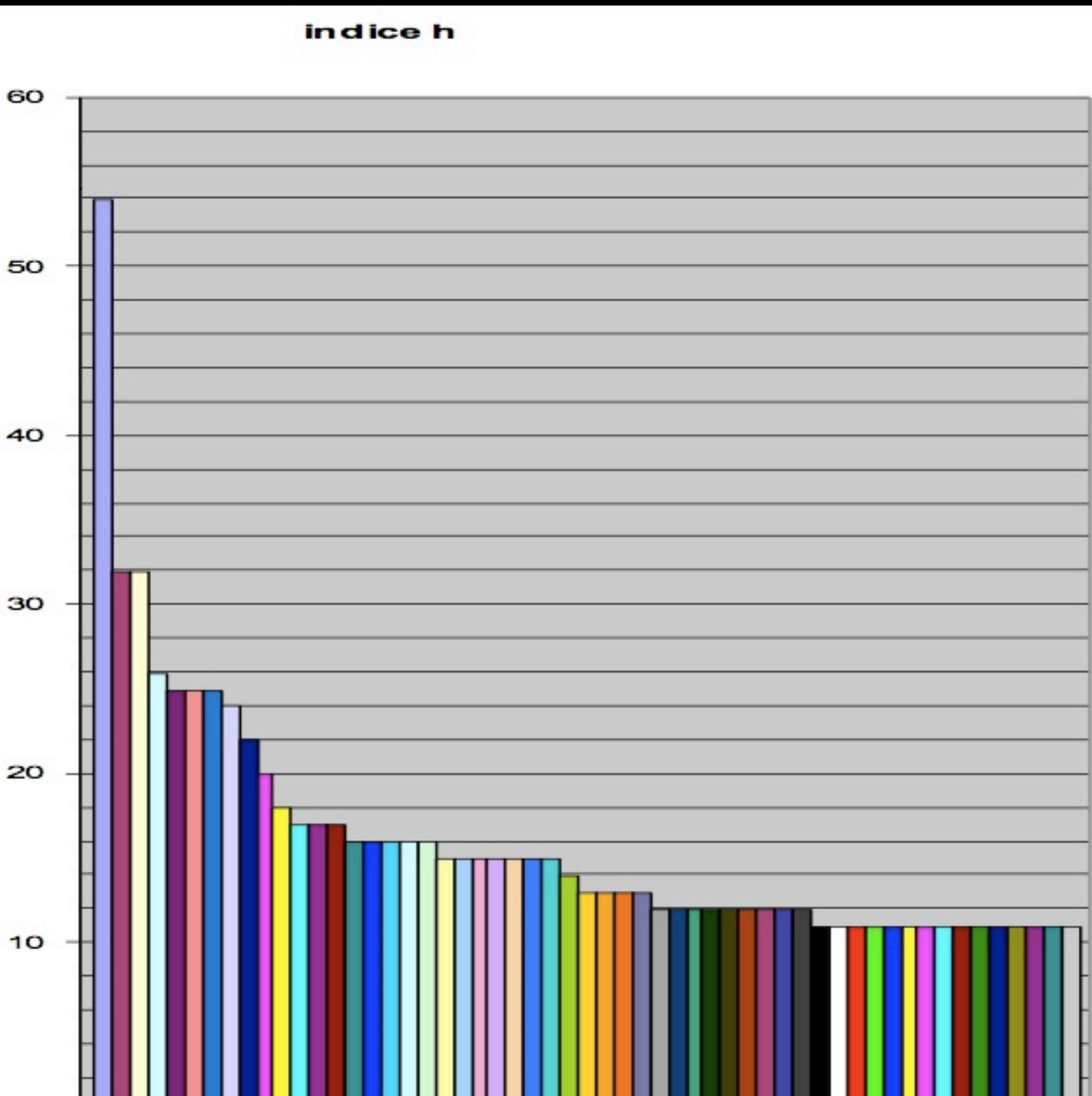
# Loi de puissance, Loi normale



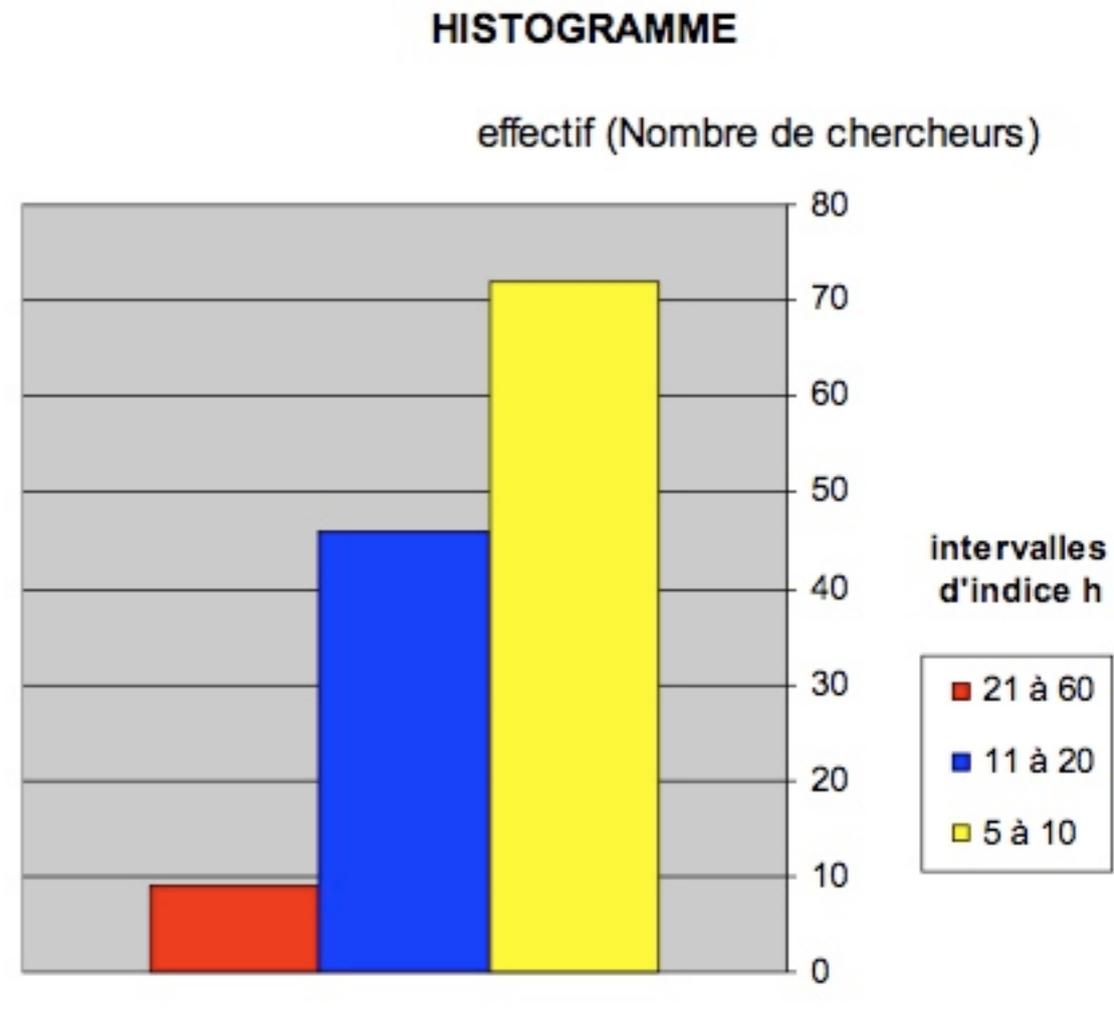
Scopus

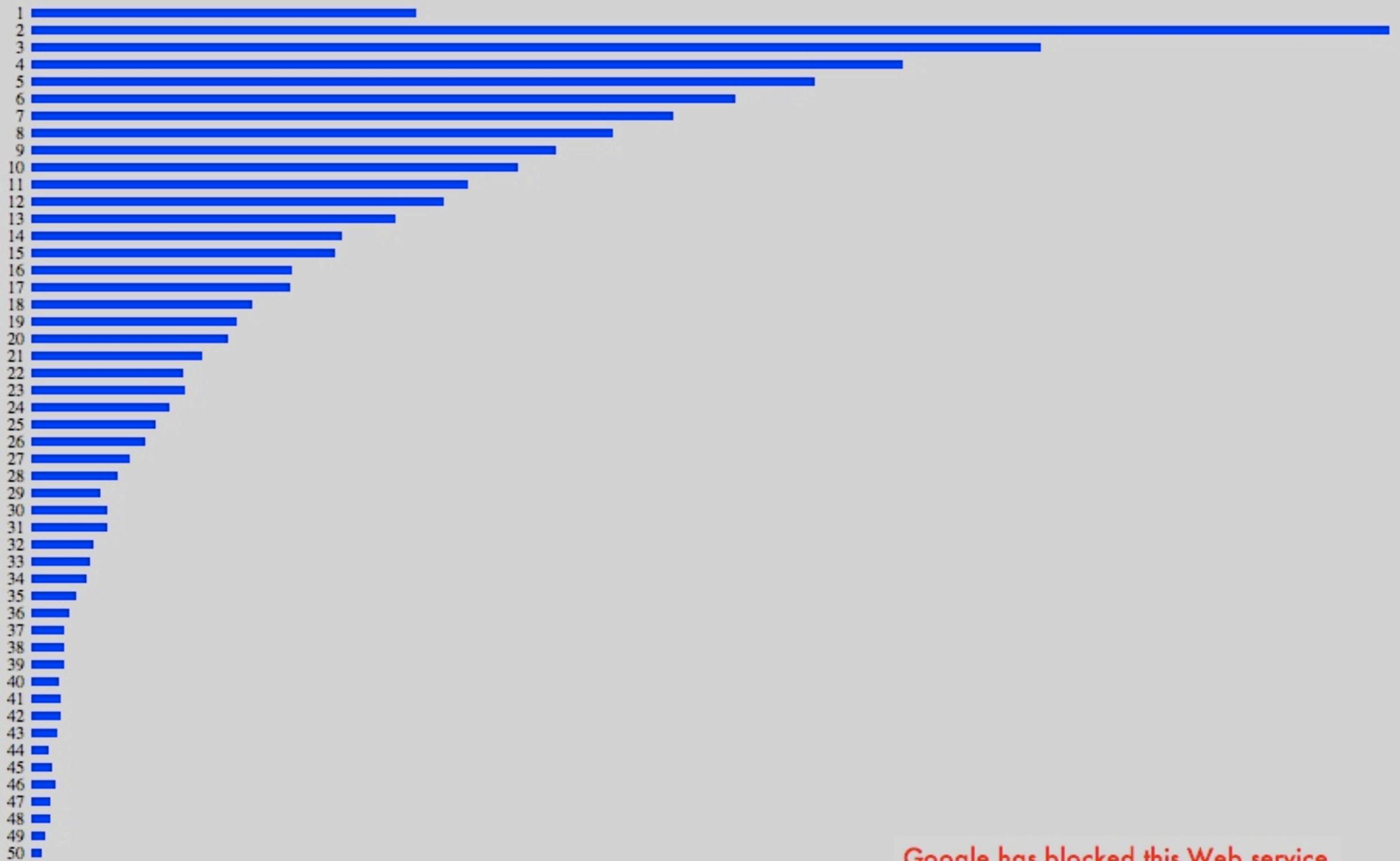
J. Hirsch - h-index des Prix Nobel

# Distribution des h-index



Digiteo





Year	Researcher	H Index <sup>2</sup>
2001	Stuart K. Card	44
2001	James D. Foley	28
2001	Morten Kyng	23
2001	Thomas P. Moran	33
2001	Donald A. Norman	49
2001	Judith S. Olson	28
2001	Ben Shneiderman	64
2002	William A. S. Buxton	15
2002	John M. Carroll	52
2002	Douglas C. Engelbart	11
2002	Sara Kiesler	39
2002	Thomas K. Landauer	38
2002	Lucy A. Suchman	10
2003	Thomas Green	39
2003	James D. Hollan	18
2003	Robert E. Kraut	34
2003	Gary M. Olson	29
2003	Peter G. Polson	21

# h-index

- Pratiques de publication très variables d'une (sous)-discipline à l'autre
  - Monotone croissant (même après la mort de l'auteur)
  - Sensible à la qualité des données
- ⇒ Peut servir *au mieux* à comparer des chercheurs de même ancienneté dans une même discipline

# Lois de puissance

- 20% des revues couvrent 80% des articles d'intérêt (loi de Bradford)
- 20% des articles représentent 80% des citations
- ...  
*=> les outils statistiques naïfs (moyenne) ne veulent rien dire avec de telles distributions*

# Facteur d'impact

Nombre de citations d'articles publiés dans  
les années N-1 et N-2 et cités l'année N

---

Nombre d'articles publiés les années N-1 et N-2

(on ne considère que les citations  
dans des revues identifiées)

# ISI Journal Citation Report

Journals 1 - 20 (of 84)

Page 1 of 5

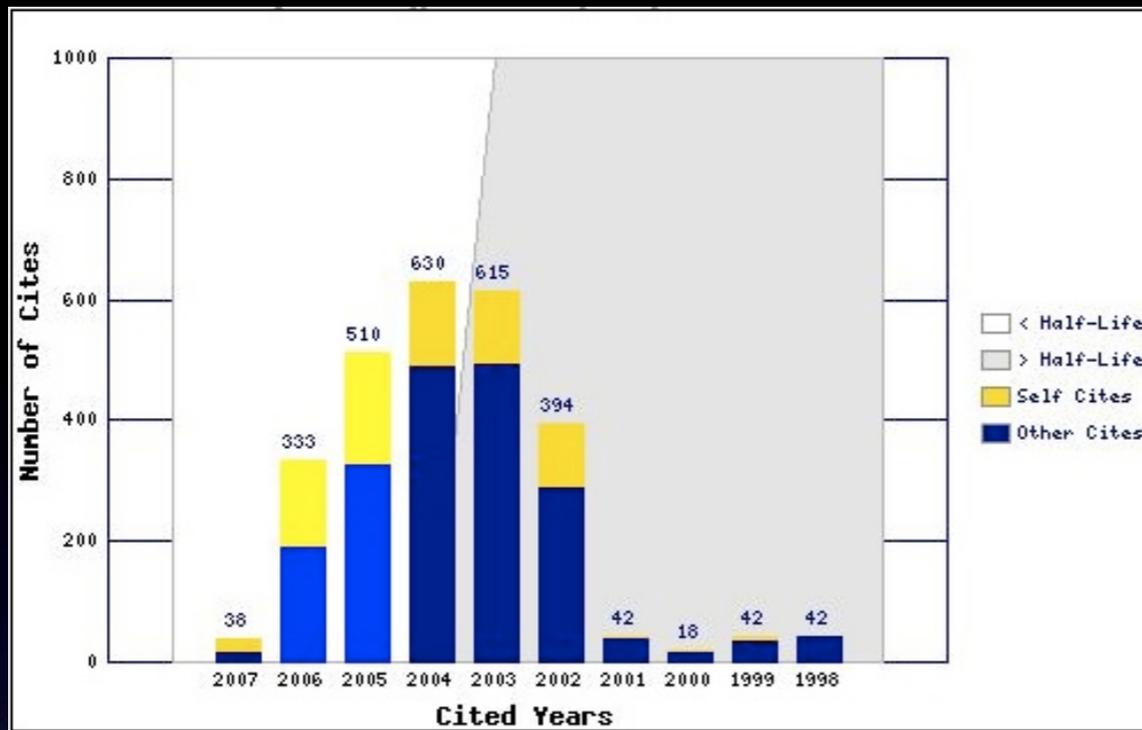
MARK ALL

UPDATE MARKED LIST

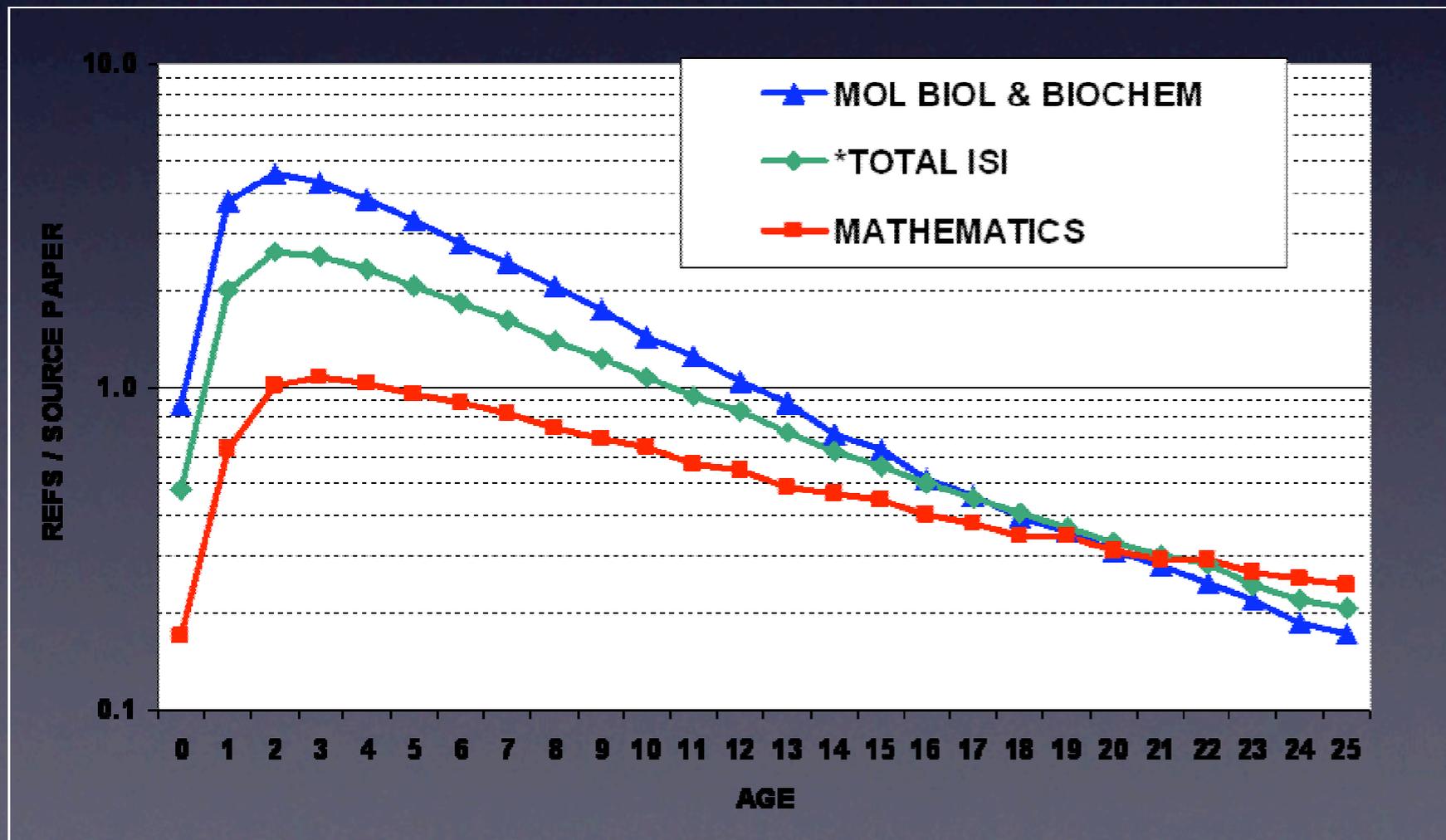
Ranking is based on your journal and sort selections.

Mark	Rank	Abbreviated Journal Title <i>(linked to journal information)</i>	ISSN	JCR Data <sup>i</sup>						Eigenfactor™ Metrics <sup>i</sup>	
				Total Cites	Impact Factor	5-Year Impact Factor	Immediacy Index	Articles	Cited Half-life	Eigenfactor™ Score	Article Influence™ Score
<input type="checkbox"/>	1	<a href="#">ACM T GRAPHIC</a>	0730-0301	3102	3.413	4.683	0.297	128	4.1	0.02259	1.833
<input type="checkbox"/>	2	<a href="#">J WEB SEMANT</a>	1570-8268	328	3.410		0.500	22	3.2	0.00147	
<input type="checkbox"/>	3	<a href="#">J ACM</a>	0004-5411	4894	3.136	3.925	0.312	32	>10.0	0.00962	2.587
<input type="checkbox"/>	4	<a href="#">ACM T SOFTW ENG METH</a>	1049-331X	534	2.792	3.667	0.417	12	7.3	0.00150	1.077
<input type="checkbox"/>	5	<a href="#">REAL-TIME IMAGING</a>	1077-2014	284	2.270	1.287			4.6	0.00187	0.537

“Half of the journal's cited items were published more recently than the half-life”



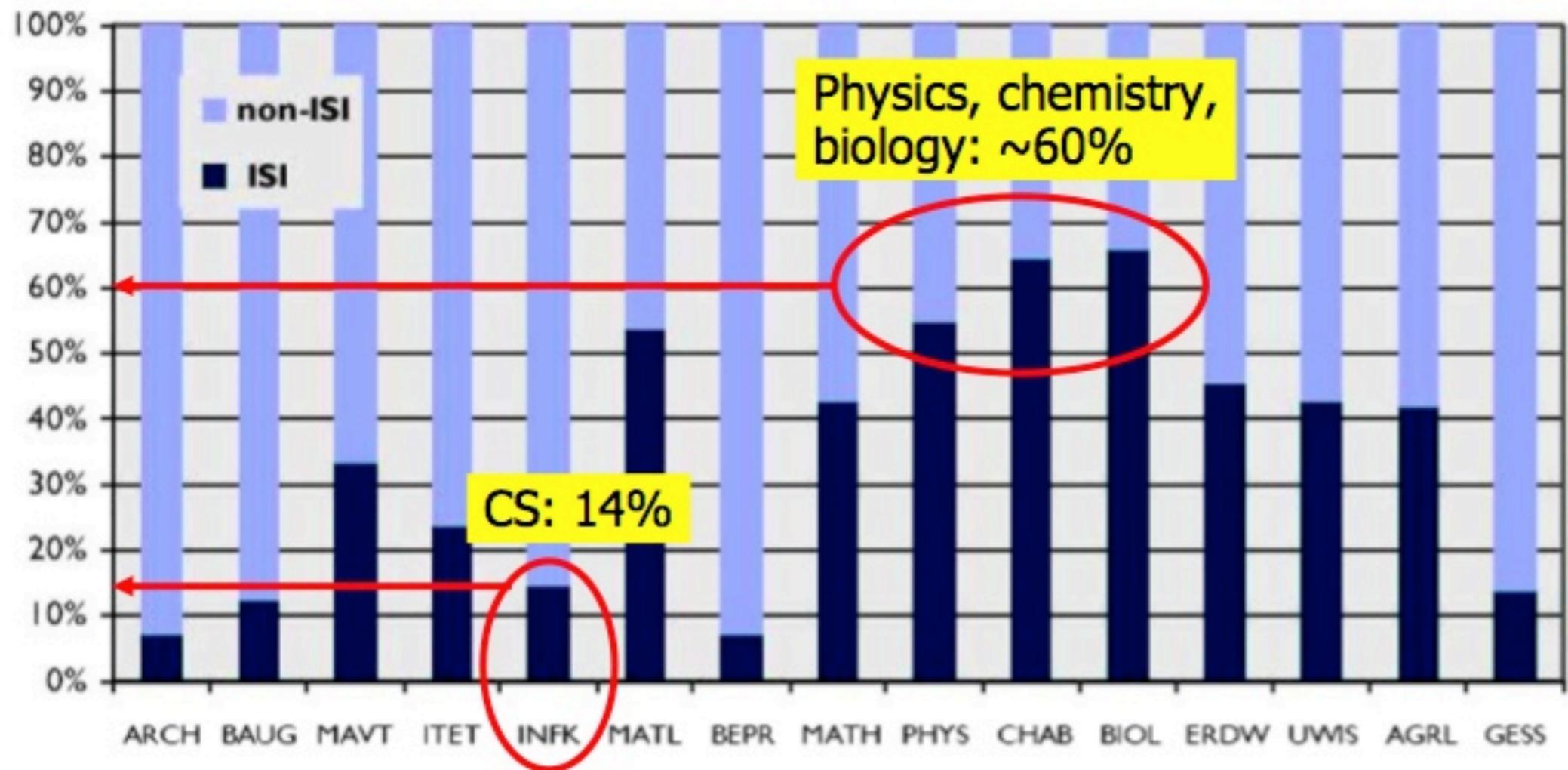
ISI JCR



Age des citations (Elsevier)

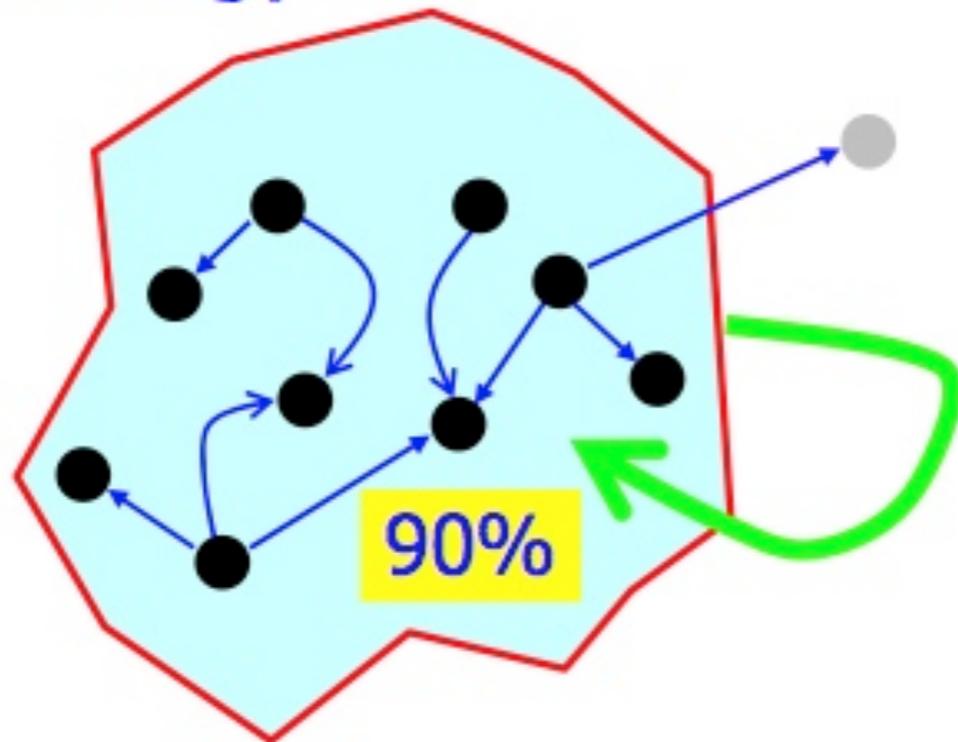
# ISI : problème de couverture

Analysis of all publications from **ETH Zurich** in 2003:

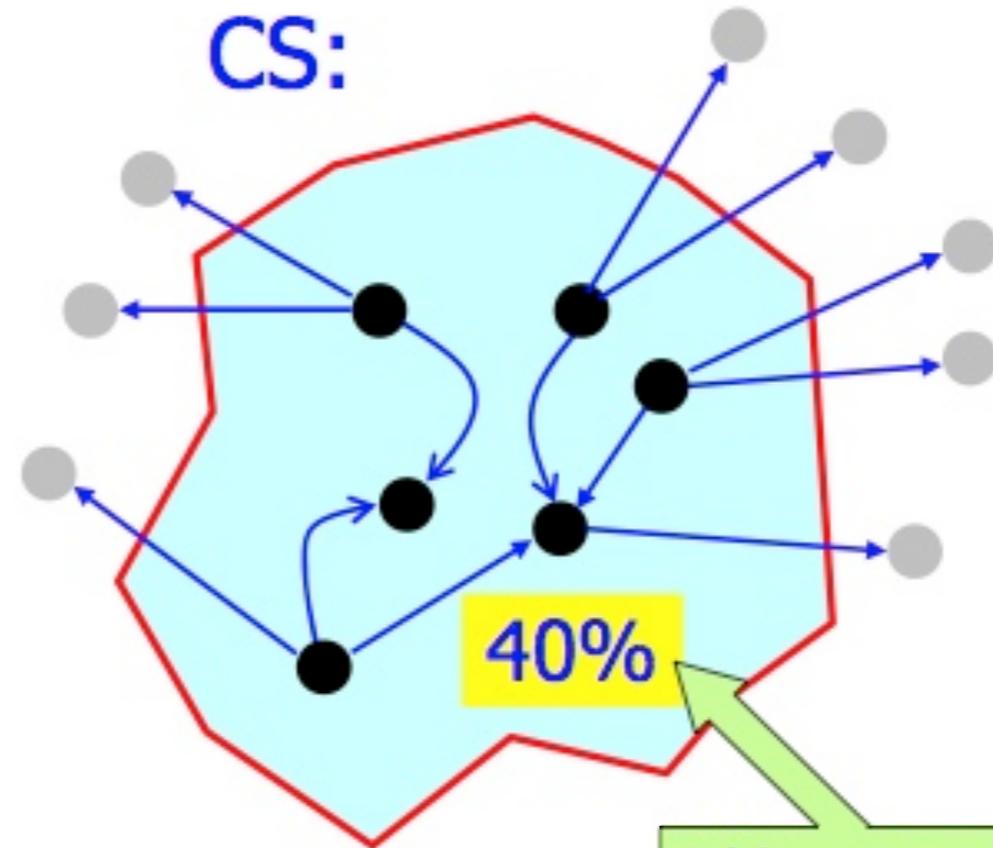


# ISI : problème de pertinence

Biology:



CS:



Based on the 2003 ETH Zurich publication pool

Adding proceedings from ACM, IEEE-CS, and LNCS yields 51%

# Pourquoi

- ISI n'indexe pas les conférences, qui sont le moyen majeur de publication en informatique, et qui ont un processus de sélection aussi sélectif que les journaux
- ISI fait preuve de mauvaise volonté pour indexer certaines revues, notamment celles éditées par les sociétés savantes

# Editeurs commerciaux

- La publication scientifique est un business, qui fonctionne grâce à une main-d'œuvre gratuite : les chercheurs eux-mêmes
- Le facteur d'impact est un argument marketing pour les éditeurs commerciaux
- Les clients sont les institutions qui achètent des abonnements

# Evolution des modèles de publications

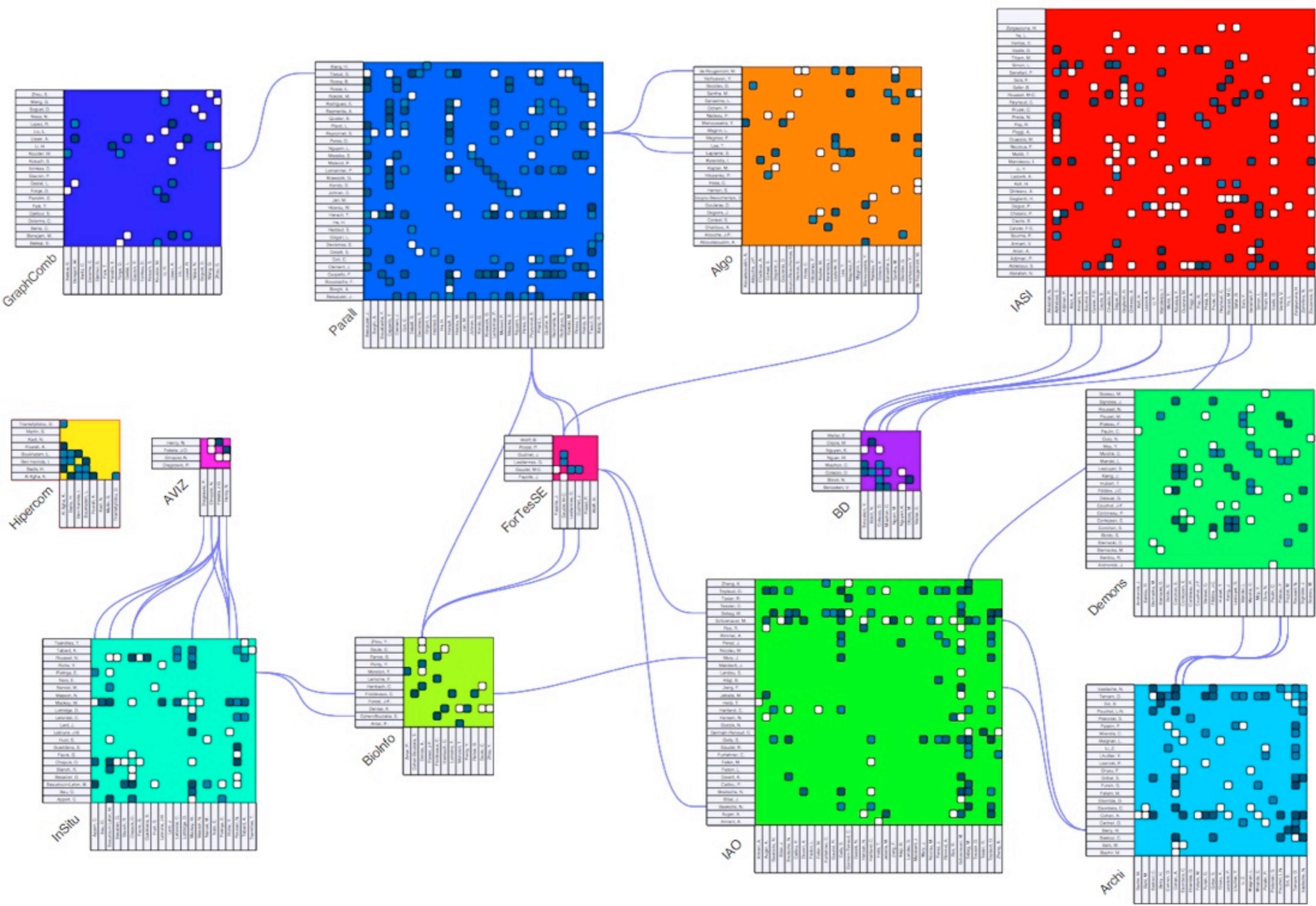
- Modèle Open Publishing
- Qui va rémunérer les éditeurs ?  
les auteurs => modèle du “author-pay”
- 1500€ à 8000€ par article !
- Les éditeurs commerciaux sont prêts à changer de modèle

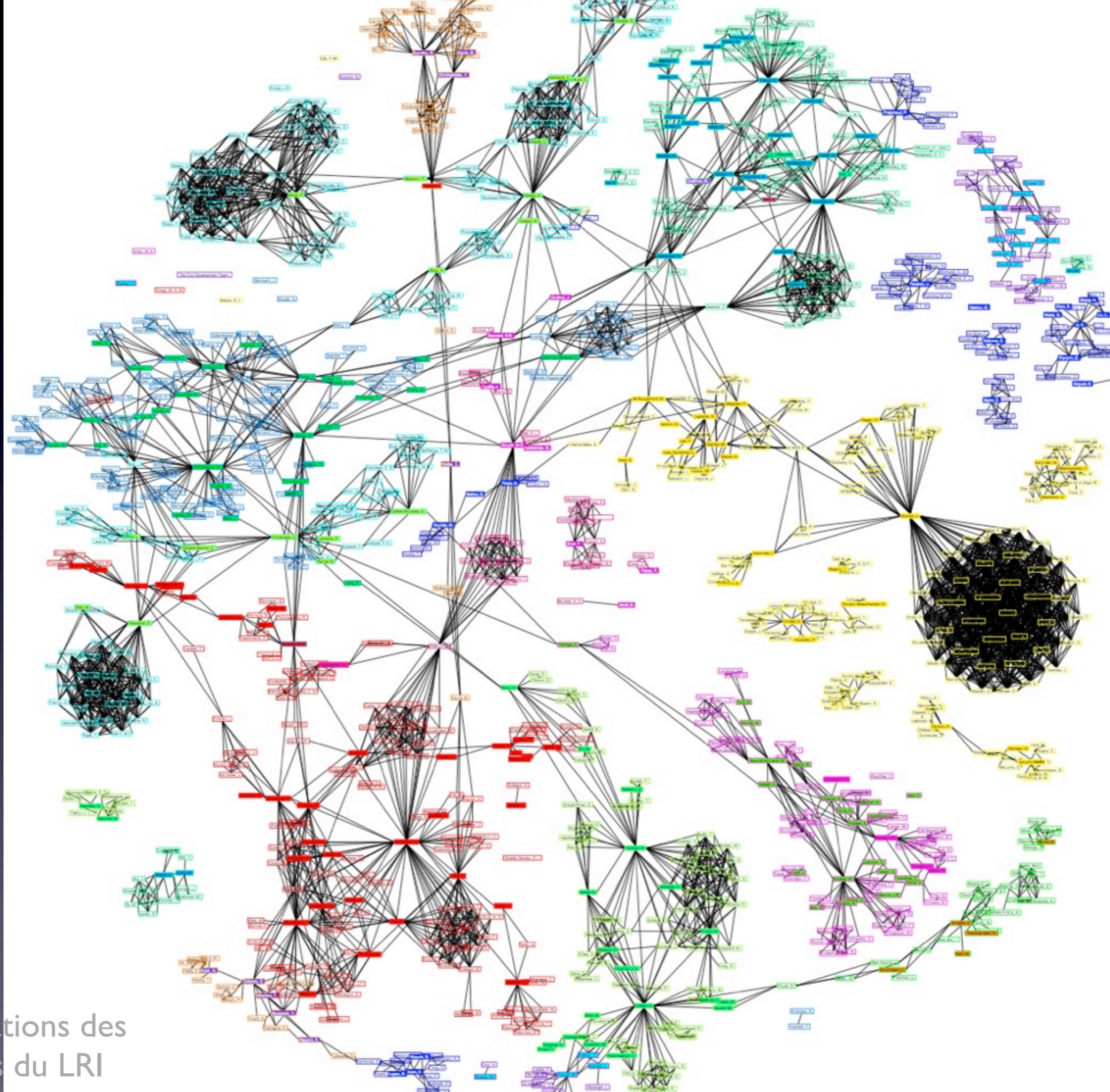
Pour terminer

# Corrélation classement-indice

- Rank **correlations of 0.22** between the peer evaluation based **quality rating** of Netherlands computer science groups and **citation impact** indicators of their papers
  - Peer rating of 42 academic computer science groups in the Netherlands in 2003 (QANU)
  - ISI database plus conference proceedings from ACM, LNCS, IEEE

- “The Mismeasurement of Science” Peter A.  
Lawrence
- La mesure perturbe le système mesuré
- Manipulation possible
- Evolution des pratiques
- Loi de puissance = loi de rareté
  - Prédiction par nature difficile
- Sortir des classements et des indices





Co-publications des  
membres du LRI

J.D. Fekete  
N. Henry