

Les indicateurs bibliométriques: généralités et étude expérimentale

J-P. Merlet

INRIA Sophia-Antipolis



1 Le constat

- une présence de plus en plus forte dans l'évaluation et dans les prises de décision
- une grande **méconnaissance de leur définition et de leur mode de calcul**
- des **avis passionnés**
 - les **partisans**: *ils sont critiqués uniquement par ceux qui sont mal classés et particulièrement les français, quelques rares exceptions où cela ne marche pas bien*
 - les **pragmatiques**: *combat d'arrière-garde*
 - les **irréductibles**: *ils n'ont aucun intérêt*

Réflexion de la Commission d'Evaluation de l'INRIA pour

- mieux comprendre ce que sont censés mesurer les indicateurs
- évaluer leur pertinence et leur fiabilité
- avoir des avis d'experts et pas seulement de bibliométriciens

en particulier dans le contexte de la recherche à l'INRIA



2 Citations: le matériel de base

L'établissement des indicateurs repose sur un matériau brut: **les citations**

Hypothèse forte: une citation est *positive*

Loet Leydesdorff, "Citations: Indicators of Quality? The Impact Fallacy":

Citation impact studies focus on short-term citation, and therefore tend to measure not epistemic quality, but involvement in current discourses in which contributions are positioned by referencing

Les sources de citations généralistes:

- **commerciales:** **Web of Science** (Clarivate), **Scopus** (Elsevier), **MEDLINE/Pubmed** (médecine)
- **gratuites:** **Google Scholar**

Les sources spécialisées pour un domaine:

- informatique: **CiteSeer**, **DBLP**,...



3 Qualité des sources

Couverture

- WoS: 12 000 journals référencés sur 27 000 (?), à partir de 1975
- MEDLINE: 5600 journaux biomédicaux, à partir de 1946
- Scopus: 21 000 journaux,

Qualité

- 7% de références incorrectes dans WoS
- taux de couverture en engineering pour Scopus et WoS: très variable selon les domaines, de 10% à 80%
- les données brutes de Google Scholar doivent être traitées: doublons, erreur d'auteurs, inconsistence, ...



4 Qualité des sources

- littérature bibliométrique: *couverture très variable selon domaine, langue, source . . .*

Pour en avoir le cœur net:

- échantillon de 4 chercheur(se)s de l'INRIA (image, réseau, math financière, info pure)
- senior et de nom non commun (homonymie)
- avec une liste de publications complètes et disponibles
- ayant fourni une liste de supports où leurs travaux peuvent être référencés

Analyse comparée: WoS, Scopus, Google, CiteSeer, DBLP

Première constatation: nécessité d'un lourd travail manuel de correction

- titre inexact, source mal identifiées, citations aberrantes, . . .



Les articles

Chercheur	1 (image)	2 (réseau)	3 (math)	4 (info)
Nombre total d'articles trouvés dans les sources	152	91	86	128
Articles trouvées dans WoS	20.3%	14.28%	18.6%	10.15%
Articles trouvées dans Scopus	34.2%	25.27%	31.3%	3.1%
Articles trouvées dans Google	92%	97.8%	89.5%	98.4%
articles \in WoS \cup Scopus	38.1%	27.5%	33.7%	12.5%

Les citations

Chercheur	1 (image)	2 (réseau)	3 (math)	4 (info)
Nombre de citations WoS	237	163	104	78
Nombre de citations Scopus	652	334	122	52
Nombre de citations Google	1859	981	565	2324
Nombre citations CiteSeer	134	103	0	818

10 à 20% des articles, pourtant cités, ne sont pas trouvés. Même pour un domaine donné le taux de récupération des citations est très variable



Cohérence des citations, chercheur 4

Numéro papier par rang de citation dans GS	1	2	3	4	5	6	7	8	9	10
Nombre de citations GS	208	160	111	101	95	85	83	74	73	70
Nombre de citations WoS	-	-	-	-	-	7	-	-	-	22
Nombre de citations Scopus	44	-	-	-	-	-	-	-	-	-
Nombre de citations CiteSeer	2	152	68	74	32	26	60	-	43	22

Affiliations

WoS	Scopus
INFIA ROCQUENCOURT (sic!)	INRIA
INRIA	Inst Natl de Recherche en (sic!)
INRIA RENNES	I.N.R.I.A.
INRIA RHONE ALPES	Inst. Natl. Rech. Info. et Automat.
INRIA ROCQUENCOURT	LORIA
INRIA SOPHIA ANTIPOLIS	LIFIA-IMAG and Iuria Rhone-Alpes (sic!)
INST NATL RECH INFORMAT & AUTOMAT	
INST NATL RECH INFORMAT & AUTOMAT ROCQUEN COURT (sic!)	
NAT RES INST COMP SCI & CONTROL	

et ne s'améliore pas avec les COMUE, Idex et autre I-site ...



5 Les indicateurs

Journaux: impact factor

- C_j : nombre de citations dans l'ensemble de la base de données d'articles publiés à l'année j dans le journal
- P_j : nombre d'articles publiés à l'année j par le journal

Impact factor à l'année n :

$$IF = \frac{C_{n-1} + C_{n-2}}{P_{n-1} + P_{n-2}}$$

Algorithmica:

- 60 articles publiés en 2004 et 2005
- en 2006 70 citations à des articles de 2004
- en 2006 32 citations à des articles de 2005

$$IF = \frac{70 + 32}{60 + 60} = 0.850$$



6 L'IF: un impact certain



springer.com



#1 Impact Factor

— ISI 2006

- ▶ All Computer Science journals
(#1 of 365)
- ▶ All Artificial Intelligence journals
(#1 of 85)
- ▶ Impact Factor = 6.085

International Journal of Computer Vision

Pourtant, San Francisco Declaration on Research Assessment (DORA, qui vient juste d'être *journal-based metrics, such as journal impact factors, should never* signé par l'INRIA): *be used to assess an individual scientist's contributions, or in hiring, promotion or funding decisions.*



7 L'IF: pourtant..

- analyse de l'IF de 140 journaux en mathématiques et en génétique
 - histogramme similaire
 - impact factor moyen **10 fois plus élevé** en génétique
- sens des 3 décimales ?
- loin d'être cohérent avec l'avis des experts (ex: classement du GDR robotique)
- une fenêtre pas toujours adaptée au temps long des citations (temps moyen de 1ère citation en MMT: 5ans)
- très influencé par des papiers très cités alors que les autres le sont peu

Propice à la manipulation

- on demande à des auteurs de rajouter une liste d'articles du journal
- on multiplie les éditoriaux (qui ne compte pas dans le dénominateur de l'IF ...) en citant beaucoup d'articles du journal



8 L'index SNIP

Proposé par Scopus pour prendre en compte la diversité des pratiques de citations selon les domaines (un exemple extrême ...)

- RIP: nombre moyen de citations reçues par article publié à $j - 3, j - 2, j - 1$
- RDCP: nombre moyen d'articles cités par les articles des revues indexées dans la base qui citent la revue sur la même période.

$$\text{SNIP} = \frac{\text{RIP}}{\text{RCDP}} \quad (\text{version initiale})$$



9 Indicateurs individuels

Le H-index: nombre h d'articles de l'auteur cités au moins h fois (Hirsh, 2005)

Chercheur	1	2	3	4
h-index WoS	8	5	6	5
h-index Scopus	14	6	6	2
h-index GS	21	12	12	25

- ne caractérise pas le niveau actuel d'activité
- sous-estimé pour les chercheurs ayant publié des livres
- ne met pas en valeur les contributions très importantes d'un auteur
- très variable selon les domaines: Sigmund Freud (273), Albert Einstein (112)
- facile à manipuler: salami slicing, auto-citation circulaire, multiplication d'auteurs, citation "honoraire" (+ de 30% en science et engineering, dont 18% pour des collègues), remorquage ...



10 Autres indicateurs individuels

- almetric: reposant sur les recherches internet, les téléchargments, les réseaux sociaux, ...
Par exemple: Research Gate RG score (algorithme inconnu)

- Research impact: N papiers publiés dans des journaux d'IF IF_j et cités C_j fois

$$RI = \sum_{j=1}^{j=N} IF_j(1 + C_j)$$

- g index: le nombre g de papiers qui ont en cumulé au moins g^2 citations:
pour prendre en compte les citations au delà de h
- e -index: pour prendre en compte les citations au delà de h

$$e^2 = \sum_{j=1}^{j=h} C_j - h^2$$

- g_r index: $C_j =$ nombre de citations du papier j , a_j années depuis la publication:
évolue avec l'activité

$$\sqrt{\sum_{j=1}^{j=N} \frac{C_j}{a_j}}$$

- h_l index: h^2/N_a où N_a est le nombre total d'auteurs des papiers considérés pour le h-index:
pour prendre en compte le nombre d'auteurs



11 Conclusion

- n'oublions pas que ce ne sont que des **indicateurs** !
- ils sont soumis à une large **incertitude** !
- il faut bien les connaître pour les calculer ... et les interpréter
- Les indicateurs sont des facteurs d'impact, pas de qualité
- ils sont facilement manipulables
- il ne donne qu'une vue partielle de l'activité scientifique



12 Recommandations

- Il est indispensable de **recouper des sources diverses**
- Il est indispensable d'utiliser plusieurs indicateurs et de les corriger par l'avis d'experts
- Il faut s'attacher, au plus, aux **ordres de grandeur**
- **Ne jamais faire de comparaison inter-domaines**
- Il faut utiliser la littérature bibliométrique pour cerner les biais
- L'usage des indicateurs est **peu favorable à la prise de risques scientifiques**
- le risque de la **gouvernance par les indicateurs**: seul le nombre devient important, pas ce qu'il est censé représenter

Document d'analyse: <https://www.inria.fr/institut/organisation/instances/commission-d-evaluation>

