

Entre LOLF et Shanghai : quel espace pour l'évaluation ?

Manuel Durand-Barthez, – Urfist de Paris

► **To cite this version:**

Manuel Durand-Barthez, – Urfist de Paris. Entre LOLF et Shanghai : quel espace pour l'évaluation ?. Bulletin des bibliothèques de France, Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques (ENSSIB), 2010, 422 (4), pp.3 - 19. <<http://bbf.enssib.fr/consulter/bbf-2010-04-0006-001>>. <sic_01426441>

HAL Id: sic_01426441

https://archivesic.ccsd.cnrs.fr/sic_01426441

Submitted on 4 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Entre LOLF et Shanghai : quel espace pour l'évaluation ?

Manuel Durand-Barthez – URFIST de Paris

« Évaluer tue » : un titre qui se voudrait provocateur, celui d'un dossier du *Nouvel âne, le magazine international lacanien*¹, reprend l'antienne d'une mouvance certes minoritaire mais bien présente dans le milieu de la recherche.

S'il est exact que les méthodes d'évaluation actuellement en vigueur en France peuvent être à l'origine d'abus dûment constatés, il n'est pas prouvé qu'une attitude simplement négatrice soit la plus adaptée pour engager une analyse critique et constructive. Au contraire.

La recherche ouverte, à risque parce que franchement innovante (de type « blue sky ») est fortement mise à mal par un milieu plutôt prédisposé aux tendances déjà affirmées de la recherche. Se démarquer des modèles ambiants pose des problèmes mis en avant par d'éminents chercheurs qui ont initié un débat significatif sur ce thème dans la revue *Nature* en 2003. D'abord sous la plume de Peter Lawrence, du Laboratoire de Biologie moléculaire de Cambridge, qui n'hésite pas à s'exprimer sur le népotisme rampant de certains comités de lecture². Son jugement est corroboré par David Colquhoun, du Département de Pharmacologie de l'University College de Londres, sous un autre titre parlant : *Challenging the tyranny of impact factors*³. Ce niveau d'analyse critique touche progressivement le milieu de la recherche française et, avec lui, les professionnels de l'information et de la documentation. C'est vers un objectif de coopération entre ces acteurs que tend la réflexion qui suit.

Elle fera le point sur les principes qui régissent actuellement en France l'évaluation des publications. Elle laissera volontairement de côté l'analyse de nombreux modèles expérimentaux souvent bénéfiques et intéressants : bases de citations liées aux archives ouvertes, indices de distribution de type H alternatifs, Eigenfactor, boîtes à moustaches (box plots) et indices de notoriété, réseaux de co-citations, ...⁴

Des référentiels qui favorisent le domaine STM (Sciences, Technologie, Médecine)

Colonne vertébrale d'un système d'information, la structure d'un référentiel oriente les résultats du traitement des données. Dans le domaine de l'évaluation bibliométrique, on peut évoquer deux éléments moteurs, l'un national (le programme 150 de la LOLF⁵, alias P150), l'autre international (le classement de Shanghai⁶).

Nous verrons qu'en France, la visée concomitante des objectifs 7 à 11 du P150 va infléchir très sensiblement le résultat des mesures effectuées dans le sens de la recherche en sciences exactes et appliquées. Quant au classement de Shanghai, son allégeance assez significative

¹ *Le Nouvel âne* n°10, février 2010, p. 3-19

² Peter Lawrence. « The politics of publication », *Nature* 422 (6929) 20 March 2003, p. 259-261
<http://www.nature.com/nature/journal/v422/n6929/full/422259a.html> doi:10.1038/422259a

³ David Colquhoun. « Challenging the tyranny of impact factors » *Nature* 423 (6939) 29 May 2003, p.479
<http://www.nature.com/nature/journal/v423/n6939/full/423479a.html> doi:10.1038/423479a

⁴ Manuel Durand-Barthez. « L'évaluation des publications scientifiques : nouvelles approches, nouveaux enjeux ». *Les Cahiers du Numérique*, 5 (2) 2009, p.123-141 et « Tendances actuelles en bibliométrie : panorama des ressources, évolution, perception ». *Documentaliste, sciences de l'information* (4) Novembre 2009, p. 44-49
⁵ Agence de mutualisation des universités et établissements (AMUE). *Programme 150, formations supérieures, recherche universitaire, projet annuel de performance*, 2008.
http://www.amue.fr/fileadmin/amue/finances/documents-publications/PAP_2008_Prog_150.pdf

⁶ Aux côtés duquel existent bien sûr des alternatives intéressantes, le classement de Leyde par exemple (<http://www.cwts.nl/ranking/LeidenRankingWebSite.html>). Bien qu'abondamment critiqué, le modèle de l'Université Jiao Tong reste ancré solidement dans les estimations les plus « populaires ».

aux critères de l'Institut for scientific Information (ISI) induit une tendance ouvertement défavorable aux sciences humaines et à une part notable des sciences sociales.

Le Programme 150 de la LOLF

Une dizaine de programmes de la LOLF (codés Pnnn)⁷ visent la recherche. Certains concernent la recherche finalisée, pouvant avoir un caractère spécialisé suivant les secteurs (énergie, spatial, risques et pollution...). La recherche académique est plus particulièrement régie par le P150 qui comporte quelques indicateurs touchant les publications, regroupés sous les objectifs n°7 et 8.

- *Objectif n°7 : Produire des connaissances scientifiques au meilleur niveau international.*
 - Indicateur 7.1 : Production scientifique des opérateurs du programme
Part des publications de référence internationale des opérateurs du programme dans la production scientifique de l'Union européenne et du monde
 - Indicateur 7.2 : Reconnaissance scientifique des opérateurs du programme
Indice de citation à deux ans des publications des opérateurs du programme
- *Objectif N°8 : Développer le dynamisme et la réactivité de la recherche universitaire*
 - Indicateur 8.1 : Part des publications de référence internationale des opérateurs du programme dans les domaines scientifiques définis comme prioritaires en référence européenne (UE25) et dans le monde

Deux autres objectifs recourent des préoccupations d'ordre documentaire: les n°9 et 11.

L'objectif n°9 s'intitule: *Contribuer à l'amélioration de la compétitivité de l'économie nationale par le transfert et la valorisation des résultats de la recherche*. Il s'agit, à certains égards, d'un autre type de littérature scientifique : les brevets (indicateur 9.1), dont le dépôt est jugé comme un élément capital dans la valorisation d'une unité de recherche.

L'objectif n°11, vise à *Consolider l'Europe de la recherche, par le biais des PCRD (programmes cadres de R & D de l'Union européenne)*. L'indicateur 11.3 : *Part des articles co-publiés avec un pays membre de l'Union européenne (UE 25) dans les articles des opérateurs du programme*, rejoint les préoccupations de l'évaluation documentaire. Cette ouverture sur l'extérieur recoupe les visées de l'Objectif n°10 relatif à la présence d'étrangers dans les équipes de recherche⁸ et, partant, à l'attractivité de celles-ci.

Associer étroitement les objectifs 9 et 11 à ceux qui visent directement les publications (7 et 8) a pour conséquence quasi incontournable une mise en avant du secteur S.T.M. On retrouvera une orientation sous-jacente identique dans les analyses proposées par l'OST.

L'I.S.I. et Shanghai

La source affichée par le programme P150 en matière de documentation s'intitule : « ISI-

⁷ France. Ministère de l'économie, des finances et de l'industrie. « Les clés de la LOLF : budget 2006 : projet de loi et architecture » http://www.minefi.gouv.fr/lof/5_1_145.htm En dépit de leur ancienneté relative, les définitions de ces programmes restent en vigueur.

⁸ Proportion d'étrangers parmi les chercheurs, enseignants chercheurs, postdoctorants et ingénieurs de recherche rémunérés par l'Etat à partir du programme ou par les opérateurs du programme.

OST »

On sait que l'ISI [voir encadré] est présent dans le troisième critère du classement de Shanghai : présence de membres en exercice de l'entité dans le fichier *Highly Cited*⁹ (20% du cumul des critères) ; dans le quatrième : présence notable dans le *Web of Science* (également 20%). Quant à l'obligation faite de publier dans *Nature* ou *Science* (cinquième critère à raison de 20 % également)¹⁰ elle concerne deux revues « généralistes d'excellence » mises en avant du fait de leur très fort facteur d'impact au sens de l'ISI. Il n'est donc pas déplacé de considérer que ce classement repose à 50% environ sur la grille de l'ISI.

Les indicateurs de l'O.S.T. et l'I.S.I.

L'OST (Observatoire des Sciences et des Techniques), G.I.P à vocation interministérielle, est géré par quinze acteurs du système français de recherche depuis 1990. Il développe des analyses de type SWOT¹¹ et fournit des indicateurs depuis 2006 à destination des entités de recherche. À côté du volet documentaire, on retrouve des rubriques relatives aux brevets et aux programmes cadres de R & D de l'Union européenne.

Les indicateurs de l'OST liés à la production scientifique sont reflétés par ceux du P150. On en dénombre huit principaux :

Part de publications scientifiques

Indice de spécialisation scientifique

Part de citations

Indice d'impact relatif

Indice d'impact espéré

Ratio de citations relatif (RCR)

Indice d'activité par classe de citations

Part d'articles en copublication

Le facteur d'impact de l'ISI est polyvalent ; il intervient dans les quatrième, cinquième et sixième indicateurs notamment. L'indice d'impact espéré est assez singulier : il « ...tient spécifiquement compte de l'impact des journaux dans lesquels l'acteur publie. Il est égal à l'indice d'impact relatif qu'obtiendrait l'acteur si, dans chaque journal où il publie, ses publications recevaient la moyenne des citations reçues par l'ensemble des publications du journal. »¹² Il est particulièrement incitatif et reflète des exigences claires au niveau normatif du classement (inter)national.

Les périodes d'analyse concernées par ces indicateurs sont généralement de deux ans, à l'image du facteur d'impact de l'ISI. La recherche appliquée est donc nettement favorisée car on est en droit d'attendre un maximum de citations à bref délai. En outre, deux disciplines sont particulièrement privilégiées : les sciences de la matière et les sciences de la nature et de la vie, notamment du fait de leur excellente couverture par le *Web of Science* auquel se réfère l'IPERU¹³ de l'OST : « La base de données source est le *Web of Science*® de Thomson

⁹ <http://isihighlycited.com/> Base librement accessible, sorte de *Gotha* ou de *Who's who* de la recherche mondiale incluant les chercheurs « hautement cités » par l'ISI.

¹⁰ Une « dérogation » est accordée pour certains domaines notamment en sciences sociales ; exemple : les sciences économiques.

¹¹ SWOT (Strengths – Weaknesses – Opportunities – Threats) ou AFOM (Atouts – Faiblesses – Opportunités – Menaces) Cf. European Union, European commission, Europeaid, Co-operation office. *Evaluation, Guidelines : l'analyse SWOT* http://ec.europa.eu/europeaid/evaluation/methodology/examples/too_swo_res_fr.pdf

¹² <http://www.obs-ost.fr/fr/comprendre-les-indicateurs/les-indicateurs-de-l-ost-s-t/indicateurs-sur-les-publications-scientifiques/indice-d-impact-espere.html>

¹³ IPERU = Indicateurs de production des établissements de recherche universitaire <http://www.scd.ups-tlse.fr/barthez/OSTpubliIperu08.pdf>

Reuters qui fait référence pour la bibliométrie dans les domaines des sciences de la matière et de la vie. »¹⁴ Il importe d'observer que les brevets générés par la recherche appliquée sont assez directement liés à l'industrie des matériaux dans le premier cas et à l'industrie pharmaceutique dans le second.

Il ne paraît pas déplacé de citer ici Yves Gingras, bibliomètre expert et réputé, directeur scientifique de l'OST de l'Université du Québec à Montréal : « L'existence du 'facteur d'impact' des revues, publié par la firme Thomson Reuters, ne constitue pas à proprement parler un *classement* officiel. Localement, il peut être utilisé par un comité qui en a décidé ainsi, croyant alors se faciliter la tâche. Il n'a aucun effet global sur la discipline ou même sur le pays. »¹⁵

Se focaliser sur le facteur d'impact d'une revue tend à faire oublier qu'existe aussi le « Facteur d'impact agrégé » (*Aggregate Impact Factor*) des catégories de sujets. Au nombre de 173¹⁶ dans la version STM du JCR et de 56 en Sciences sociales¹⁷, ces *Subject Categories* se chevauchent plus ou moins, du fait qu'une revue n'est pas forcément circonscrite sur le plan disciplinaire avec une rigueur parfaitement « étanche », transversalité oblige. Si l'on aligne un certain nombre de catégories de sujets présentant des affinités en tant que disciplines connexes, visées dans le champ d'exercice d'une équipe de recherche, on peut analyser les facteurs d'impact agrégés des ensembles de revues recouvertes par chaque catégorie en utilisant cette agrégation comme une référence de cotation. Si, au fil des ans, telle ou telle catégorie de sujet voit sa cote baisser, on peut imaginer qu'une agence de moyens privilégie le financement de tel ou tel axe de recherche au détriment d'un autre.

Les bases complémentaires

S'il est relativement clair que les instances administratives françaises semblent privilégier les bases de l'ISI, certaines institutions doivent compléter les opérations de caractérisation de leurs unités avec *Scopus*¹⁸ d'Elsevier. En 2008 des analyses « montrent que 7434 revues sont indexées par les deux bases et que 6256 titres de Scopus ne sont pas indexés par WoS alors que 1467 revues indexées par WoS ne le sont pas par Scopus. Ainsi, 54 % seulement des revues traitées par Scopus se trouvent aussi dans WoS et 84 % des revues de WoS sont dans Scopus. »¹⁹

¹⁴ OST, *Note méthodologique sur les indicateurs de production scientifique en Sciences de la matière et de la vie mesurée par les publications* http://icb.u-bourgogne.fr/plus/halub/OST/Note_methodo OST_2009.pdf II.

Rappel méthodologique ; a) Sources des données

¹⁵ Yves Gingras. « Les systèmes d'évaluation de la recherche » *Documentaliste, sciences de l'information* (4) Novembre 2009, p. 35

¹⁶ http://www.accesowok.fecyt.es/wp-content/themes/Wok/archivos/noticias2007/CATEGORY_S.txt

¹⁷ Le JCR (*Journal Citation Reports*) de l'ISI publie les facteurs d'impact des revues. Il analysait, en 2008, 6620 titres en STM et 1985 en Sciences sociales. La base *Highly Cited* de l'ISI comporte 21 disciplines, réduisant ainsi drastiquement les cas de chevauchement ; l'OST synthétise les 173 catégories du JCR en 8 disciplines STM, plus une neuvième, « pluridisciplinaire », visant des titres auxquels Thomson n'a pas affecté de spécialité scientifique particulière ; exemple : les nanosciences. L'OST compte 15 disciplines en SHS. Cf. OST, *Mémento à l'usage des opérateurs : indicateurs de production scientifique mesurés par l'OST pour les opérateurs du programme 150 "Formations supérieures et recherche universitaire" Vagues 2008*. http://www.obs-ost.fr/fileadmin/medias/tx_ostdocuments/Memento_operateursP150Vagues2008 OST.pdf, p.5

¹⁸ http://info.scopus.com/citationtracker/CTO_author/CTO_author.html

¹⁹ Ylva Gavel, Lars Iselid. « Web of Science and Scopus : a journal title overlap study » *Online Information Review*, 32 (1) 2008, p. 8-21. Cité par Francine Courtial. « Sélectionner des revues pour les chercheurs : l'exemple de l'IGR [Institut Gustave Roussy] » *Documentaliste, sciences de l'information* (4) Novembre 2009, p. 43

Deux sous-domaines STM, notamment, recourent à d'autres sources en vue de comptabiliser les publications. L'informatique d'une part, comme en témoigne un rapport de l'INRIA devenu classique²⁰ sur la comparaison *Web of Science / Scholar Google / Scopus / DBLP*²¹ et la recherche en milieu hospitalier, d'autre part, avec *PubMed*, relayé localement par l'outil d'évaluation de la recherche médicale des établissements hospitaliers universitaires *SIGAPS*²².

La problématique des Sciences humaines

L'enquête du CNRS de 2004

Le problème des Sciences humaines réside en l'incompatibilité relative de ses éléments d'appréciation avec les modalités d'évaluation du domaine STM.

En 2004, le Département Sciences humaines du CNRS a entrepris de réévaluer la qualité des revues qu'il subventionnait. Il a inscrit sa démarche en aval du projet de l'European Science Foundation²³ visant à la constitution de l' *European Citation Index in Humanities*. Dans un article²⁴ exposant les prémisses de ce projet, l'ESF ne remet pas fondamentalement en cause l'autorité de l'ISI dans le secteur des Sciences exactes et appliquées. Il conteste cet institut essentiellement vis-à-vis de la prééminence anglo-saxonne qu'il suppose, à l'instar de la réaction européenne qui a pu suivre l'annonce du projet *Google Print*²⁵.

C'est l'AHCI (Arts and Humanities Citation Index) publié par le même ISI qui fait l'objet d'un « contre-modèle ». Le Département SHS du CNRS reprend les standards de l'ESF en Sciences humaines bien sûr, mais également en Sciences sociales :

- 1) le nombre d'années de référence passe de deux à dix, ce qui peut sembler naturel dans les disciplines SHS.
- 2) les sources de citations sont générées par un noyau de « revues coeur » à l'intérieur de 18 champs disciplinaires. On compte en moyenne 6 « revues coeur » par discipline - moitié anglophone, moitié francophone - jugées prééminentes dans le domaine au niveau international.
- 3) Les 18 groupes sont divisés en *quartiles* dont l'étendue varie en fonction du nombre des titres constituant le corpus. Les trois premières strates correspondent aux trois « rangs » en vigueur au CNRS :
 - A : revues internationales de très haut niveau
 - B : revues internationales de haut niveau
 - C : revues d'audience nationale

La quatrième strate se situe « hors classement ».

Le rang C (et la remarque n'est pas superflue) accorde une importance non négligeable aux revues d'intérêt national.

²⁰ Anne-Marie Kermarrec et al. *Que mesurent les indicateurs bibliométriques ? : document d'analyse de la commission d'évaluation de l'INRIA*, Validé par la Commission d'Évaluation du 12 septembre 2007

http://www.inria.fr/inria/organigramme/documents/ce_indicateurs.pdf, p. 27-28 en particulier

²¹ DBLP = Digital Bibliography & Library Project <http://www.informatik.uni-trier.de/~ley/db/>

²² SIGAPS = Système d'Interrogation, de Gestion et d'Analyse des Publications Scientifiques

<http://www.sigaps.fr/>

²³ <http://www.esf.org/>

²⁴ « Building a European Citation Index in Humanities » *Sciences de l'Homme et de la Société* n°69, mai 2004, p.8-10

²⁵ Lancé aux Etats-Unis en octobre 2004, devenu *Google Books*. En France : *Google Livres*

Le Département a dépouillé 125 revues sur la décennie 1992-2001. À travers elles, 38.973 bibliographies terminales d'articles ont permis le recensement de 584.398 citations²⁶.

Les critères précédemment énoncés contribuent positivement à viser une certaine équité dans l'évaluation. Néanmoins, deux observations s'imposent :

- 1) Ces champs disciplinaires font état de *très nombreuses citations de monographies*, alors que l'article de périodique est quasiment généralisé en sciences exactes et appliquées. On n'est donc plus dans le même environnement documentaire et la comparaison s'en trouve faussée
- 2) La méthode ayant présidé au choix des 125 revues est essentiellement fondée sur la libre appréciation de chercheurs CNRS, délégués par le Département dans les 18 champs disciplinaires, pour proposer des listes de revues paraissant prééminentes dans le domaine, tant au niveau national qu'international. C'est du moins ce qui résulte d'un sondage effectué par nos soins auprès de certains responsables. Ceci posé, même si la méthode peut sembler souffrir de l'empirisme, on ne peut contester à tel ou tel titre le fait qu'il soit effectivement considéré comme une référence dans la discipline.

Configuration actuelle en SHS

En dépit d'inévitables oublis ou erreurs, et de l'objectif affiché d'une réévaluation des revues dans la perspective d'une restructuration de l'aide du CNRS, on a voulu corriger un certain nombre d'anomalies propres au modèle de l'ISI et, ne serait-ce que vis-à-vis de cette méthode, il convient d'en prendre acte favorablement.

Néanmoins Yves Gingras porte un jugement critique sur l'évaluation française dans ce domaine: « Lorsque l'AERES ou l'European Science Foundation créent littéralement un classement des revues en sciences humaines, elles officialisent et figent un état de la structure des relations entre revues, qui est toujours dynamique et changeant. Elles génèrent ainsi des effets en retour qui désavantagent les revues plus jeunes et mouvantes ou plus spécialisées. Sans compter que la procédure de classement est opaque et arbitraire. »²⁷

Face à cette situation, il convenait de réunir un maximum d'éléments propres à répertorier les revues SHS. Ainsi, dans le cadre des actions entreprises par le Très grand Equipement Adonis, M. Dassa et C. Kosmopoulos ont-elles élaboré *JournalBase*²⁸, fichier qui comprend plus de 21000 entrées et près de 10000 titres de revues différents. Il agglomère :

- l'*Arts & Humanities Citation Index* (ISI)
- le *Social Science Citation Index* (ISI)
- l'*European Reference Index for the Humanities* (ERIH)²⁹
- la *liste de l'AERES*³⁰
- *Scopus*

L'ERIH résulte de la coopération de l'ESF susnommée et de l'HERA (*Humanities in European Research Area*)³¹, consortium d'agences de financement de la recherche en sciences humaines : il répertoriait 5000 titres en 2008.

En juillet de la même année, l'AERES en recensait 6000.

²⁶ Sur les détails techniques, voir entre autres : l'*Introduction* à cette enquête par Christian Henriot, *ibid.* p. 7, *La Méthodologie de l'enquête. Ibid.* p. 53 et les *Commentaires sur les résultats de l'enquête. Ibid.* p. 55

²⁷ Yves Gingras. *Id. Ibid.*

²⁸ Michèle Dassa et Christine Kosmopoulos. « JournalBase – Une étude comparative internationale des bases de données des revues scientifiques en sciences humaines et sociales (SHS) », *Cybergeographie : revue européenne de géographie*, mis en ligne le 25 juin 2009, modifié le 14 janvier 2010. <http://www.cybergeographie.eu/index22492.html>

²⁹ <http://www.esf.org/research-areas/humanities/erih-european-reference-index-for-the-humanities/erih-initial-lists.html>

³⁰ <http://www-aeres-evaluation-fr.aw.atosorigin.com/Listes-de-publications-en-Sciences>

³¹ <http://www.heranet.info/>

JournalBase recouvre 27 catégories disciplinaires qui doivent composer avec des périmètres très différents en fonction des bases et des pays.

Perspectives : coopérer, analyser, restructurer

Les alternatives aux principes que nous venons d'énoncer sont nombreuses, comme nous le rappelions au début de cette réflexion synthétique, mais sont rarement agréées aujourd'hui par les instances administratives d'évaluation, du moins en France.

Ainsi que nous le suggérons en introduction, il appartient à la communauté scientifique de réfléchir sur les modalités d'une restructuration des principes d'évaluation et de caractérisation des unités de recherche au plan documentaire. La constitution de groupes de travail associant des enseignants-chercheurs et des professionnels de la documentation permettrait de fédérer les énergies et de mutualiser les expériences. Dans l'absolu, s'y adjoindre des membres d'agences d'évaluation serait évidemment bienvenu. Enfin, la synergie STM/SHS est clairement indispensable.

Vaste projet sans doute, mais d'une brûlante actualité...