

Science, Métriques et Éthiques

Marc Bertin

marc.bertin@univ-lyon1.fr

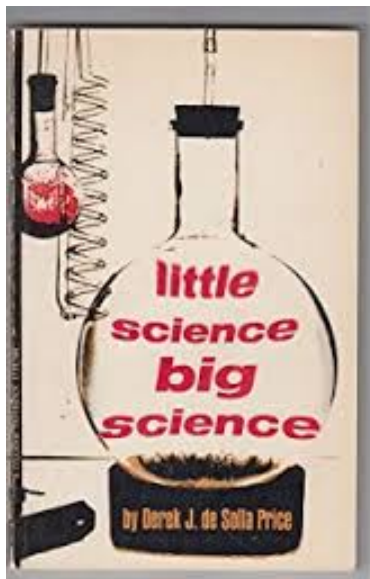
Deux questions pour introduire la problématique

- 1 Qu'est ce que la science?
- 2 Qu'est ce qu'un (bon) chercheur?

- ❶ **Universalisme** : Les énoncés soumis à la communauté scientifique doivent être évalués selon des critères impersonnels, sans égard aux caractéristiques sociales ou institutionnelles de la personne qui propose une découverte ou une théorie.
- ❷ **Communisme** : Toute découverte scientifique est un bien commun qui est le produit de la collaboration entre scientifiques et appartient donc à l'ensemble de la communauté

- ❶ **Désintéressement** : Le savant cherche avant tout la vérité pour elle-même et pour l'ensemble de la communauté scientifique et non pour son profit ou sa gloire personnelle.
- ❷ **Scepticisme organisé** : Norme à la fois méthodologique et institutionnelle qui enjoint le savant à avoir une attitude critique face à tout énoncé nouveau qui doit alors être scruté, vérifié et reproduit avant d'être accepté comme valide et intégré au savoir accumulé.

La thermodynamique du chercheur



On étudie le comportement d'un gaz à différentes conditions de température et de pression. On ne s'intéresse pas à une molécule appelée Georges, se déplaçant à une vitesse spécifique et située en un endroit spécifique à un instant donné; on considère seulement la moyenne de l'ensemble total des molécules où certaines sont plus rapides que d'autres, où elles sont situées au hasard et se déplaçant en différentes directions.

L'infométrie est parfois synonyme de bibliométrie, mais la distinction suivante peut être faite, à savoir que l'infométrie traite par essence, de tous les modèles en relation avec l'information, impliquant qu'elle englobe la bibliométrie ainsi que la scientométrie.

- La scientométrie applique des techniques bibliométriques la science. Il faut souligner que, selon le terme anglo-saxon, Scientometrics traite, des sciences naturelles, de la physique et des mathématiques excluant les sciences sociales.
- La scientométrie va au-delà des techniques bibliométriques lorsque, par exemple, elle compare des politiques de recherche entre États où le nombre de scientifiques par nation.

- Scientométrie = mesure de l'activité scientifique
- Bibliométrie = mesure de l'activité scientifique à partir des publications

Qu'est ce que la science ?

Pour Solla Price : La science est ce qui se publie dans les revues, les articles, les communications et les ouvrages scientifiques. Bref, elle est ce que matérialise la Littérature ...

Ainsi définie, toute littérature scientifique se prête au dénombrement, à la classification et à la représentation sous forme de séries temporelles ; et dans de nombreux cas, c'est même chose faite

Définition du chercheur

- ❶ Définition du chercheur : « *personne qui a publié un article scientifique* »
- ❷ Il ajoutera qu' « *un scientifique n'a rien à faire avec une personne ayant reçu une formation ou non, s'il est employé comme scientifique académique ou industriel, ou s'il réalise un travail scientifique de quelque nature que ce soit, bien que significatif, si ceci n'arrive pas à être publié* » .
- ❸ Enfin, de Solla Price conclura que « *la science est ce qui est publié dans les articles scientifiques* ».
- ❹ Ce type de raisonnement le conduira à conclure que « *le produit final majeur du travail d'un scientifique est l'article qu'il publie* »

Hypothèse de Solla Price

- La science est une population de publications.
- Un document écrit est une sorte d'atome de connaissance.
- Un article représente au moins un quantum d'information scientifique.

La conclusion proposé par de Solla Price sur la nature de la science est contestable. Cependant, elle était une justification suffisante pour envisager la construction d'indicateurs bibliométriques. Ces derniers n'ont été rendus possibles qu'à la condition que l'article soit réduit à une unité de matière dénombrable.

Selon Polanco, dans son travail de synthèse sur l'œuvre de de Solla Price, Aux sources de la scientométrie , les articles scientifiques sont devenus « un outil de définition de la science »

Il nommera cette réduction de la science à la littérature scientifique, le réductionnisme bibliométrique :

« J'appelle réductionnisme bibliométrique le point de vue par effet duquel l'article scientifique devient un outil de définition de la science et l'on fait de la publication écrite un indicateur privilégié de l'activité scientifique, considérant que le produit final de la recherche scientifique est la publication d'un texte écrit. »

❶ Premier postulat

- ▶ Une publication est le produit objectif de la pensée. Dans un contexte scienti que, une publication est une représentation de l'activité de recherche de son auteur.
- ▶ Le plus grand effort du chercheur est de persuader les autres scientifiques que ses découvertes, ses méthodes et techniques sont particulièrement pertinentes.
- ▶ Le mode de communication écrit fournira donc tous les éléments techniques, conceptuels, sociaux et économiques que le chercheur doit dépeindre dans son argumentation.

❶ Deuxième postulat

- ▶ *L'activité de recherche est la confrontation de connaissances acquises par des travaux émanant d'autres auteurs avec les propres réflexions du chercheur.*
- ▶ *La publication devient par conséquent le fruit d'une communion de pensées individuelles et de pensées collectives.*
- ▶ *Ainsi, les chercheurs, pour consolider leur argumentation, font souvent référence à des travaux d'autres chercheurs qui font l'objet d'un certain consensus dans la communauté scientifique.*

Quelques définitions de la bibliométrie

- Pritchard, 1969 « *c'est-à-dire l'application des mathématiques et des méthodes statistiques aux livres et aux différents médias de communication.*»
- Lancaster et Baker, 1977 « *l'application de diverses analyses statistiques afin d'étudier les différents modèles de la paternité des articles, des publications et de leurs utilisations.* »
- White et McCain 1989 « *La bibliométrie est aux publications ce que la démographie est aux populations. [...] étude quantitative de la littérature telle qu'elle est utilisée dans les bibliographies.*»

Quelques définitions de la bibliométrie

- Hawkins et Miller, 1993 « *Les analyses quantitatives des caractéristiques bibliographiques d'un corps de littérature.* »
- Rostaing 1993 « *un outil de mesure auquel on fait appel pour aider à la comparaison et à la compréhension d'un ensemble d'éléments bibliographiques...*

La caractéristique de la bibliométrie est d'établir les études de publications sur des données quantitatives et non plus simplement subjectives (avis des pairs). »

- Loi de Lotka (1926) : Nombre de scientifiques produisant n articles par unité de temps est proportionnel à $1/n^2$
- Solla Price : nombre de chercheurs double tous les 20 ans (1964) les trentes glorieuses.
- Collèges invisibles (activité de discussions et d'échanges informels et de rencontres)
- Merton : Effet Matthieu (1968)

A quoi sert la bibliométrie ?

Utilisation pour :

- Bibliothéconomie et science de l'information
- Histoire des sciences
- Sociologie des sciences
- Évaluation de la recherche
- Politique scientifique
- Économie de la science

Le rôle des indicateurs bibliométriques

Pour prendre des décisions aussi éclairées que possible les décideurs publics doivent avoir une bonne compréhension des activités scientifiques et technologiques. Par exemple, les indicateurs bibliométriques sont les seuls à pouvoir fournir une vision globale de la production scientifique d'un pays.

A quoi sert la bibliométrie ?

Publications aujourd'hui utilisées comme instrument de mesure pour évaluer qualité scientifique

- d'un auteur,
- d'une revue,
- d'un groupe de recherche,
- d'un pays

Les niveaux d'évaluation

Niveau	Domaine d'application
Macro	Mesure la production nationale par pays, province ou même par ville. L'analyse de la production nationale sert à établir des comparaisons entre les systèmes de recherche nationaux et déterminer les liens qui unissent les différentes institutions d'un même système national.
Meso	Ils décrivent la production scientifique des institutions ou des laboratoires et/ou groupes de recherche. Ils peuvent être aussi définis en fonction de programmes de subventions données et donc, de servir à l'évaluation des programmes.
Micro	Le niveau micro s'applique aux individus. Il est possible de mesurer la production d'un chercheur dans le cadre d'une évaluation.

TABLE 2.5 – Niveau du domaine d'application.

Type d'évaluation	Micro	Meso	Macro
Organisation	Personne, équipe	Institut, département	Instituts, groupes de pays, monde
Thématique	Projet	Sous-domaine de recherche	Discipline scientifique, nature de la science
Publication	Un article	Ensemble de publications	Toutes les publications

TABLE 2.6 – Niveau d'évaluation.

- **Au niveau national** : indicateurs de l'OST
- **Au niveau local** : systèmes 3 volet " Budget + Carrières + Publications

Les 8 principaux indicateurs de l'OST

- Part de publications scientifiques
- Indice de spécialisation scientifique
- Part de citations
- Indice d'impact relatif
- Indice d'impact espéré
- Ratio de citations relatif
- Indice d'activité par classe de citations
- Part d'articles en co-publication

Les indicateurs "descriptifs" vs "relationnels"

Les indicateurs bibliométriques se divisent en deux grandes catégories : les indicateurs descriptifs et les indicateurs relationnels qui ont une fonction analytique.

Le dénombrement des articles et des citations, le dénombrement des brevets et des citations dans les brevets sont les indicateurs descriptifs les plus courants. Ils mesurent le volume et l'impact de la recherche à divers niveaux d'agrégation.

Les indicateurs "descriptifs"

Lorsqu'utilisés sur de longues périodes de temps, ils permettent d'identifier des tendances. La méthode du dénombrement est basée sur le calcul du nombre de publications scientifiques attribuables à un acteur, dans un domaine donné.

Il peut s'agir d'un auteur, d'une institution, d'un secteur d'activité regroupant diverses institutions – université, laboratoire public, industrie –, ou encore d'une unité géographique – ville, province, pays.

Le niveau d'agrégation du domaine de recherche peut être une discipline ou une sous-discipline scientifique, une technologie ou encore un créneau technologique spécifique.

Enfin, il est intéressant de rappeler que les indicateurs descriptifs peuvent être appliqués aux publications et aux brevets, selon que l'analyse porte sur la production scientifique ou la production de technologie.

Les indicateurs "relationels"

Dans un document de recherche sur les indicateurs de la science et de la technologie, Godin dresse un portrait de la situation actuelle de la bibliométrie : Les cosignatures sont l'indicateur relationnel le plus courant. L'analyse des cosignatures sert à mettre en lumière les liens et les interactions entre les acteurs des systèmes nationaux et internationaux de science et technologie. Ce sont ces interactions que nous désignons par le concept de flux de connaissances.

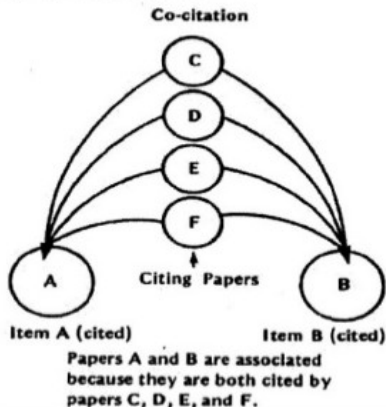
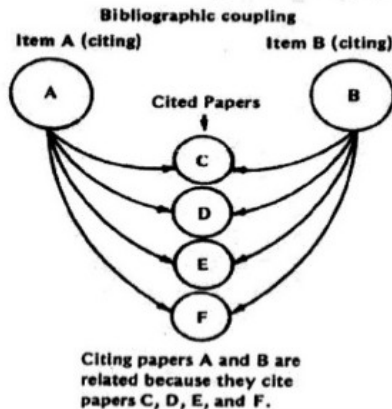
Les indicateurs "relationels"

La méthode des mots associés et celle des cocitations sont aussi des indicateurs relationnels. Ils permettent de dresser des portraits de l'activité scientifique fondés sur le contenu des publications. Ces indicateurs permettent de suivre l'évolution de la science et de la technologie et d'identifier les thèmes de recherche émergents et les acteurs qui y contribuent.

Les indicateurs de cocitations et de mots associés sont toutefois rarement utilisés dans un cadre politique, contrairement aux indicateurs descriptifs et à l'analyse des cosignatures qui sont couramment utilisés pour les tâches de description et d'évaluation de la recherche.

"Co-citation" et "Bibliographic-coupling"

Figure 1: Bibliographic coupling vs. co-citation.



Les indicateurs de flux mettent en évidence les **relations entre les chercheurs, institutions et domaines de recherche**. Pour cette raison, ils sont parfois désignés sous le terme d'indicateur relationnel.

Science et technologie résultent des échanges de connaissance et de la collaboration entre les acteurs de la recherche. De cette prémisse découle le besoin de créer des indicateurs pour identifier ces réseaux d'échange.

L'analyse des flux cherche donc à rendre compte des relations existantes entre les divers acteurs d'un système donné.

La mesure des flux peut porter sur des relations entre des individus, des institutions, des secteurs d'activités ciblées, comme les relations université-industrie par exemple, les liens entre science et technologie, ou encore dresser un portrait de la collaboration entre des provinces et des pays.

L'analyse des flux donne une mesure importante de l'**intégration des acteurs de la recherche** et de l'**étendue du réseau national de collaboration** :

Deux informations qui permettent de mieux saisir l'efficacité de l'intervention publique et ses effets.

Les échanges entre pays sont sans doute le flux de connaissances le plus connu. Pour obtenir un portrait de la collaboration scientifique d'un pays avec l'étranger, il s'agit de compiler, pour une année donnée, toutes les publications où le pays à l'étude est mentionné dans le champ «adresse». On compile ensuite le total des cosignatures avec les autres pays figurant dans le champ «adresse».

L'identification des pays se fait à partir des affiliations institutionnelles des auteurs. Si l'indicateur est relativement simple, le dénombrement des parts de chaque pays participant à une publication reste une opération délicate.

Combien de chercheur en France?

Combien de chercheur en France?

Combien de chercheur en France ?

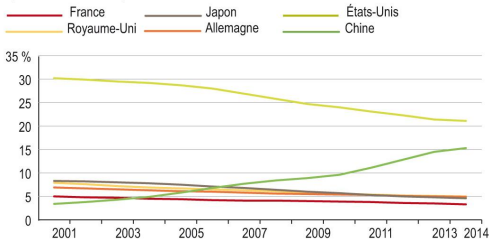
Les activités de recherche sont réalisées soit dans les établissements d'enseignement supérieur ou les organismes de recherche (recherche publique), soit dans les entreprises (recherche privée).

- **295 754 chercheurs** dont 28 % des chercheurs femmes.
- En termes de densité de chercheurs (nombre de chercheurs pour 1 000 emplois), la France "se place au 6e rang mondial"
- L'emploi affecté à la recherche en France s'établit à 482 850 ETP recherche

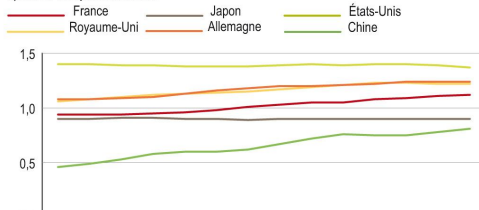
Publications scientifiques des six premiers pays producteurs

02 Publications scientifiques des six premiers pays producteurs (évolution de 2001 à 2014, toutes disciplines confondues)

a) Part mondiale des publications

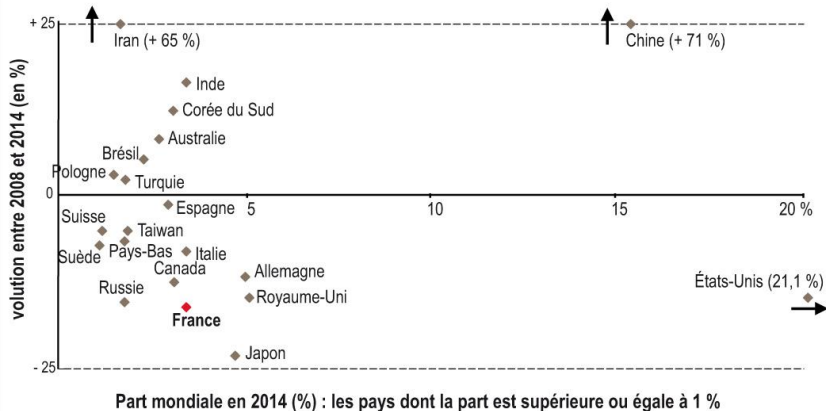


b) Indice d'impact immédiat



Part mondiale des publications

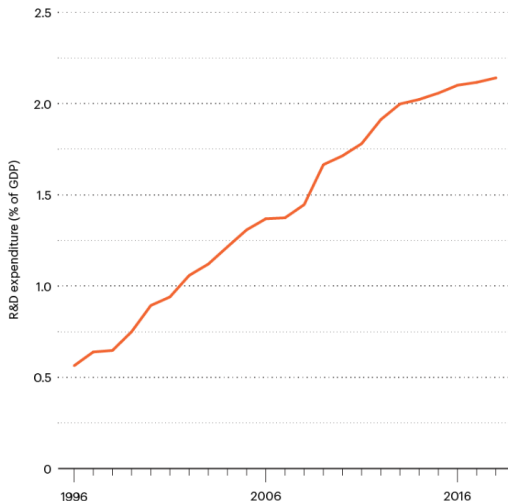
01 Part mondiale de publications, toutes disciplines confondues, des premiers pays producteurs (2014 et évolution de 2009 à 2014)



Source : Thomson Reuters, traitements OST du HCERES.

STEADY SPENDING

China's research and development (R&D) spending as a share of its gross domestic product (GDP) has been rising steadily. This has probably contributed to the recent strong performance of its institutions in the Nature Index.



Évaluation qualitative vs Évaluation Quantitative

Évaluation par les pairs :

Évaluation qualitative du contenu qui conditionne la publication

Originalité et qualité de la contribution

Évaluation en simple ou en double aveugle ou publique

Évaluation quantitative :

Après publication, évaluation de la notoriété d'une publication

Présumé : article beaucoup cité est un très bon article

Les classifications internationales

Classement			Commentaires
Nom	Pays	Date	Description
Classement de Shangai	Chine	2003	Les 100 meilleurs universités sont présentées par domaine de spécialisation (depuis 2007)
Centre for Higher Education Research Ranking	Allemagne	1998	Le CHE Research Ranking révèle la performance de la recherche universitaire allemande
Professional ranking of World universities	France	2007	Nombre d'anciens étudiants figurant parmi les dirigeants exécutifs des 500 plus grandes entreprises mondiales
Times Higher Education	Royaume-Uni	2004	Palmarès de 200 universités mondiales
Global MBA Ranking	Royaume-Uni	2006	Classement synthétique autour du management

TABLE 2.7 – Les classifications internationales

- **La qualité de l'enseignement (10% de l'évaluation)**

Le critère « *Alumni* » : ce critère prend en compte le nombre de prix Nobel (hors littérature et paix) et de médailles Fields reçues par les anciens élèves ayant validé un bachelor, un master ou un doctorat (les distinctions les plus récentes valant plus de points).

- **La qualité de l'établissement (40% de l'évaluation)**

Le critère « *Award* » : ce critère est lié au nombre de prix Nobel (toujours hors littérature et paix) et de médailles Fields obtenus cette fois-ci par les chercheurs (une fois de plus le poids des chercheurs du début du siècle est moins important que les chercheurs en poste).

Le critère « *HiCi* » : le deuxième élément utilisé pour mesurer la qualité de l'établissement est le nombre de chercheurs les plus cités dans leurs disciplines.

- **Le volume de publications scientifiques (40% de l'évaluation)**

Le critère « N&S » : ce critère classant prend en compte les articles publiés dans les revues scientifiques « Nature » et « Science »

Le critère « PUB » : il comptabilise les articles indexés dans le « Science Citation Index », et le « Arts & Humanities Citation Index ».

- La taille de l'établissement (10% de l'évaluation)

Calculé sur les données du WoS qui présentent déjà de nombreux biais (notamment la langue, type de doc, qualité de la base, ...)

- Pb multiples liés à l'affiliation des chercheurs (institutions de rattachement, mixité des labos, articles en collaboration, émérite univ A<>B, ...)
- Pb de nos univ dans la mondialisation : éparpillement , petites tailles, pluridisciplinaires, ...

Les classifications : de Shangai à Taiwan

Système de classement	Indicateurs	Pondération
Classement Académique des Universités Mondiales – SJT « SHANGAI » (Cf dia suivante)	<ul style="list-style-type: none"> Qualité de l'enseignement <ul style="list-style-type: none"> Qualité des membres du corps professoral Nombre de Prix Nobel / Médailles Fields Nombre de chercheurs les plus cités («HiCi») Production de la Recherche <ul style="list-style-type: none"> Nombre d'Articles dans «Nature/Science» Nombre d'Articles dans le «Citation Index» Taille de l'Institution 	10% 20% 20% 20% 20% 10%
Classement Mondial des Universités de Times QS « THE- QS »	<ul style="list-style-type: none"> Évaluation par les Pairs Employabilité des Diplômés Qualité de l'enseignement / rapport élèves-enseignants Étudiants Internationaux Personnel enseignant international Qualité de la Recherche / Citations par Faculté 	40% 10% 20% 5% 5% 20%
Classement des Performances des Articles Scientifiques pour Universités de Recherche « Taiwan »	<ul style="list-style-type: none"> Productivité de la Recherche <ul style="list-style-type: none"> Nombre d'Articles dans les 11 dernières années Nombre d'Articles dans l'année en cours. Impact de la Recherche <ul style="list-style-type: none"> Nombre de Citations dans les 11 dernières années Nombre de Citations dans les 2 dernières années Nombre moyen de Citations dans les 11 dernières années Excellence de la Recherche <ul style="list-style-type: none"> Classement HiCi des 2 dernières années Nombre d'articles HiCi dans les 10 dernières années Nombre d'Articles dans des Revues à fort facteur d'impact Nombre de disciplines dans lesquelles l'Université fait preuve d'Excellence 	10% 10% 10% 10% 10% 20% 10% 10% 10%

Le classement de Leiden - indicateurs d'Impact

- Le **score moyen de citation** (MCS) . Le nombre moyen de citations des publications de l'université.
- Le **score moyen citation normalisée** (MNCS) . Le nombre moyen de citations des publications d'une université, compte tenu des différences sur le terrain, année de publication, et le type de document. Une valeur MNCS de 2, par exemple, signifie que les publications d'une université ont été cités 2 fois supérieur à la moyenne mondiale.
- **Proportion des « top » publications 10%** (PP top 10%) . La proportion des publications d'une université qui, par rapport à d'autres publications similaires, appartiennent à la tranche supérieure de 10% les plus fréquemment cités. (indicateur le plus stable)

- **Proportion des publications de collaboration** (PP collab) . La proportion des publications d'une université qui ont été co-écrit avec un ou plusieurs autres organisations.
- **Proportion publications internationales de collaboration** (PP int collab) . La proportion des publications d'une université qui ont été co-écrit par deux ou plusieurs pays.
- **La distance moyenne collaboration géographique** (MGCD) . La collaboration à distance géographique moyenne des publications d'une université)
- **Proportion des publications à longue distance de collaboration** (PP > 1000 km) . La proportion des publications d'une université qui ont une distance géographique de la collaboration de plus de 1000 km.

Classement de Leyde (Leiden) - Relation comptage / discipline

- **CPP** (Citations per Publication) nombre moyen de citations par article
- **JCS** (Journal Citation Score) impact moyen d'une équipe rapporté à celui d'une revue
- **FCS** (Field Citation Score) impact moyen d'une équipe rapporté à celui d'un champ, i.e. d'une catégorie de sujet
- **MNCS** (mean normalized citation score) = Moyenne des citations des publications d'une entité, pondérée en fonction de la discipline et de l'année de publication
- **PP-Top10** = proportion des publications d'une entité qui, comparée avec les autres publications du même domaine et de la même année, appartiennent au Top 10% les plus citées par ailleurs.

<https://www.leidenranking.com/>