



Évaluation des productions scientifiques

9-10 juin 2011 Paris (France)

Des innovations en SHS ?



Actes du colloque

Evaluation des productions scientifiques Des innovations en SHS ?

Sous la direction de

Christine Kosmopoulos et Michèle Dassa

9-10 juin 2011
CNRS, Paris, France

Table des matières

Discours d'inauguration, F Ruggiu.....	1
L'évaluation de la qualité des publications en économie, C Bosquet [et al.]	3
Témoignage : le cas des revues de psychologie, J Pétard.....	10
RIBAC : un outil au service des acteurs de la recherche en SHS, M Dassa [et al.]	19
L'évaluation en Sciences Humaines et Sociales : Comment mesurer ce qui compte, M Maurel	31
Le classement des revues en SHS : nouvelles perspectives européennes., G Mirdal.....	37
Open Access et évaluation des productions scientifiques dans l'espace européen de la recherche, C Ramjoué.....	79
Introduction au libre accès dans la recherche, C Kosmopoulos.....	83
JournalBase ? Une étude comparative internationale des bases de données des revues scientifiques en sciences humaines et sociales (SHS), M Dassa [et al.]	87
Les indicateurs de la recherche en SHS, J Dubucs.....	104
L'évaluation scientifique en SHS : les questions méthodologiques et perspectives de solutions, G Filliatreau	131
Les SHS au prisme de l'évaluation par l'AERES, P Glaudes	143

Discours d'inauguration

François-Joseph Ruggiu

Directeur adjoint scientifique à l'Institut des Sciences Humaines et Sociales du CNRS

Chers collègues,

Mes premiers mots seront pour vous présenter les excuses de Patrice Bourdelais, directeur de l'INSHS, qui est en mission à l'étranger, et qui regrette vivement de ne pouvoir être présent aujourd'hui.

Je suis très heureux d'accueillir en son nom les chercheurs et les spécialistes de l'information scientifique venus parler et assister à ce colloque international intitulé « Evaluation des productions scientifiques : des innovations en SHS ? ». C'est un colloque original à plus d'un titre et je voudrais, dans les quelques minutes dont je dispose, souligner l'exemplarité de la démarche qui nous a réunis ici.

Il est, d'abord, le fruit, d'une collaboration étroite menée depuis plusieurs années entre l'INSHS et le laboratoire "Géographie-cités". Cette collaboration témoigne des échanges nombreux entre l'Institut et les laboratoires et, donc, de la volonté de la direction scientifique de l'Institut d'être au plus près des préoccupations des chercheurs. Elle a émergé lors d'une rencontre entre Michèle Dassa, chargée de mission à l'institut pour les indicateurs et Christine Kosmopoulos, ingénieure de Recherche au laboratoire Géographie-cités, qui sont les deux organisatrices de notre rencontre. En discutant de l'utilisation et de la validité des bases bibliométriques en SHS, elles ont, en effet, conçu l'idée d'une étude approfondie et détaillée de ces outils, étude essentielle à entreprendre pour combler l'absence d'informations précises sur les sources des bases de données bibliométriques et tenter ainsi de mettre un terme à une confusion vivement ressentie dans la communauté scientifique nationale et internationale.

Un projet commun est alors né et il a été soumis en 2007 lors du premier appel à projets du TGE Adonis. Intitulé *Constituer des outils collaboratifs pour évaluer et signaler des ressources en sciences humaines et sociales*, il a été retenu, parmi les 16 premiers, et a démarré en janvier 2008. L'objectif était donc d'apporter des données précises sur le contenu des outils bibliométriques existants, qu'une partie de la communauté scientifique était tentée d'utiliser pour évaluer les sciences humaines et sociales. Il s'agissait, avant tout de fournir un recensement précis des revues, de mesurer leur représentativité, et de mettre en évidence les manques en comparaison des publications réelles des chercheurs.

En 2009, le TGE a renouvelé sa confiance dans ce projet qui a abouti à un recensement exhaustif des revues contenues dans les différents outils bibliométriques étudiés : le Web of Science, Scopus, ERIH et les listes de l'AERES. Présenté à plusieurs reprises depuis 2008, mis et remis sur le métier, ce travail a abouti à la réalisation d'un site web "JournalBase" et qui va vous être présenté cette après-midi.

La direction scientifique de l'Institut a apporté son soutien à ce projet pour plusieurs raisons. L'enthousiasme communicatif de ses deux porteuses est ici un élément essentiel – et le fait que soient présents ici les acteurs majeurs de l'évaluation scientifique en SHS en France en est une nouvelle et brillante preuve - mais il s'ajoute à deux raisons plus profondes.

La première est que la direction scientifique est convaincue de l'importance des revues dans les dispositifs de publications en SHS. A la différence d'autres communautés scientifiques, les

chercheurs SHS ont la chance de disposer d'une très vaste gamme de supports de publication et de diffusion des connaissances : le livre personnel – qui reste à juste titre un élément fondamental pour beaucoup de disciplines SHS ; le livre collectif ; la communication dans un colloque ; l'édition commentée de textes... Cette diversité est une richesse qu'il convient de préserver mais qui ne doit pas masquer le fait que la publication dans les revues à comité de lecture est, sans doute, un des moyens les plus efficaces de publier des résultats de recherche qui ont été évalués par les pairs et qui peuvent, grâce aux abonnements des bibliothèques, mais aussi grâce à la diffusion électronique, atteindre un lectorat étendu : le sens même de notre métier est d'être lu.

La seconde raison qui a amené la direction scientifique de l'Institut à encourager la démarche de nos deux collègues, est que la certitude que le débat très complexe autour d'évaluation doit se nourrir de discussions sur ses principes – et par exemple sur la place de la bibliométrie en son sein qui est une question légitime – mais aussi d'initiatives pratiques qui fournissent des matériaux fiables à partir desquels la réflexion peut progresser : l'idée de comparer les bases de données bibliométriques, d'en exposer les mécanismes et les ressorts est un excellent exemple de ces initiatives.

Il me reste à présent, et c'est une tâche très agréable, à remercier très chaleureusement tous les intervenants à ce colloque qui ont pris le temps, dans un agenda que nous savons particulièrement chargé, de venir présenter le contexte général de l'évaluation en France, en Europe et au-delà, et de venir faire avancer notre réflexion commune dans le sens de l'intérêt collectif scientifique. Je voudrais aussi remercier les participants dont le nombre dit bien les attentes des différentes communautés SHS sur les thématiques abordées aujourd'hui et demain.

Nos remerciements vont également à :

- l'Université Paris 1,
- l'Association des professionnels de l'information et de la documentation (ADBS-Recherche)
- Thomson Reuters, qui commercialise la plate-forme Web of Science
- Elsevier qui commercialise la base de données SCOPUS
- Le TGE ADONIS, évidemment,

Qui ont tous soutenu financièrement ce colloque et, enfin, au CNRS qui nous accueille aujourd'hui.

Enfin, je voudrais remercier le comité d'organisation du colloque, ainsi que toutes les personnes de l'INSHS et du laboratoire Géographie-cités qui se sont impliquées dans la préparation et donc dans le succès de ces deux journées, depuis l'accueil d'aujourd'hui jusqu'aux conclusions de demain, en passant par le cocktail de ce soir. Je vous souhaite à tous et à toutes, un excellent colloque et des échanges particulièrement fructueux. Je vous remercie de votre attention.

L'évaluation de la qualité des publications en économie

Clément Bosquet

Sciences Po

Aix-Marseille School of Economics, CNRS et EHESS

clement.bosquet@sciences-po.org

Pierre-Philippe Combes

ppcombes@univmed.fr

Aix-Marseille School of Economics, CNRS et EHESS

Résumé

Nous synthétisons ici une série de travaux veillant à évaluer la qualité de la recherche en économie en France. Nous prenons soin de présenter les enjeux de ces évaluations bibliométriques et de leurs choix méthodologiques ainsi que les prolongements possibles utiles à des fins de politique publique. L'accent est principalement mis ici sur l'évaluation de la qualité des revues scientifiques puis des institutions de recherche (centres et universités), cette dernière pouvant être réalisée à partir des évaluations moyennes des articles des revues ou bien à l'aide d'indicateurs de citations individuelles dont nous comparons les avantages respectifs. Nous résumons également les résultats d'études utilisant les nouvelles données de citations Google Scholar qui permettent d'élargir le type de support traditionnellement considéré (les revues) aux livres et documents de travail, à la fois du côté des cités et des citants.

Mots clés : évaluation de la recherche, production scientifique, qualité des publications

Abstract: This article summarizes some studies that evaluate the quality of research in economics in France. We pay particular attention to the methodological choices of these bibliometric assessments. We conclude by underlying a number of possible extensions useful for public policy. Most of the article is devoted to the evaluation of the quality of scientific journals and of research institutions (research centers and universities). Publication measures can be obtained from average scores of published articles in scientific journals or from individual citation indexes of each publication. As regards the latter, we present some results using Google Scholar citations, which allow us to consider a larger variety of publications (journal articles but also books and working papers) as regards both cited and citing items.

Introduction

Cette intervention n'a pas pour objet de revenir sur les bonnes ou mauvaises raisons de l'existence d'évaluation de la production scientifique. Cependant, si évaluation il doit y avoir, il semble légitime de se demander s'il existe des façons de la mener plus pertinentes que d'autres. Nous résumons ici une série de travaux évaluant l'impact de certains choix méthodologiques relatifs aux critères d'évaluation des publications académiques en économie. Au delà de la quantité publiée, la difficulté principale est relative à l'évaluation de la qualité des travaux de recherche.

Depuis plus de dix ans, les économistes proposent des évaluations quantitatives systématiques de la production scientifique.

Deux grandes questions émergent : quels supports (revues, ouvrages, etc.) sont considérés ? Comment est mesurée la qualité de la production ? La réponse la plus fréquente que l'on peut par exemple trouver dans Combes et Linnemer (2001) ou Combes et Linnemer (2003) est la suivante. L'activité scientifique est mesurée à l'aide des publications dans les revues à comité de lecture et la qualité moyenne de la revue est attribuée à chacun des articles considérés.

Il découle de ce type de choix une question subsidiaire. De quelle façon mesure-t-on la qualité des revues ? Deux approches s'opposent pour répondre à cette question. D'une part, il est possible d'utiliser des mesures strictes de la qualité des revues, issues d'une bibliométrie purement quantitative. D'autre part, il est possible d'utiliser des mesures issues de l'appréciation par les pairs, en général plus subjectives et multicritères. Cette opposition des deux méthodes ne doit pas laisser penser qu'elles sont totalement exclusives. Par exemple, la catégorisation des revues en économie et en gestion de la section 37 du CNRS

a opté officiellement pour la seconde solution tout en précisant « La section n'a cependant pas ignoré les indicateurs bibliométriques dans son analyse champ par champ, revue par revue. »¹ Néanmoins, dès que l'on s'éloigne de critères purement bibliométriques, il est tout de suite très difficile de justifier tel ou tel choix et de parvenir à des choix faisant l'unanimité. Le choix d'un indicateur bibliométrique fait rarement l'unanimité mais il présente l'avantage d'être complètement transparent quant à sa méthode de calcul.

Qu'elle soit faite par les pairs ou par la bibliométrie, l'évaluation moyenne d'une revue n'a, à l'origine, pas vocation à être unique ou à juger de la pertinence d'une ou de quelques publications données mais de refléter la qualité moyenne d'un ensemble évidemment bien plus large comme l'intégralité des articles de la revue, et ce souvent sur plusieurs années. Afin de mener une entreprise d'évaluation de publications particulières, s'appuyer sur la qualité moyenne des revues peut mal évaluer, en positif comme en négatif, le véritable impact de ces publications. La variation intra-revue de l'impact de ses articles est en effet tout à fait considérable. Cependant, il existe désormais des approches plus directes qui permettent à la fois d'étendre le support de publications considérées (à des ouvrages par exemple) et d'individualiser la mesure de l'impact scientifique. Sont alors utilisés directement des indices de citation du chercheur ou de ses publications, sans passer par l'impact moyen de la revue dans lequel elles sont publiées. Google Scholar, par exemple, ne considère non seulement plus uniquement les revues à comité de lecture mais tout type de supports « académiques » présents sur internet incluant les ouvrages et les documents de travail. Il permet aussi d'attribuer à chaque support non pas la qualité moyenne des supports équivalents mais son propre nombre de citations. Ainsi, deux articles publiés dans la même revue mais avec des nombres de citations différents peuvent être considérés comme d'impact différent.

Le reste de cet article est organisé de la manière suivante. Dans la section 2, nous revenons sur les méthodes d'estimation bibliométriques de la qualité des revues, avec une illustration sur l'économie. La section 3 montre brièvement les utilisations qu'il peut en être fait pour évaluer les institutions et présente une approche directe alternative. La section 4 présente des exemples de travaux de prolongement que permettent ces processus d'évaluation.

Evaluation des revues

Il existe de nombreux classements des revues scientifiques en économie issus d'évaluations quantitatives. Les écarts d'impact entre revues obtenus sont souvent tout aussi importants pour l'évaluation des institutions que la hiérarchie des revues elle-même. Ainsi peut-on considérer qu'il vaut mieux être troisième avec 80% de la note du premier que deuxième avec 70% de celle-ci. Un consensus relativement large, même si on peut le discuter, s'étend au delà des frontières de la science économique : les citations reçues par les publications de la revue constituent une mesure objective de son impact. Dès lors, le principal reproche que l'on peut faire aux classements existants est de considérer un nombre trop restreint de revues. Ainsi, pour l'économie, Laband et Piette (1994) en considèrent 130, Kalaitzidakis, Mamuneas et Stengos (2003), 159, Kodrzycki et Yu (2006), 181, et Ritzberger (2008), 271. Or, il existe plus de 1200 revues référencées dans EconLit, la base de données bibliographique la plus utilisée en économie. La contribution importante de Combes et Linnemer (2010) consiste à donner un classement et un score à l'ensemble des revues EconLit en utilisant une méthodologie en trois étapes décrite ci-après.

Comme le font la plupart des approches précédemment citées, Combes et Linnemer (2010) considèrent que la note d'une revue est une moyenne pondérée d'indices de citations. Le choix de ces indices, et leur poids dans la moyenne, est relativement ad hoc, mais transparent, cherchant à capter des dimensions différentes de l'impact d'une revue. Il peut facilement être modifié si d'autres postulats souhaitent être effectués. La contrainte importante provient du fait que la source principale de citations existant est le Journal of Citation Reports (JCR) de Thomson-Reuters qui ne contenait, en 2007 date de l'étude, que 304 revues en économie. Pour les 898 autres revues recensées dans EconLit, aucun comptage de citations n'est effectué. Combes et Linnemer (2010) proposent d'estimer sur les 304 revues recensées dans le JCR un modèle économétrique reliant leur impact à certaines de leurs caractéristiques observables. Celles-ci étant disponibles non seulement pour ces revues mais également pour les 898 autres, une fois le modèle estimé, il peut être réutilisé pour prédire l'impact des revues pour lesquelles il ne peut être calculé directement. L'indice d'impact choisi est une moyenne pondérée de $2 \times 3 = 6$ indices. Les trois indices de citations sont constitués de deux indicateurs d'impact fondés sur les citations JCR des articles (tels que proposés par

¹ Catégorisation des revues en Economie et en Gestion, Section 37 (Economie/Gestion), du Comité National de la Recherche Scientifique, Version 3.01, octobre 2011.

deux projets Internet : RedJasper et EigenFactor²) et d'un indice-H de citations Google Scholar.³ De façon cruciale, chacun de ces indices est introduit en deux variantes dans la moyenne générale : Sa valeur brute et sa valeur normalisée par la valeur de l'indice prise au niveau mondial par chacun des domaines représentés dans la revue. Ce deuxième élément corrige ainsi totalement pour les différences entre domaines de recherche en termes de citations reçues qui existent au niveau mondial.⁴ Cela tend à prendre en compte la critique principale des indices de citation, à savoir qu'ils sont nécessairement plus élevés pour les domaines surreprésentés dans la profession. A l'inverse, comme l'on peut aussi argumenter que cette surreprésentation est justement un aspect de l'impact scientifique, les deux aspects sont considérés dans l'indice total du fait que la moyenne des deux est calculée.

Les variables observables utilisées pour prédire l'indice d'impact de la revue sont de deux natures. Il s'agit d'une part d'indices fondés sur les citations Google Scholar des revues qui sont disponibles pour l'ensemble des 1202 revues. La deuxième famille est constituée de l'impact moyen des auteurs de la revue, mesuré via les indices RedJasper et Eigenfactor. Autrement dit, on utilise le fait que si une revue non référencée dans le JCR a des auteurs publiant dans des revues référencées par le JCR, l'impact de cette revue est vraisemblablement lié à l'impact des autres revues dans lesquels ses auteurs publient. Cette hypothèse n'est pas forcément valide a priori, mais il s'avère qu'a posteriori la qualité de prédiction du modèle est très élevée, puisque 98% de la variance de l'indice est expliquée par ces variables. Il ne reste plus qu'à prédire les indices des 898 autres revues à l'aide du modèle estimé.

Les inconvénients de cette procédure sont de deux types. Le premier est commun à l'ensemble des méthodes pondérant plusieurs indices pour construire un score : les indices de citation choisis, et leur poids dans le calcul de la moyenne (choisis ici égaux), sont arbitraires. Le deuxième inconvénient tient à l'acceptation de l'hypothèse de validité du modèle estimé pour les revues référencées par le JCR pour les revues qui ne le sont pas. Notons que pour quelques revues qui depuis ont intégré la base JCR, comme le *Journal of the European Economic Association*, le modèle s'est avéré tout à fait pertinent.

Le premier avantage de cette approche est tout d'abord que les critères sont « objectifs ». Dans la mesure où il s'agit de moyennes pondérées d'indices de citations, il est possible de discuter facilement du choix des indices comme du poids qui leur est donné. Il est également possible d'extraire des sous-ensembles de revues par domaine pour obtenir des classements « intra-domaine ». Un classement cardinal continu étant obtenu, on peut appliquer ensuite une fonction choisie pour obtenir un degré de sélectivité plus ou moins fort, choix encore laissé à l'analyste. Finalement, et encore une fois, le support couvert est considérablement élargi (plus de 1200 revues, dont de nouvelles revues créées récemment) par comparaison avec les autres exercices du même type.

A titre d'exemple de différences possibles entre des classements purement bibliométriques des revues et les classements issus des pairs, même lorsque ceux-ci se réclament au moins partiellement d'une approche en termes d'impact bibliométrique, nous proposons une succincte comparaison des classements des revues selon Combes et Linnemer (2010) et selon le comité de la section 37 du CNRS.⁵ L'analyse prend pour donné le nombre de revues appartenant à chacune des 5 catégories considérées dans le classement du Comité CNRS. Elle alloue ensuite les revues entre ces catégories à partir du classement Combes et Linnemer (2010). Autrement dit, sachant qu'il y a 6 revues appartenant à la plus haute catégorie par exemple, les 6 revues les mieux classées par Combes et Linnemer (2010) sont mises dans cette catégorie et l'on recommence avec les catégories suivantes. Ainsi, chaque revue appartenant à l'intersection d'EconLit et du classement du Comité CNRS (qui considère beaucoup moins de revues en économie⁶) appartient donc à deux catégories, celle selon le classement du Comité CNRS et celle selon le classement Combes et Linnemer (2010). Le Tableau 1 est une matrice de transition qui donne pour chaque catégorie du classement Combes et Linnemer (2010) (lignes) le nombre de revues dans chacune des catégories du classement de la section 37 du CNRS (colonnes), et vice et versa. Ainsi, parmi les 6 revues classées dans la catégorie la plus élevée par le Comité CNRS, 5 l'auraient aussi été en utilisant la hiérarchie issue de l'indice Combes et Linnemer (2010) mais une revue n'aurait été que dans la catégorie suivante. A l'inverse, une revue classée dans la deuxième catégorie par le CNRS aurait dû se trouver dans la première. De façon générale, on s'aperçoit que les deux hiérarchies des revues sont très loin de se correspondre. Si le terme diagonal de la matrice est toujours le plus élevé, signifiant que pour un certain nombre de revues celle-ci

² Il s'agit d'indices récursifs valorisant davantage les citations provenant elles-mêmes des revues les plus citées.

³ Une revue obtient un indice-H de valeur h s'il a h articles ayant reçus au moins h citations chacun.

⁴ A cette fin, l'économie présente le grand avantage d'avoir un système normalisé de classification des domaines, utilisé par toutes les revues.

⁵ Version 3.01, octobre 2011.

⁶ Il considère à l'inverse des revues de gestion non prises en compte ici.

ont été placées par le Comité CNRS au niveau où le classement Combes et Linnemer (2010) les aurait aussi situées, les termes sur et sous diagonaux sont aussi très importants. Un grand nombre de revues sont surclassées selon le classement du Comité CNRS par rapport à la hiérarchie Combes et Linnemer (2010) alors qu'un grand nombre d'autres sont déclassées. Au delà des différences dans le nombre de revues considérées (bien plus faible dans le classement du Comité CNRS) et du choix ad hoc du nombre de classes et de revues par classe alors que les indices Combes et Linnemer (2010) sont continus, l'avis des pairs semble s'éloigner relativement fortement d'une hiérarchie fondée uniquement sur des indices de citation. Il est difficile d'interpréter ce résultat mais on peut lui donner deux grandes familles d'explications. La première, que le regard des pairs prend en compte des critères plus larges que celui des citations, ce qui conduit à des différences. La seconde que les pairs distordent, consciemment ou pas, la réalité de l'impact des revues. A ce stade, il est difficile de trancher entre ces deux explications.

Tableau 1 – Matrice de transition entre classements

		Classement CNRS				
		1*	1	2	3	4
Clst. CL	1*	5	1	0	0	0
	1	1	43	19	7	0
	2	0	23	50	31	5
	3	0	3	36	77	37
	4	0	0	4	38	42

Evaluation des institutions

Évaluer la quantité et la qualité des publications des chercheurs est une étape préalable indispensable à l'évaluation de leurs institutions d'appartenance. Une fois disponibles des indices de qualité des revues, ceux-ci peuvent être utilisés pour définir des indices de quantité et qualité des publications des institutions, ce que nous présentons dans une première sous-section. Dans une deuxième, nous proposons une stratégie alternative fondée sur les citations reçues par chaque publication, et non sur les indices moyens d'impact des revues.

Evaluation à partir de la qualité moyenne des revues

Dans Bosquet, Combes et Linnemer (2010), nous calculons un ensemble d'indices de la production scientifique en économie pour l'ensemble des universités et centres de recherche français (105 centres regroupés dans 76 universités), en 1998 et 2008. Cela nous permet de réaliser des études de sensibilité quant aux choix méthodologiques possibles.

Pour effectuer ce travail, notre méthodologie a été la suivante. Nous avons d'abord recensé l'ensemble des articles publiés par tous les enseignants-chercheurs (Universités) et chercheurs (CNRS, INRA, Ecoles) en économie français (soit 2832 personnes équivalent temps plein) dans les 1202 revues EconLit. Ensuite, nous avons calculé des indices prenant ou pas en compte le nombre d'auteurs par article, le domaine (Codes JEL), la qualité de la revue (comme expliquée dans la section précédente) et la longueur relative des articles, et ce pour différentes périodes de temps : 5 dernières années, 40 dernières années, avec un facteur de décompte dans le temps, par année depuis la thèse pour chaque chercheur, etc.

Le premier résultat que l'on peut alors observer est la relative insensibilité de la hiérarchie des universités à l'indice de publication choisi. Les autres tendances qui se dégagent de cette étude sont les suivantes. On observe une augmentation des taux de publiant dans le temps (entre 1998 et 2008), une augmentation de la part des publications des chercheurs français dans le monde, une baisse des disparités entre chercheurs/centres/universités et une courbe de productivité en cloche en fonction de l'âge. Chaque université ou centre de recherche peut également se situer dans la hiérarchie française en volume total et en production moyenne par chercheur et peut également évaluer comment sa position évolue dans le temps. Une analyse effectuée pour chaque grand domaine de recherche pris séparément permet aussi de déterminer ceux pour lesquels l'université ou le centre sont les plus visibles.⁷

Evaluation à partir des citations individuelles

Les approches fondées sur la prise en compte de la qualité des publications à partir de la qualité moyenne des revues présentent le risque de masquer des différences potentiellement importantes entre publications d'une même revue. De nouveaux outils permettent cependant désormais de mesurer l'impact individuel de chaque publication, et donc de ne plus avoir besoin d'utiliser les indices moyens d'impact des revues. Dans Bosquet et Combes (2011a) et Bosquet et Combes (2011b), nous ne calculons plus des taux de publiant ou des volumes totaux et moyens d'articles publiés en prenant en compte la qualité moyenne des revues, mais

⁷ Pour plus de détails, se référer directement au rapport ou à sa note de synthèse, disponibles en ligne aux liens <http://www.vcharite.univ-mrs.fr/pp/combes/RapportRanking010310.pdf> et <http://www.vcharite.univ-mrs.fr/pp/combes/SyntheseRapportRanking.pdf>, respectivement.

des indicateurs de citations totales, par article, des indices-H et des indices-G des institutions à partir de ces mêmes indicateurs calculés pour chacun de leurs membres.⁸ Ces indices de citation sont issus de Google Scholar. Ici également, de nombreux tests de robustesse sont effectués selon le fait que l'on prend en compte ou pas le nombre d'auteurs par article ou l'âge de ces articles, et les évaluations sont également effectuées sur différentes périodes de temps.

Les avantages de cette nouvelle perspective par rapport à une analyse des publications sont les suivants. D'une part, le type de supports considérés pour chaque chercheur est bien plus large, puisqu'étendu à tous ceux présents sur des sites Internet académiques (articles mais aussi ouvrages, documents de travail, etc.). L'extension est également très importante du côté des citations, puisque sont considérées non pas seulement celles émanant de revues académiques mais également de tous ces mêmes supports présents sur des sites académiques. Pour l'ensemble des économistes exerçants en France en 2008, on compte par exemple 42 448 entrées GS ayant au total 265 578 citations. Les contributions interdisciplinaires sont également mieux prises en compte puisque l'on n'est plus obligé de se restreindre au champ économique que représente EconLit. Ainsi, les publications des économistes en sociologie, histoire, statistiques ou gestion sont par exemple considérées, et l'on pourrait élargir aux mathématiques pures, à la physique etc. La conclusion principale de cette étude est qu'il n'y a pas de fortes différences dans la hiérarchie des institutions françaises entre cette approche utilisant les citations et l'approche précédemment présentée utilisant la qualité moyenne des revues, même si, au niveau d'un chercheur donné, les différences sont plus importantes. Néanmoins, les institutions dont le cœur de métier est davantage orienté « gestion » par exemple progressent légèrement dans les classements.⁹

Au delà de la mesure, des guides possibles de la politique scientifique

Dans notre démarche scientifique, nous ne considérons pas que l'évaluation des revues, chercheurs ou institutions est une fin en elle-même mais une étape préalable à des études pouvant être instructives en termes de politique de la recherche. Ainsi, dans Bosquet et Combes (2012) et Bosquet et Combes (2011c), nous nous intéressons à l'évaluation des déterminants de la production de recherche des universités. Nous tentons alors de répondre au type de questionnement suivant : quelle est le rôle pour l'activité de publication de la taille des laboratoires, de leur diversité thématique, de leur composition en termes de ratio maîtres de conférence / professeurs, de la charge d'enseignement des chercheurs à l'université, de leurs liens avec des centres de recherche étranger, etc. ? Ce type de quantification devrait être susceptible de guider la politique de recherche et d'éventuellement identifier des stratégies plus efficaces que d'autres en termes de publications ou de citations reçues. Dans un autre travail en cours, Bosquet, Combes et García-Peñalosa (2012), nous cherchons à évaluer les origines des différences de publication des hommes et des femmes et les comparons à leur statut (maître de conférences versus professeur par exemple) sur le marché académique. Cela peut par exemple permettre de déceler une éventuelle discrimination de ces dernières, à niveau de publication donné. De nombreuses autres utilisations des indices de publication peuvent être envisagées.

Références

Clément Bosquet et Pierre-Philippe Combes : *Comparaison des mesures Econlit et Google Scholar de la production de recherche en économie en France en 2008*. Direction Générale de la Recherche et de l'Innovation (DGRI) du Ministère de l'Enseignement Supérieur et de la Recherche, 2011a.

Clément Bosquet et Pierre-Philippe Combes : Un panorama de la recherche française en économie comparant les approches Google Scholar et Econlit. *GREQAM Working Paper*, (2011-56), à paraître dans *Revue d'Economie Politique*, 2011b.

Clément Bosquet et Pierre-Philippe Combes : *Déterminants de la production d'articles de recherche en France*. Direction Générale de la Recherche et de l'Innovation (DGRI) du Ministère de l'Enseignement Supérieur et de la Recherche, 2011c.

⁸ Un chercheur a un indice-G de valeur g s'il a g articles ayant reçus en moyenne g citations chacun, ou g^2 citations au total.

⁹ Pour plus de détails, se référer directement au rapport ou à sa note de synthèse, disponibles en ligne.

Clément Bosquet et Pierre-Philippe Combes : Do large departments make academics more productive? Agglomeration and peer effects in research. *Mimeo Greqam*, 2012.

Clément Bosquet, Pierre-Philippe Combes et Cecilia Garcia-Peñalosa : Gender differences in a micro labour market : promotions amongst academic economists in France. *Mimeo Greqam*, 2012.

Clément Bosquet, Pierre-Philippe Combes et Laurent Linnemer : *La publication d'articles de recherche en économie en France en 2008. Disparités actuelles et évolutions depuis 1998*. Direction Générale de la Recherche et de l'Innovation (DGRI) du Ministère de l'Enseignement Supérieur et de la Recherche, 2010.

Pierre-Philippe Combes et Laurent Linnemer : La publication d'articles de recherche en économie en France. *Annales d'Economie et de Statistiques*, 62:5-47, Avril/Juin 2001.

Pierre-Philippe Combes et Laurent Linnemer : Where are the economists who publish ? Publication concentration and rankings in Europe based on cumulative publications. *Journal of the European Economic Association*, 1(6):1250-1308, December 2003.

Pierre-Philippe Combes et Laurent Linnemer : Inferring missing citations. A quantitative multi-criteria ranking of all journals in economics. *GREQAM Working Paper*, (2010-25), 2010.

Pantelis Kalaitzidakis, Theofanis P. Mamuneas et Thanasis Stengos : Ranking of academic journals and institutions in economics. *Journal of the European Economic Association*, 1(6):1346-1366, December 2003.

Yolanda K. Kodrzycki et Pingkang Yu : New approaches to ranking economics journals. *B.E. Journals in Economic Analysis and Policy : Contributions to Economic Analysis and Policy*, 5(1):1-42, 2006.

David N. Laband et Michael J. Piette : The relative impacts of economics journals : 1970-1990. *Journal of Economic Literature*, 32(2):640-666, June 1994.

Klaus Ritzberger : A ranking of journals in economics and related fields. *German Economic Review*, 9(4):402-430, November 2008.

Témoignage : le cas des revues de psychologie¹

The ranking of the HSS journals by the AERES: the case of the psychological journals

Jean-Pierre Pétard

directeur de publication du *Bulletin de psychologie*

Résumé

S'appuyant sur le cas d'une revue généraliste de psychologie en langue française, cette contribution souligne les errements du dispositif mis en place par l'Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES). Les effets néfastes de facteurs de comparaison utilisés pour classer les revues, notamment la langue de publication et le facteur d'impact, sont soulignés. L'usage, la diffusion et la valorisation de tels classements menacent l'existence de nombre de revues scientifiques et de domaines de recherche des sciences humaines et sociales.

Mots-clés : sciences humaines, sciences sociales, revues scientifiques, bibliométrie, évaluation de la recherche

Abstract

Leaning on the case of a generalist psychological journal in French language, this contribution emphasizes the misguided ways of the provision made by the Agency for the evaluation of the research and higher education (AERES). The harmful effects of the comparison factors used to classify the journals, in particular the language of publication and the impact factor are emphasized. The use and valorization of such rankings threaten the existence of many scientific journals and fields of human and social sciences research.

Key-words : humanities, social sciences, bibliometry, scientific journals, science assessment

Des quatre missions dont est chargée l'Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES) (Code de la recherche, Article L114-3-1, 2008), aucune ne mentionne un quelconque classement des revues scientifiques. Cependant, parmi les quinze domaines des Sciences humaines et sociales (SHS), la commission chargée du domaine "Psychologie – Éthologie – Ergonomie" a publié un classement, en 2008, puis en 2009. Ce classement a été retiré du site de l'AERES en janvier 2012².

C'est en suivant ce qui est advenu à une revue généraliste de psychologie que nous proposons d'exposer certains aspects des classements de l'AERES, d'abord en situant les classements des revues de psychologie parmi ceux des autres domaines des SHS, puis en

1 Pratiques d'évaluation de la recherche en SHS. Table ronde

2 En juin 2011, le classement dans le domaine "Psychologie-Éthologie-Ergonomie" était tel que décrit dans cet article. Ce n'est qu'en décembre 2011, après qu'une mise en demeure eut été adressée, par notre association "Conférence des publications de psychologie de langue française", au président de l'AERES, sous peine d'action par voie de droit, que la liste de classement des revues, d'octobre 2009 a été transférée dans une rubrique "Archives des listes antérieures n'étant plus en usage" où elle demeure consultable. En janvier 2012 une nouvelle liste de revues, sans classement, a été mise en ligne, datée de juillet 2011 (sic).

analysant les effets de ces classements, qui portent sur 2 838 revues, au regard de deux critères : le facteur d'impact et la langue de publication.

Le contexte

Je suis directeur de publication d'une revue, généraliste, de psychologie, Le *Bulletin de psychologie* créé en 1947. C'est une revue de langue française qui vit grâce à ses abonnements, et qui est largement diffusée en France³ et à l'étranger.

En 2003, Philippe Jeannin remettait son rapport de mission « La "revuemétrie" de la recherche en SHS »⁴, au ministère délégué à la recherche et aux nouvelles technologies. Sur les 325 périodiques retenus dans son étude⁵, le *Bulletin de psychologie* figurait au 28^e rang et au 4^e rang des revues de psychologie de langue française. Pour mémoire, les données issues de cette étude avaient été recueillies auprès des enseignants-chercheurs de la discipline.

Aussi, en juillet 2008, avons-nous été stupéfaits par la publication, sur le site de l'AERES, d'un classement des revues en psychologie. Le *Bulletin* était classé au plus bas de l'échelle (C)⁶, lors du premier travail de la commission ad-hoc de l'AERES. Un deuxième classement était publié le 1^{er} septembre 2008, accompagné d'une explication de la commission : « Un travail détaillé, titre par titre, a été effectué sur les revues de la classe C. Certaines de ces revues ont été relevées par la commission dans la catégorie dite B' ». Un troisième classement a été publié en octobre 2009. C'est celui qui fait référence en juin 2011, et dont je vais traiter ici.

Je propose d'exposer ici, après avoir présenté certains aspects du classement des revues de psychologie, quelques réflexions et questions à propos des effets probables de ce genre de classement.

3 . Sur l'ensemble des revues de psychologie de langue française, le *Bulletin de psychologie* est celle qui a le plus grand nombre de localisations fournies par le Sudoc Abes (216) <www.sudoc.abes.fr>.

4 Pour télécharger le rapport : <http://www.inrp.fr/vst/Rapports/DetailRapport.php?id=52>

5 La méthode utilisée pour cette étude « repose fondamentalement sur un tri raisonné des bases de données qu'utilisent les chercheurs. Pour conduire une évaluation disciplinaire pertinente, la méthode suivie se doit d'être aussi scientifique, éthique, globale, flexible, transparente, accessible, stimulante, productive, actualisable et internationalisable. Cette méthode consiste à croiser les bases qui font autorité (ISI – Institute for Scientific Information – ou autres), lister les titres de revues qu'elles contiennent, et à demander, de façon large, à chaque communauté, ce qu'elle en pense. La scientificité de la revue est alors définie ainsi : est scientifique la revue qui est jugée telle par la communauté » (extrait du rapport cité dans la note 2). Cette enquête, réalisée auprès de 526 personnes (les responsables de laboratoires et d'équipes de chercheurs en psychologie, et, par tirage au sort, à un professeur sur deux et à un maître de conférences sur trois de la 16^e section du CNU), a reçu 112 réponses exploitables. Chaque revue était évaluée par 4 choix : – « Oui, niveau 1 », si vous estimez que la revue est de niveau 1 (revues au cœur de la discipline, dans lesquelles il est particulièrement recommandé aux chercheurs de publier). – « Oui, niveau 2 », si vous estimez que la revue est de niveau 2 (revue de psychologie spécialisée, ou revue au cœur d'une autre discipline, accueillant des contributions de psychologie ; il est recommandé aux chercheurs de publier dans ces revues). – « Non », si vous estimez qu'elle n'est pas de type 1 ou 2. – « Ne sais pas », si vous ne savez pas si la revue est de type 1 ou 2, ou bien si vous ne la connaissez pas.

6 « classement des revues : – A : une revue A est une revue classée A au moins dans une de ses catégories d'appartenance (voir PsycINFO), en général elle a un facteur d'impact élevé dans sa catégorie ; – B : une revue B est une revue qui a un facteur d'impact (voir PsycINFO), mais qui n'est classée A dans aucune de ses catégories d'appartenance ; – C : une revue C est une revue répertoriée » (AERES, juillet 2008).

Les listes de l'AERES pour les SHS

Examinons, d'abord, sur la base des classements publiés par l'AERES, tels qu'ils apparaissaient, en juin 2011, sur son site internet⁷, la situation de la psychologie par rapport aux autres disciplines.

Les SHS sont réparties selon 15 domaines [voir tableau 1]. Pour chaque domaine, la commission en charge du classement fait précéder son travail d'un court préambule, à peu près identique pour tous, soulignant l'importance de la « production » de livres et la multiplication des revues, « outil essentiel d'animation et de diffusion permanente ». Ensuite, chaque commission expose ses choix en matière de classement des revues.

Tableau 1. Les 15 domaines des SHS et la date de publication des classements

Domaine	Mise en ligne
Anthropologie - Ethnologie	10/02/10
Arts	01/10/10
Droit	30/06/10
Économie - Gestion	16/06/10
Géographie – Aménagement – Urbanisme	21/02/11
Histoire - Histoire de l'art - Archéologie	mars-09
Histoire et philosophie des sciences	09/11/09
Philosophie	nov. 2009
Psychologie – Éthologie - Ergonomie	14/10/09
Sciences de l'éducation	17/11/09
Sciences de l'information et de la communication	oct. 2010
Sciences et techniques des activités physiques et sportives	14/12/09
Science politique	29/01/10
Sociologie - Démographie	24/06/10
Théologie	23/11/09

Sur les 15 domaines, 10 commissions optent pour la publication d'une liste dite « périmètre », c'est-à-dire sans classement [voir tableau 2] ; 3 domaines classent les revues selon 2 rangs ; 1 domaine, les Staps, selon 3 rangs. Enfin, le domaine, « Psychologie - Éthologie – Ergonomie », classe les revues en 6 catégories.

Tableau 2. Types de classement retenus, par domaine

Domaine	Types de classement
Anthropologie - Ethnologie	périmètre
Arts	périmètre
Droit	périmètre
Économie - Gestion	périmètre
Géographie – Aménagement – Urbanisme	périmètre
Histoire - Histoire de l'art - Archéologie	périmètre
Histoire et philosophie des sciences	1 ^{er} et 2 ^e rang
Philosophie	périmètre
Psychologie – Éthologie - Ergonomie	N/D, excellent, très bon, bon, assez bon, modéré

⁷ <http://www.aeres-evaluation.fr/Publications/Methodologie-de-l-evaluation/Listes-de-revues-SHS-sciences-humaines-et-sociales>

Sciences de l'éducation	référent, interface
Sciences de l'information et de la communication	périmètre
Sciences et techniques des activités physiques et sportives	1 ^{er} , 2 ^e et 3 ^e rang
Science politique	périmètre
Sociologie - Démographie	périmètre
Théologie	1 ^{er} et 2 ^e rang

Les bases du classement de la commission Psychologie sont présentées dans l'encadré 1

Encadré 1 : Arguments présentés par la commission de l'AERES pour le classement en psychologie

« (...) La liste est constituée de l'ensemble des revues référencées dans *PsycINFO*, d'une partie de celles d'*Ergonomics Abstracts*, de revues d'éthologie (sélection constituée par des pairs), d'une partie des revues de *PubMed* (sélection par le CNU 16e section) et d'une partie des revues de neurosciences *Web of Science* (sélection par des pairs). La liste est donc large, intégrant de très nombreuses revues.

La commission a décidé de ne pas caractériser ces revues en fonction d'un critère de qualité scientifique. C'est donc une liste unique, une liste "périmètre", qui est adoptée. Toute publication dans une de ces revues est retenue pour caractériser la production scientifique de l'unité de recherche évaluée.

Toutefois, afin d'aider la communauté scientifique à situer ces revues les unes par rapport aux autres, **il leur a été associé un critère descriptif**, leur niveau d'impact exprimé en terme relatif et non pas absolu, lorsqu'un tel descripteur était disponible. Le **facteur d'impact** retenu est celui calculé sur 5 ans afin de tenir compte de la "durée de vie" plus longue des articles en psychologie que dans certaines autres sciences. **Il a été calculé de manière relative**, c'est-à-dire rapporté à la distribution des impacts de toutes les revues relevant d'une même sous-discipline (telles qu'elles sont définies dans les catégories de *Journal Citation Reports (JCR)– Science Edition* et *Social Sciences Edition*). (...)

(...) On trouvera ainsi associé à chaque revue, l'un des **six descripteurs** suivants : impact non disponible (N/D), impact excellent, très bon, bon, assez bon et modéré. Le niveau excellent qualifie les 10% de revues présentant le plus haut facteur d'impact relatif (pour les sous disciplines concernées), le niveau très bon correspond aux revues situées entre les 10% et 25%, bon entre 25% et 50%, assez bon entre 50% et 75% et modéré pour le reste. Lorsqu'une revue relevant de plusieurs disciplines ou sous-disciplines se trouve classée différemment selon celles-ci, le meilleur classement a été retenu. Enfin, la liste des revues publiées ici n'étant pas exhaustive, il est important d'ajouter que les revues relevant du domaine de la psychologie de la santé (référencées Medline) sont prises en considération lorsqu'elles sont indexées dans le JCR Sciences ou Social Sciences. »

Outre le fait que les rédacteurs de la commission ne précisent pas ce qu'il faut entendre par la « qualité scientifique » d'une revue, ils persistent dans l'utilisation, pour le classement des revues, de facteurs d'impact auxquels ils appliquent des correctifs raffinés. Ainsi chaque revue aura l'un des six descripteurs suivants : « impact non disponible, impact excellent, très bon, bon, assez bon et modéré ».

Ces explications (voir encadré 1) nous conduisent à nous poser quelques questions :

- Plusieurs outils fournissent des Facteur d'Impact (FI). Pourquoi avoir privilégié la référence au Journal Citation Report (JCR) ?
- Sur quelles analyses, des distributions comme les seuils 10%, 25% etc. sont-ils fondés ?
- Pourquoi choisir de passer d'un ensemble continu (le FI) à un ensemble discret (excellent, très bon, bon, etc.), réduisant ainsi les qualités métriques du Facteur d'Impact ?

La transparence déclarée par l'AERES n'est pas au rendez-vous. Du moins, pas dans cette commission (Laurens, 2009).

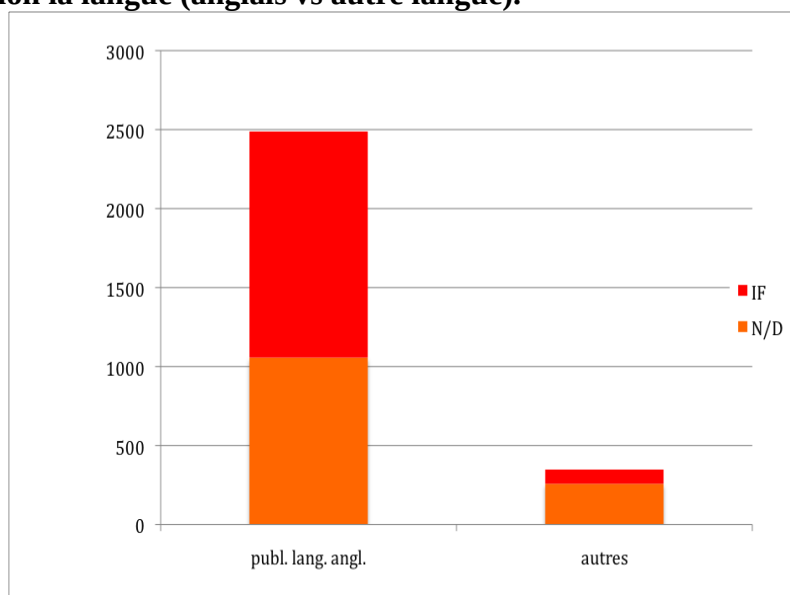
En prenant l'ensemble des 2 838 revues retenues dans le domaine « Psychologie – Éthologie – Ergonomie », nous nous sommes interrogés sur la langue de publication [voir tableau 3 et graphique 1].

Tableau 3. Répartition des revues, selon qu'elles ont ou non un facteur d'impact (FI), et selon la langue (anglais vs autre langue)

	publications langue anglaise		autres langues		total	
N/D*	1064	42 %	251	82 %	1315	46 %
FI**	1468	58 %	55	18 %	1523	54 %
Total n	2532		306		2838	
Total %	89 %		11 %			

Il n'est pas surprenant de constater que 89% des revues du domaine sont des publications en langue anglaise et 11% dans d'autres langues. Il apparaît, également, qu'une part non négligeable de revues, quelque soit leur langue de publication n'a pas de FI (46%). Pour les publications de langue anglaise, la proportion est sensiblement la même (nécessairement, du fait de la forte proportion de ces revues dans l'ensemble) : 42% n'ont pas de FI. En revanche, 82% des revues, dans la colonne « autres langues », n'ont pas de FI.

Graphique 1. Répartition des revues, selon qu'elles ont ou non un facteur d'impact, et selon la langue (anglais vs autre langue).

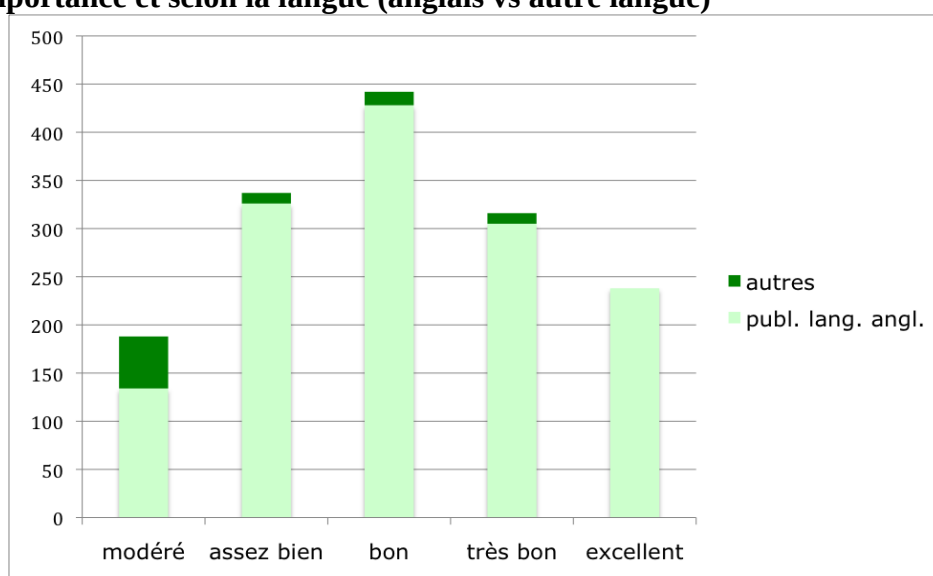


Si l'on s'intéresse maintenant aux seules revues qui ont un FI, soit 54% des 2 838 revues – rappelons qu'elles ont été classées selon un critère qu'il faut lire comme « descriptif », selon le préambule de la commission, mais qu'on tendrait, plutôt, à désigner comme « qualitatif » –, la grande majorité des revues de langue anglaise est classée en « bon », « assez bien », « très bon » et « excellent », relativement peu en « modéré ». Alors que, pour les autres langues, l'essentiel des revues ayant un facteur d'impact se retrouvent sous la caractéristique « modéré » [tableau 4 et graphique 2].

Tableau 4. Répartition des revues dotées d'un facteur d'impact, selon son importance et selon la langue (anglais vs autre langue)

classement	publications langue anglaise	autres langues	total	% selon catégories
modéré	145	43	188	12,34%
assez bien	332	5	337	22,13%
bon	439	5	444	29,15%
très bon	314	2	316	20,75%
excellent	238	0	238	15,63%
	1468	55	1523	

Graphique 2. Répartition des revues dotées d'un facteur d'impact, selon son importance et selon la langue (anglais vs autre langue)



Quelques remarques

La liste de revues du domaine « Psychologie-Éthologie-Ergonomie » fait apparaître une présence écrasante de revues de langue anglaise. L'anglicisation de la communication scientifique est un fait qui n'a pas, bien sûr, que des inconvénients.

En revanche, la valorisation excessive, et absurde, du FI des revues, lorsqu'on prétend s'intéresser à la qualité du travail des chercheurs, des enseignants-chercheurs et des laboratoires, outre qu'elle méprise le travail des experts qui évaluent les articles, menace la vie même et des revues et des travaux de recherche.

En psychologie, et, sans doute, également, dans beaucoup de secteurs des SHS, les auteurs des articles sont enclins, sinon fortement encouragés par leur laboratoire, à publier dans des revues de langue anglaise à fort FI. C'est-à-dire, sur des objets, selon des paradigmes conformes, dans une perspective plus productiviste qu'inventive, valorisant la quantification et l'expérimentation, ou, plus exactement, leur simulacre.

Si bien que, faute de propositions d'articles, les revues publiant en langue française et sans facteur d'impact, sont, à terme, menacées de disparition.

Dans le domaine des SHS, pour la psychologie du moins, dans beaucoup de spécialités, les modélisations sont complexes, leur formalisation est faible, ce qui ne facilite pas la présentation d'une étude dans une autre langue.

On observe, dans certaines études, en psychologie pathologique, notamment, des tactiques, pour entrer dans la norme, tactiques qui consistent à travailler sur des données quantifiées, données de type épidémiologique, ou en référence aux classements du DSM IV, par exemple, plutôt que de travailler sur des cas de façon clinique. Ainsi, les auteurs se donnent une scientificité apparente, par l'usage d'une quantification, souvent mal ou insuffisamment fondée, et abandonnent toute analyse psychologique qualitative.

Or, sur des questions comme l'adolescence et la violence, les effets du handicap, le suicide, ou encore les apprentissages scolaires, les phénomènes étudiés ont une valence culturelle forte. De nombreux chercheurs et praticiens de la psychologie – dans le champ de l'éducation, de la santé, des entreprises – qui participent à la production des connaissances et à leur diffusion, risquent de ne plus avoir de support pour accueillir leurs travaux.

Un orateur a fait référence, lors du colloque, au ministère australien de l'Innovation, de l'industrie, de la science et de la recherche, qui a pris la décision, en mai 2011, de supprimer les classements des revues, établis sur l'avis des experts. Ces classements, selon le ministre (déclaration reproduite dans le quotidien *The Australian* du 30 mai 2011, voir Rowbotham J., 2011) conduisent, dans certains domaines de la recherche rattachés à des questions d'intérêt national (économie, démographie), à des comportements indésirables : les responsables de la recherche mettent en garde contre la publication dans des revues classées B ou C. En conséquence, des revues ont mis la clé sous le paillason. Tout ceci, selon le ministère australien, nuit à la recherche.

Voici trois ans, notre collègue Roger Brumback, rédacteur en chef du *Journal of Child Neurology*, Professeur de pathologie et de psychiatrie, faisait remarquer, dans un article publié en avril 2008, « Worshipping false idols: The impact factor dilemma » « [...] Unfortunately, over the past decade, the journal impact factor has been transformed into a singular rating to determine not just the value of journals but also the quality of scientists, institutions, and even scientific research. The journal impact factor is now being used extensively by committees and government in Europe and to a lesser extent in North America to make decisions concerning the awarding of grant funding and about the promotion and tenure of individual scientists » (Brumback R. A., 2008).⁸

En France, l'Académie des sciences, dans le rapport remis, le 17 janvier 2011 à la Ministre de l'Enseignement supérieur et de la recherche, « Du bon usage de la bibliométrie pour l'évaluation individuelle des chercheurs »⁹, met sérieusement en garde contre l'usage du FI « souvent pris en compte pour évaluer la qualité d'un article. Cette pratique, très utilisée dans certaines disciplines comme la biologie et la médecine, est dangereuse car les revues les plus prestigieuses, à facteur d'impact très élevé, contiennent un pourcentage significatif d'articles de qualité moyenne (...) ». Ce même rapport souligne la faiblesse des bases de données pour les SHS, ce que met clairement en évidence l'étude de Michèle Dassa, Christine Kosmopoulos et Denise Pumain¹⁰

8 Je traduis : « Vénération de fausses idoles : la question du facteur d'impact ». « [...] Malheureusement, depuis plus d'une décennie, le facteur d'impact a été transformé en une note spécifique pour déterminer non pas la valeur des revues mais également la qualité des scientifiques, des établissements et même la recherche scientifique. Le facteur d'impact d'une revue est maintenant en train d'être utilisé à profusion par les commissions et les gouvernements en Europe et, dans une moindre mesure, en Amérique du Nord, pour fonder les décisions d'attribution de fonds, l'emploi et la promotion des chercheurs »

9 Pour télécharger le rapport : <http://www.academie-sciences.fr/activite/rapport/avis170111.pdf>

10 Pour accéder à l'étude : <http://cybergeogeo.revues.org/22864>

Pour conclure

Il importe, en premier lieu, de souligner que l'évaluation des revues n'entre pas dans les missions de l'AERES. Rappelons que ce classement des revues de psychologie a été remplacé par une liste « périmètre », depuis mon intervention de juin 2011.

De plus, et sauf erreur, parmi les membres de la commission du domaine « Psychologie - Éthologie - Ergonomie », la présence de spécialistes de la bibliométrie ou de la scientométrie n'est pas avérée.

Pour ce qui relève des missions de l'AERES, prenons la liberté de suggérer qu'on peut très bien, facilement et de manière pertinente, évaluer une large part des productions des chercheurs, le nombre de fois où ils sont cités... sans faire ce détour par les revues. Il suffit, par exemple, de regarder dans une base combien de fois apparaît le nom d'un chercheur, combien de fois ses articles sont cités...

Que ceux qui ont la charge d'évaluer la recherche et les chercheurs portent leur attention sur les travaux mêmes qui sont publiés et non sur leurs supports.

Un répertoire de revues, conçu comme un « périmètre » peut avoir une utilité, pour nos collègues, tant évaluateurs qu'évalués – ou devant l'être –, si le titre de chaque revue est accompagnée d'un profil descriptif de son fonctionnement. *La Société française de psychologie* avait édité, sous la direction de Marie Santiago, un tel répertoire, en 1999. S'il s'agit de repérer la position d'une revue dans un ensemble, la plate-forme de comparaison en ligne, pour les revues du secteur SHS, que nous propose *JournalBase*¹¹, constitue un outil remarquable. Élaborée à partir des travaux de Dassa et Kosmopoulos (2009, 2010), cette application en ligne a mis en évidence le manque de fiabilité en SHS des bases de données les plus réputées. Veut-on comparer ces sources, et c'est leur faible recouvrement qui est dévoilé. Autrement dit, l'amateurisme des agences d'autorité, en matière de classement des revues, se conjugue à celui des sources sur lesquelles elles s'appuient.

Aujourd'hui, avec une cinquantaine de responsables de revues, associés en « Conférence des publications de psychologie de langue française », en collaboration avec la *Fédération française des psychologues et de psychologie*, et l'*Association des enseignants-chercheurs de psychologie des universités*, nous avons lancé un travail de réflexion sur la situation actuelle, avec le projet d'élaborer une charte définissant des principes auxquels adhèrent les revues signataires, pour l'information des personnes et instances concernées par la publication.

Bibliographie

Académie des sciences, *Du bon usage de la bibliométrie pour l'évaluation individuelle des chercheurs*. Rapport remis le 17 janvier 2011 à Madame la ministre de l'Enseignement supérieur et de la recherche. Pour télécharger le rapport : <http://www.academie-sciences.fr/activite/rapport/avis170111.pdf>, 67 p.

Brumback (Roger A.), "Worshiping false idols: The impact factor dilemma", *Journal of Child Neurology*, 23, 4, avril 2008, p.365-367.

Code de la recherche, 2008, "L'Agence d'évaluation de la recherche et de l'enseignement supérieur", in : Code de la recherche, Livre 1^{er}, Titre 1^{er}, chapitre IV, section 2, Article L114-3-1, consultable sur <http://www.legifrance.gouv.fr/>

Dassa (Michèle), Kosmopoulos (Christine), 2009, "JournalBase – Une étude comparative internationale des bases de données des revues scientifiques en sciences humaines et sociales (SHS)", *Cybergeos*, Dossier : Les revues en sciences humaines et sociales, mis en ligne le 25 juin 2009, URL : <http://www.cybergeos.eu/index22492.html>

¹¹ <http://journalbase.cnrs.fr>

Dassa (Michèle), Kosmopoulos (Christine), Pumain (Denise), 2010, "JournalBase. Comparer les bases de données scientifiques internationales en sciences humaines et sociales (SHS)", *Cybergeos*, <http://cybergeos.revues.org/22864>

INIST-CNRS, 2011, JournalBase, URL : <http://www.inist.fr/spip.php?article85>.

Jeannin (Philippe), 2003, *Revue de la recherche en sciences humaines et sociales*, Direction de la recherche, ministère délégué à la Recherche et aux nouvelles technologies. Pour télécharger le rapport : http://www.revues.org/cost/images/1/12/JEANNIN_-_Rapport_final2003a.pdf, 41 p.

JournalBase, <http://journalbase.cnrs.fr/>

Laurens (Stéphane), "L'étrange classement des revues de psychologie et le cas mystérieux du Bulletin de psychologie", *Bulletin de psychologie*, 62, 1, janvier-février 2009, p. 67-77.

Rowbotham (Jill), "End of an ERA: journal rankings dropped", *The Australian*, May 30, 2011 <http://www.theaustralian.com.au/higher-education/end-of-an-era-journal-rankings-dropped/story-e6frgcjx-1226065864847>

Santiago Delefosse (Marie), *Répertoire des revues francophones de psychologie. Un outil à destination des auteurs*, Paris, Société française de psychologie, 1999.

Présentation du système d'information RIBAC du CNRS

Un outil de valorisation de l'activité des acteurs de la recherche en SHS

Michèle Dassa (CNRS-Institut des Sciences Humaines et Sociales) et Isabelle Sidéra (CNRS-Laboratoire Préhistoire et Technologie)

Résumé

L'outil RIBAC est un formulaire normalisé très complet, descriptif de l'activité des chercheurs en sciences humaines et sociales. Les chercheurs le complètent en ligne via le web, avec un accès personnel et sécurisé. Les informations recueillies sont gérées dans une base de données. Elles permettent de décrire le paysage de la recherche en SHS grâce à la construction d'indicateurs relatifs à l'ensemble des activités des chercheurs : publications, enseignement, thématiques de recherche, participations aux congrès, programmes de recherche, outils employés, activités éditoriales, activités d'administration, de transfert et de valorisation ; en bref, toute action de recherche, d'enseignement et de transmission de la culture scientifique. Cet outil constitue également, pour les chercheurs qui le complètent, un outil personnel de gestion de leur activité et de leurs publications.

L'outil est présenté ici et associé à quelques exemples de résultats du traitement des données de l'enquête 2009.

Introduction

Le formulaire Ribac permet de recueillir annuellement les informations relatives à l'activité des chercheurs, via un questionnaire diffusé à chacun par le WEB. Il est accessible tout au long de l'année. Aussi, les informations relatives à l'année en cours sont-elles très vite disponibles. Le questionnaire a été récemment mis au point par les auteurs pour les sciences humaines et sociales et tout juste institutionnalisé par le CNRS en 2011. Il est actuellement géré par l'INSHS. L'un de ses objectifs est de combler les lacunes de la couverture bibliographiques des bases de données internationales pour les sciences humaines et sociales (Web of Science et Scopus) et de fournir des indicateurs de la production scientifique en temps réel ou presque. Il faut savoir qu'aujourd'hui encore les statistiques de production scientifique du CNRS ne prennent pas en compte les publications des chercheurs en SHS, faute d'outils représentatifs et fiables de recensement. Il devenait donc urgent de proposer une solution, afin de faire « exister » les sciences humaines et sociales du CNRS au même titre que les autres disciplines ; d'où l'élaboration originale de cet outil.

La conception de RIBAC a été réalisée en 2008 en collaboration entre le conseil scientifique du département SHS du CNRS (CSD) et la direction scientifique du département SHS et avec l'appui d'un comité de pilotage composé de chercheurs, d'enseignants-chercheurs et de spécialistes de l'évaluation (AERES, ANR, CoNRS, OST) (Dassa, Sidéra *et al.*, 2008). La maquette de Ribac a été testée par une dizaine de laboratoires pilotes en 2009 (Dassa, Sidéra *et al.*, 2009) avant d'être définitivement opérationnelle et mise en œuvre auprès de tous les chercheurs

des SHS ; ceci fin 2009. A la différence des outils de gestion du CNRS, RIBAC a été placé sous l'égide scientifique de Joël Bertrand, Directeur Général Délégué à la Science.

I - Conception, caractérisation et mise en œuvre de l'outil RIBAC

La construction de RIBAC a tout d'abord consisté à élaborer un questionnaire fondé sur une nouvelle typologie de l'activité, qui permette la construction d'indicateurs simples et représentatifs de l'activité. Fondées sur une connaissance empirique, les activités des chercheurs en SHS ont d'abord été listées puis ordonnées, avec l'idée de décrire, caractériser et quantifier les pratiques dans toute leur richesse et leur diversité, bien au-delà des activités strictement scientifiques. L'activité des chercheurs, en effet, ne se réduit pas à l'acquisition de données et à la publication de résultats, mais concerne une très large palette d'actions différentes : activités éditoriales, enseignement, valorisation, élaboration de collections de comparaison ou thématiques et de bases de données, etc. A des fins de traitements statistiques, RIBAC est une base de données, avec des informations normalisées au sein de champs définis. Des choix pré-établis sont privilégiés et sont présentés sous forme de cases à cocher, des menus déroulants et des listes circonscrites de mots-clés.

Les chercheurs accèdent à leur fiche personnelle via le Web selon un mode sécurisé avec un mot de passe personnalisé. Ils doivent au préalable s'inscrire avec leur adresse électronique et leur date de naissance pré-enregistrées à partir des informations présentes dans le système d'information du CNRS LABINTEL.

Le formulaire RIBAC est organisé en 4 grandes rubriques : identité et insertion professionnelle, activités de recherche et d'enseignement, activités de transfert et de valorisation et expression libre. Chaque rubrique est structurée en sous-rubriques et en sous-sous-rubriques. Elles permettent aux chercheurs de décrire finement leurs activités. Ils disposent d'un mode d'emploi accessible via le menu « documentation » pour les aider à naviguer et à saisir leurs informations (cf figure 1).

RIBAC V11.0 - Mozilla Firefox

https://www.ribac.cnrs.fr/index.php?id_form=4

Recueil d'Informations pour un observatoire des Activités des Chercheurs en SHS

Activités du 1er janvier au 31 décembre 2011

Identité et insertion professionnelle

Identité

☐ Monsieur ☒ Madame

Nom :

Nom de naissance :

Prénom :

Date de naissance :

☒ Chercheur ☐ Enseignant - Chercheur

Etablissement de rattachement : ☒ CNRS ☐ Université ☐ autre

Numéro d'agent CNRS :

Adresse électronique :

Grade :

Statut particulier :

Section du CN :

CID (commission interdisciplinaire) :

Section(s) CNU :

Remarques / commentaires

Figure 1 : Formulaire RIBAC – sous rubrique « identité ».

La base de données RIBAC a été conçue pour être évolutive et modulaire. Les besoins remontés par la communauté scientifique et validés par un comité d'utilisateurs peuvent facilement être pris en compte pour ajouter, modifier ou supprimer des zones de saisie. En 2009, les utilisateurs ont demandé d'avoir la possibilité de compléter les fiches RIBAC au fil de l'eau, tout au long de l'année. Ceci a été pris en compte dans la version 2010 de RIBAC. Grâce à cette nouvelle fonction, RIBAC est ainsi devenu un outil de mémoire et de gestion de l'activité. Les fiches produites peuvent être facilement mises à contribution pour la réalisation d'un curriculum vitae ou d'un rapport d'activité.

Ajoutons que les chercheurs sont de plus en plus sollicités pour répondre à différentes enquêtes correspondant à la diversité des objectifs d'évaluation des organismes et des tutelles. RIBAC a

donc eu dès le départ l'objectif de réduire les redondances de saisie. Par exemple, l'import des références des publications entrées dans l'archive ouverte Hal SHS a été une amélioration récemment apportée dans RIBAC, moyennant un travail d'adaptation minimal, grâce à la correspondance des champs. Il reste cependant à progresser encore pour automatiser davantage la saisie et, surtout, pour parvenir à une sorte de document unique, qui pourrait servir à plusieurs services, organismes, ou objectifs collectifs ou individuels.

A - Rubrique « compétences individuelles »

Une disparité de pratiques existe entre les chercheurs des SHS, qui induit une véritable différenciation de leurs activités. Certains vont plus que d'autres aux sources. D'autres se livrent plus particulièrement à l'édition ou encore au management de la recherche, à l'enseignement, etc. Différents profils de chercheurs se distinguent alors, qui ne sont pas toujours exclusifs, mais peuvent se succéder dans le temps ou se juxtaposer dans une carrière professionnelle. De même, les objets de recherche sont de la plus grande variabilité, des plus généralistes au plus spécialisés. Par exemple, un individu peut détenir une compétence pointue, rare ou unique, comme au contraire être expert de plusieurs domaines. S'ajoute à cela la diversité du contexte de la recherche et des outils employés, variable selon les individus et les courants de pensée. Les chercheurs se sont emparés des nouvelles technologies et se sont appropriés une vaste palette d'outils pour leur recherche, dont de nombreux issus d'autres disciplines, qui renouvelle l'environnement technique et les compétences. En parallèle, le développement de l'interdisciplinarité, l'élargissement du champ des disciplines et le renouvellement des problématiques scientifiques ont élargi les connaissances. Ceci a bouleversé les pratiques récentes de recherche et le cadre dans lequel les chercheurs exercent leur travail. Il valait donc la peine d'intégrer la diversité et les nouveautés pour caractériser les pratiques, les méthodologies, les environnements et les thématiques de recherche ainsi que leurs évolutions. Les rubriques de RIBAC ont donc pris en compte ces différents aspects dans la partie « compétences individuelles ».

Cette rubrique « compétences individuelles », est organisée avec des cases à cocher et des champs à compléter à l'aide de mots-clés libres, car le domaine d'expertise d'un individu ne peut pas toujours être traduit par des valeurs prédéfinies et arrêtées. Cependant, à côté de ces champs « libres », une liste fermée de 27 disciplines est proposée pour cadrer le domaine de recherche. Ces 27 catégories ont été traduites et adaptées de la classification de l'ESF (European Science Foundation) et sont accessibles dans la zone « grand champ disciplinaire » (Figure 2).

L'idée au travers de la cotation d'informations relatives aux thématiques de recherche et aux outils employés est aussi le repérage des expertises, de manière à ce que le CNRS puisse identifier et recommander, à la demande de journalistes par exemple ou encore pour ses besoins propres, les spécialistes les plus pointus sur une question d'actualité ou autre (figures 2 & 3).

Recueil d'Informations pour un Observatoire des Activités des Chercheurs en SHS
Activités du 1er janvier au 31 décembre 2011

RIBAC V11.0 ©CNRS-INSHS

Import de HAL Synthèse PDF Documentation

Identité et insertion professionnelle

Identité
Affectation
Période d'inactivité
Formation continue
Distinction
Qualification
Thèmes de recherche de l'unité
Mots clés des sections / Objets d'étude

Compétences individuelles

Compétences linguistiques

Activités de recherche et d'enseignement

Acquisition de données
Enseignement et encadrement
Programmes de recherche
Congrès, colloques et tables rondes
Missions dans d'autres laboratoires
Responsabilités collectives
Expertises scientifiques
Productions multimédia
Activités éditoriales
Publications scientifiques

Activités de transfert et de valorisation

Expertise / Conseil
Contrats avec des organismes privés
Contrats avec des organismes publics
Publications de transfert et de valorisation
Exposition
Presse orale
Sites WEB, Blogs, Wikis
Animation

Compétences individuelles

Grand champ disciplinaire :

Autre grand champ disciplinaire SHS :

Autre grand champ disciplinaire non SHS :

Domaine :

Période :

Zones géographiques :

Thème interdisciplinaire :

Ex : santé

Travaux en liaison avec les instituts du CNRS :

☐ INSHS ☐ INC ☐ INEE
☐ INP ☐ IN2P3 ☐ INSB
☐ INSIS ☐ INSMI ☐ INSU
☐ INS2I

Spécialité et Méthodologie :

Figure 2 : Copie d'écran « compétences individuelles » : domaines, disciplines, géographie et période.

Animation d'une association
Création d'entreprise

Remarques / commentaires

Validation finale de toutes les fiches

Profil du chercheur :

☐ Activité éditoriale
☐ Administration de la recherche
☐ Analyse statistique
☐ Constitution d'une documentation à usage collectif
☐ Direction d'un projet collectif de recherche
☐ Encadrement d'étudiants M1, M2, D
☐ Enseignement entre 30h et 60h
☐ Expérimentation
☐ Laboratoire
☐ Médiation scientifique
☐ Mise au point de méthode
☐ Modélisation
☐ Publication d'un état de l'art ou d'un travail de synthèse significatif sur un domaine
☐ Simulation
☐ Acquisition de données (enquête/fouille/dépouillement d'archive/étude de collection/bases de données)
☐ Théorie
☐ Valorisation de la recherche
☐ Autres

Outils de recherche employés :

☐ Analyses chimiques
☐ Analyse d'images
☐ Analyses physiques
☐ Analyses statistiques
☐ Bases de données
☐ Bibliothèque
☐ Calculateur
☐ Capteurs
☐ Grand équipement
☐ Logiciel spécialisé
☐ Matériel d'enregistrement d'images et de son
☐ Microscopie électronique
☐ Microscopie optique
☐ Modélisation/simulation
☐ Ordinateur individuel
☐ Outils de numérisation
☐ Référentiels ou collections de comparaison
☐ Spectrométrie
☐ Système d'information géographique
☐ Télédetection
☐ Autres

Valider

Figure 3 : Copie d'écran « compétences individuelles » : profil du chercheur et outils employés.

B - Typologie des activités

Une même activité recouvre en réalité souvent un ensemble d'actions distinctes (figure 4). Aussi, ces actions se déclinent-elles dans différentes rubriques et catégories du formulaire, en fonction de leur nature. Par exemple, une communication à un colloque est entrée dans la rubrique « congrès, colloques et tables rondes ». Si elle donne lieu à une publication, elle est aussi entrée dans la rubrique « publications », comme un article ou un chapitre d'ouvrage. Il en va de même pour les missions. Celles-ci peuvent comprendre un déplacement, une acquisition de données et une valorisation de ces données sous quelque forme que ce soit.

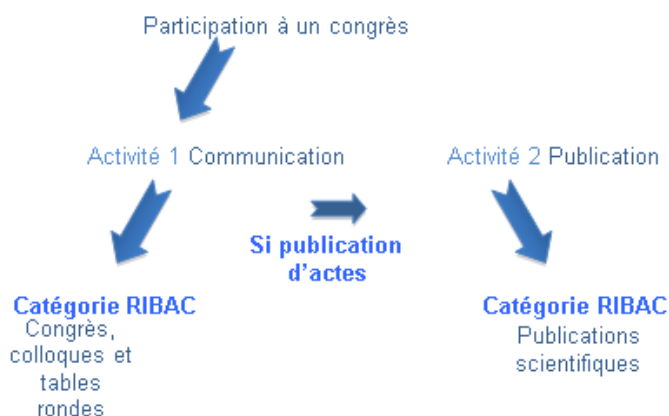


Figure 4 : Distribution des activités au sein du formulaire RIBAC

C - Utilisation des informations saisies les années précédentes

La plateforme RIBAC permet de construire en dynamique une synthèse des informations saisies chaque année par les chercheurs et d'accéder à celle des années précédentes grâce à l'onglet « synthèse PDF ». Toujours dans l'optique de faciliter le transfert des informations pour d'autres usages, un export des données en formats Word et Excel est maintenant possible (cf figure 5).

De plus, grâce à l'onglet année (2010, sur la figure 5), les informations saisies les années précédentes peuvent être aisément reportées, via une fonction copié-collé, sur le formulaire en cours. Ceci est intéressant pour les activités récurrentes d'une année à l'autre, tel l'enseignement, les activités éditoriales ou les responsabilités de tout ordre.

Tel que cela a déjà été souligné plus haut, l'allégement de la saisie et le recyclage des informations sont l'objet d'un souci constant et d'un travail régulier.

Figure 5 : copie d'écran du formulaire RIBAC

II - Résultats des premières enquêtes RIBAC

A - Taux de participation

La première enquête RIBAC lancée en grandeur réelle auprès des chercheurs CNRS des unités SHS a eu lieu à l'automne 2009. La même année, les chercheurs SHS avaient été invités à renseigner simultanément le nouveau système d'information et le compte rendu annuel d'activité officiel « CRAC ». Ils devaient donc effectuer une double saisie. Malgré ce surcoût de travail, 1006 chercheurs, soit 51 % d'entre eux ont participé à l'enquête RIBAC 2009. Ceci laisse apprécier le soutien de la communauté (cf figure 6).

En 2010, par décision de la Direction du CNRS, le choix a été laissé aux utilisateurs de compléter soit le formulaire CRAC soit le formulaire RIBAC. Ils ont choisi RIBAC à 59 % (1100 chercheurs). Le taux de participation a donc été légèrement supérieur en 2010 par rapport à 2009 (cf figure 6).

La participation aux enquêtes 2009 et 2010 a été assez homogène en fonction des sections du comité national, mais au-dessus de la moyenne pour les sections 32 (Mondes anciens et médiévaux) et 38 (Anthropologie, Mondes contemporains), où elle se situe entre 60 et 66 % (cf figure 6).

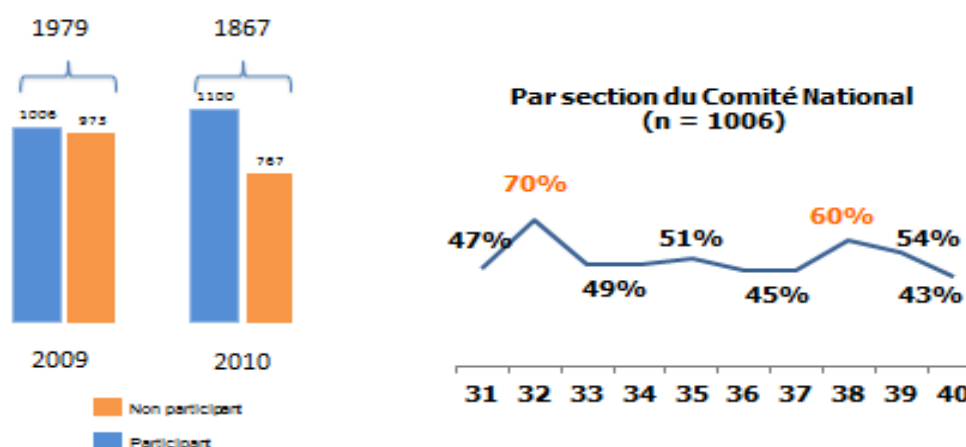


Figure 6 : Participants aux enquêtes RIBAC 2009 et 2010

En 2011, pour la 3^{ème} année consécutive, la fiche RIBAC est proposée aux chercheurs CNRS de l'INSHS. Mais, suite à la décision de Joël Bertrand, Directeur général délégué à la science, en concertation avec l'ensemble des acteurs du CNRS (Direction des Ressources Humaines, Comité National et Conseil Scientifique d'Institut), RIBAC devient le compte rendu annuel officiel des chercheurs du CNRS évalués par les 10 sections du Comité national SHS (n^{os} 31 à 40).

B - Quelques exemples d'indicateurs issus de la campagne 2009

a - Les langues utilisées

La diversité linguistique est l'une force des SHS par rapport aux autres communautés scientifiques, qui utilisent surtout l'anglais. Aussi, cet attribut a-t-il fait l'objet d'une attention particulière au sein du questionnaire. Les langues de communication dans les congrès, colloques et tables rondes, ainsi que celles des publications sont indiquées.

Le dépouillement de 1015 articles de revues déclarés en 2009 par des chercheurs évalués par une section du comité national SHS fait ressortir l'emploi de 15 langues différentes. Il montre également l'effort important des SHS concernant la publication en anglais, qui représente un taux non négligeable de 31,5 % des publications (figure 7).

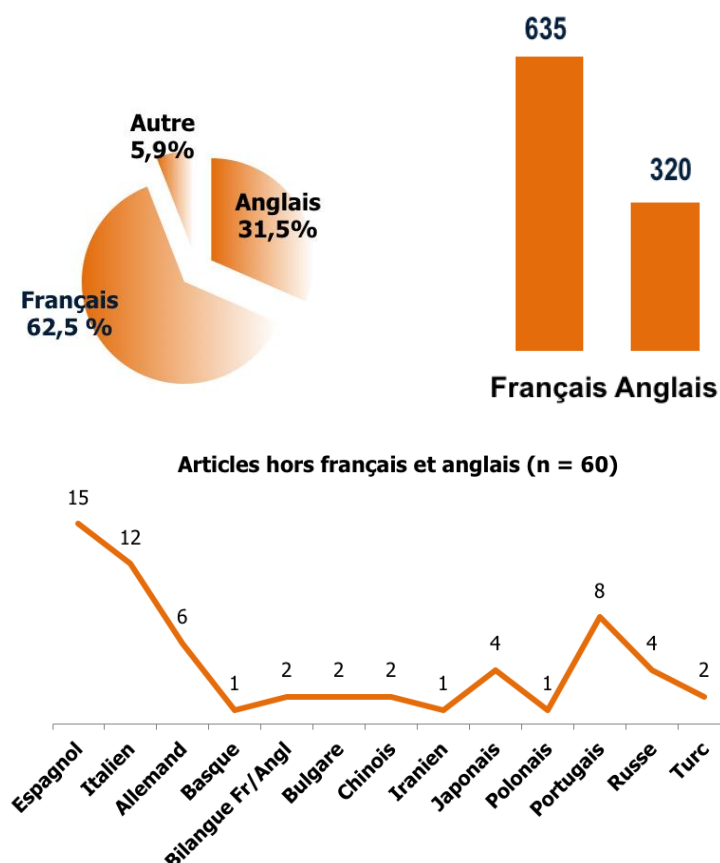


Figure 7 : Langues des articles de revues déclarés en 2009 dans les 10 sections SHS (n = 1028).

b – Les publications

Dans RIBAC, les publications sont classées selon leur support éditorial, proposant une typologie de 11 catégories descriptives : ouvrages, directions d'ouvrages collectifs, chapitres d'ouvrage, articles avec ou sans comité de lecture, comptes rendus, notices, traductions, lexiques, thèses et HDR, brevets et rapports non publiés. Un article publié dans les actes d'un colloque est considéré comme un « chapitre d'ouvrage » et distingué d'un « article de revue », inscrit dans la rubrique « article ». Ce classement neutre, car indépendant de toute évaluation, permet de compter les publications de l'ensemble de la communauté SHS à l'aune de critères communs.

Soulignons que la communication orale dans un colloque n'est pas comptabilisée comme une publication. Cette rubrique ne recense que l'écrit.

L'examen de la répartition des trois catégories principales de publication - chapitre d'ouvrage, articles de revues et ouvrage – au sein des sections du Comité National, fait apparaître trois comportements distincts de publication en fonction des sections.

- 1) Les chapitres d'ouvrage dominent la publication : sections 32, 33, 35, 36 et 38.
- 2) Les articles de revues sont majoritaires : sections 34, 37 et 39.
- 3) Chapitres d'ouvrage et articles sont en nombre équivalents ou à peu près : sections 31 et 40 (figure 8).

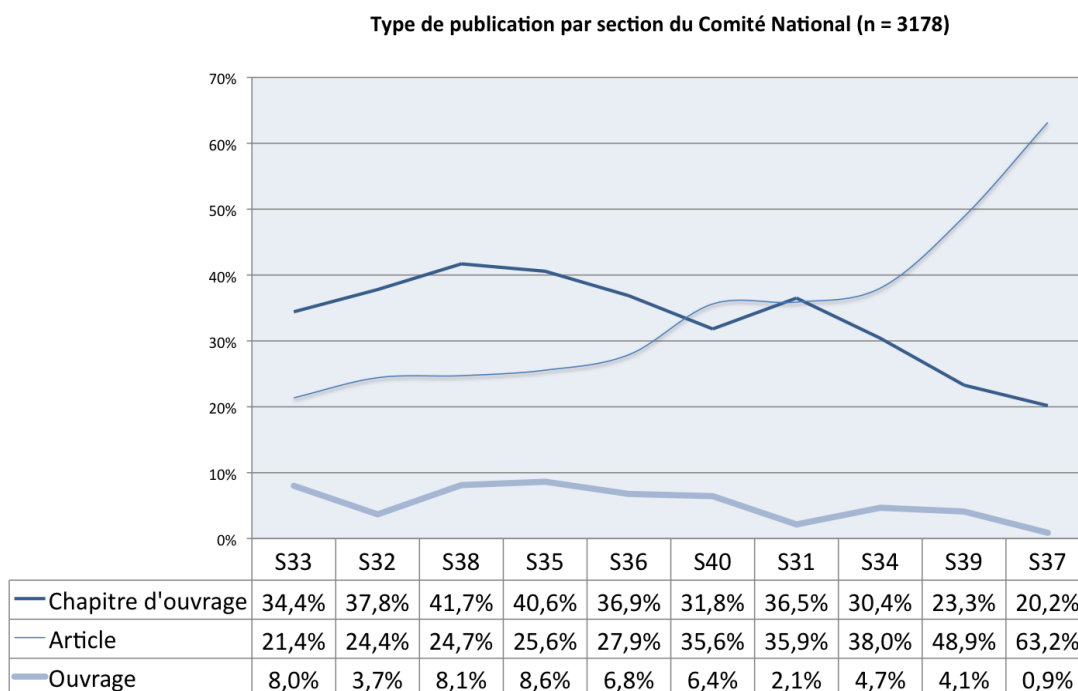


Figure 8 : Publications scientifiques par sections du Comité National (n = 3178).

Par ailleurs, on observe une grande dispersion des publications des chercheurs SHS par rapport aux supports, puisque les 1028 articles déclarés en 2009 ont été publiés dans 693 revues différentes. Ajoutons que l'ouvrage reste un support d'expression scientifique important et incontournable pour les SHS.

Si l'on poursuit l'enquête, l'examen montre que 55 % des articles publiés par les chercheurs de l'INSHS se trouvent parmi les listes de revues 11000 revues que recense *JournalBase* issues du Web of Science, Scopus et les listes d'autorité ERIH et AERES (journalbase.cnrs.fr/, Dassa M, Kosmopoulos C et Pumain D, 2010).

Enfin, 49 % de ces articles sont publiés dans une revue française contre 51 % dans une revue non française, dont 13 % dans une revue américaine et 11 % dans une revue anglaise (cf figure 9).

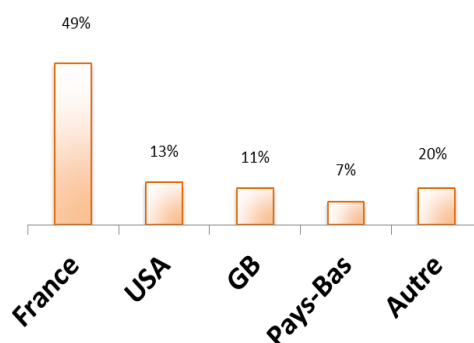


Figure 9 : Répartition des articles par pays d'édition – enquête 2009

Les articles publiés dans des revues non répertoriées dans *JournalBase* sont des articles juridiques¹, des articles publiés dans des revues non SHS (mathématiques, informatique) ou encore dans des revues SHS non recensées dans les listes prises en compte (WOS, SCOPUS, ERIH et AERES).

Les fonctions éditoriales définies dans RIBAC (éditeur, responsable scientifique, membre d'un comité de lecture), montrent une forte implication dans l'édition puisque 437 chercheurs ont déclaré une fonction éditoriale (43 %). Cette fonction concerne 515 revues différentes (dont près de 50 % répertoriées dans *JournalBase*).

Conclusion

A terme, l'enquête RIBAC pourrait être élargie à l'ensemble des chercheurs du CNRS. Dans cet objectif, l'INSHS est actuellement en relation avec d'autres services du CNRS pour construire une nouvelle application qui remplacerait le rapport d'activité annuel CRAC et serait adressée à l'ensemble des instituts du CNRS.

Il est également souhaitable que l'enquête RIBAC soit étendue à l'ensemble des acteurs de la recherche : enseignants-chercheurs et ingénieurs. En effet, les travaux de ces acteurs de la recherche ne sont pas toujours pris en compte dans les systèmes actuels de recensement ou d'évaluation, et leur contribution à la recherche ainsi qu'à l'enseignement et à toutes les autres activités que recense RIBAC, est méconnue. Cet élargissement permettrait d'avoir une vision exhaustive des activités liées à la recherche en Sciences Humaines et Sociales en France et de mieux cerner l'implication et la contribution des SHS ainsi que ses spécificités institutionnelles.

Comme cela a déjà été pointé et montré de nouveau dans d'autres communications (Dassa, Kosmopoulos *et al.*, ce volume), les bases de données bibliométriques actuelles WOS ou SCOPUS sont loin d'être exhaustives, notamment en ce qui concerne les Sciences humaines et sociales. Il en résulte une très large sous-estimation de la production des chercheurs et enseignants-chercheurs français dans ces outils commerciaux. Le système d'information RIBAC permet de dresser rapidement un panorama réaliste des publications de l'année, au plus près de la réalité de l'activité des chercheurs en SHS. Les listes de revues dans lesquelles les chercheurs SHS publient seront aussi aisément disponibles.

Il convient pour conclure de souligner le caractère valorisant du système d'information RIBAC pour la recherche en sciences humaines et sociales, puisqu'il permet de mettre en avant toute l'activité des chercheurs, non pas seulement leurs publications. Ajoutons également un élément important de ce système, qui n'a pas été créé pour juger, mais pour décrire et compter certains éléments utiles pour aligner les SHS sur le même plan que les autres disciplines scientifiques. RIBAC rassemble des faits relatifs à l'activité. Il ne contient aucun élément lié à la qualité de la recherche menée. Celle-ci reste du ressort de l'évaluation et d'une structure *ad hoc*.

Bibliographie

¹ la liste des revues « droit » de l'AERES n'a pas été intégrée dans *JournalBase* 2010.

Dassa M., Sidéra I., coll. PESCHANSKI D. et REGNIER P., 2008, *Caractérisation et quantification de l'activité individuelle des chercheurs SHS du CNRS. Préparation à l'évaluation*. Travaux du Conseil scientifique de département SHS du CNRS et du Département SHS du CNRS, 28 p. <http://www.cnrs.fr/inshs/recherche/evaluation-production-scientifique/introduction.htm>

Dassa M., Sidéra I. & Régnier P. 2009, "Mise en place d'une veille scientifique sur les SHS : Caractérisation et quantification de l'activité des chercheurs en sciences humaines et sociales", *Lettre d'information de l'INSHS* 3, p.5.

Dassa M., Sidéra I. & Régnier P. 2009, "Phase de test du RIBAC – Recueil pour un observatoire des activités des chercheurs", *Lettre d'information de l'INSHS* 4, p.9.

Dassa M., Sidéra I. & Régnier P. 2009, "Nouvelles de l'observatoire des activités en SHS – INSHS. Test du Recueil pour un observatoire des activités des chercheurs en SHS" (RIBAC), *Lettre d'information de l'INSHS* 6, p.8-9.

Dassa M., Sidéra I., avec la coll. de Régnier P. & Lievin G., 2009, "Ribac : un outil pour caractériser et quantifier l'activité des acteurs de la recherche en SHS". ADBS Association des professionnels de l'information et de la documentation, 46 (4), p. 67-68.

Dassa M, Kosmopoulos C et Pumain D., 2010, "JournalBase : Comparer les bases de données scientifiques internationales en sciences humaines et sociales" *Cybergeog*, 2010, <http://cybergeog.revues.org/index22864.html> (version française et anglaise).

L'Évaluation en sciences humaines et sociales : Comment mesurer ce qui compte ?

Marie-Claude Maurel (Directrice d'étude à l'EHESS)

*« Mesurer ce qui compte plutôt que compter ce qui peut être facilement mesuré »
(Rapport AUBR, Mackiewicz, 2009)¹*

Résumé

De nouvelles procédures d'évaluation ajustables aux SHS ont émergé ces dernières années. Les principes et les objectifs de l'évaluation doivent être clairement identifiés si l'on veut asseoir la légitimité de la démarche. Que veut-on évaluer ? Qui veut-on évaluer ? Qui est en mesure d'évaluer ? Autant de questions qui invitent à relier les dispositifs et les procédures afin de maîtriser « l'inflation évaluative ».

Comment mesurer ce qui compte, tel est l'enjeu de l'évaluation, que celle-ci s'applique à l'activité scientifique du chercheur individuel ou à celle du collectif de recherche. Cette contribution a pour objet de développer quelques éléments de réflexion sur des procédures d'évaluation en constante expansion dans le système d'enseignement supérieur et de recherche. Elle se veut ni un plaidoyer pour les outils d'évaluation existants ni une défense des indicateurs bibliométriques. L'intention est tout à fait autre. Dans l'esprit qui a animé les travaux du Conseil pour le développement des humanités et des sciences sociales (CDHSS)², l'accent est mis sur les finalités et les modalités de l'évaluation, sur le « bon usage de la bibliométrie pour l'évaluation individuelle des chercheurs », pour reprendre le titre du rapport de l'Académie des sciences et la proposition qu'il contient³.

A quelles fins évaluer ?

Une exigence de politique publique

L'évaluation est devenue une exigence permettant de mesurer les effets d'une politique publique afin de mettre les moyens consentis en regard des attentes de la société. Le financement de la recherche a besoin d'être fondé sur des systèmes d'évaluation et des indicateurs de performance de manière à justifier la meilleure allocation possible des ressources. S'agissant de l'activité de production scientifique, l'amélioration de la performance répond à un enjeu d'importance stratégique pour le développement économique social et culturel. La qualité du système d'enseignement supérieur et de recherche est en effet reconnue comme un facteur déterminant de la compétitivité globale, au plan

¹ Citation extraite du rapport « *Assessing Europe's University Based-Research* » préparé par le groupe d'experts nommé par la DG recherche de la Commission européenne, 2009.

² Pour des sciences humaines et sociales au cœur du système d'enseignement supérieur et de recherche, Rapport remis à Valérie Pécresse, Ministre de l'Enseignement supérieur et de la Recherche, octobre 2010. Voir en particulier le chapitre 5, L'évaluation en sciences humaines et sociales (77-94).

³ Du bon usage de la bibliométrie pour l'évaluation individuelle des chercheurs, 17 janvier 2011, <http://www.academie-sciences.fr/activite/rapport/avis170111.pdf>

international. C'est dans ce contexte que les universités, les organismes de recherche et leurs personnels sont soumis à une évaluation multidimensionnelle (multicritères) impliquant la prise en considération de l'ensemble des missions qui leur sont confiées. L'objectif est de promouvoir la qualité de l'activité des enseignants-chercheurs et des chercheurs et de baser sa reconnaissance sur des principes qui s'imposent à tous. La sélection de l'excellence individuelle et/ou collective constitue un volet spécifique de cette politique qui participe à l'objectif d'ensemble. Qui fait ou qui produit de la recherche en SHS en France (les « producteurs » au sens défini par l'AERES), pour quels résultats et à quelles conditions ?

Les responsables de la politique d'enseignement supérieur et de recherche n'accordent-ils pas une confiance imméritée à cette évaluation nouvelle manière, qui trouve sa source d'inspiration dans le « *management* par la qualité » ? L'adoption de nouveaux outils et de nouvelles méthodes de l'évaluation ne risque-t-elle pas de se combiner à un esprit de système – qui finirait au demeurant par rendre l'évaluation inopérante ?

Une fonction inhérente à l'activité scientifique

L'acte d'évaluer est une réalité ancienne. En tant que fonction inhérente à l'activité scientifique, il participe à la validation des savoirs, en ce qu'il contribue à la définition des critères relatifs aux formes et au degré de véracité de ceux-ci. L'acte d'évaluer se réfère à des principes d'action et/ou à des objectifs qui doivent être énoncés à l'avance, et par rapport auxquels sont mesurés les écarts de réalisation des actions. Le respect de ces principes permet de fonder la légitimité de l'acte d'évaluer. Les objectifs de l'évaluation doivent être définis et appréciés en relation avec de multiples facteurs. Les dispositifs d'évaluation sont à mettre en cohérence avec les objectifs qui ont été assignés aux chercheurs, tels que produire des connaissances nouvelles, développer des pratiques de recherche interdisciplinaires et intégrées, etc. Enfin, il s'agit de relier les dispositifs et les procédures mises en œuvre aux objectifs et de manière plus générale aux attentes.

Comment évaluer ?

Dans ce qui suit, on gardera présent à l'esprit les raisons pour lesquelles les activités de recherche et de formation à la recherche sont soumises à évaluation, pour cibler la réflexion sur les procédures adoptées pour ce faire. C'est la question du « comment évaluer ». Cette question se pose avec d'autant plus d'acuité lorsque sont appliquées de nouvelles modalités d'évaluation et introduits des outils d'évaluation rénovés. Plusieurs risques sont mis en avant : la lourdeur du processus et son caractère répétitif ; la crainte qu'il ne soit pas conduit de manière impartiale ; le danger de déséquilibrer les procédures de recherche, un point particulièrement sensible en sciences humaines et sociales.

Pour comprendre ces craintes qui affectent la communauté universitaire, il convient de poser la question de l'évaluation dans sa globalité, c'est-à-dire en prenant en compte non seulement le statut, la composition, les missions et les modalités de fonctionnement des instances d'évaluation, en particulier en sciences humaines et sociales, mais également les pratiques et les critères propres aux divers champs disciplinaires⁴.

⁴ L'analyse des craintes des enseignants-chercheurs concernant les procédures nouvelles d'évaluation ont été recensées dans le rapport ci-dessus mentionné, p.80.

A quelles conditions l'évaluation peut-elle être adéquate et efficace ?

Le respect de quelques principes devrait permettre de mieux répondre aux critères d'adéquation et d'efficacité. Ces principes consistent à :

- impliquer les chercheurs et les enseignants-chercheurs dans la préparation des modalités d'évaluation (par un véritable processus de négociation sociale)
- exiger la transparence et l'objectivité à la fois des procédures d'évaluation et des résultats ;
- assurer la définition explicite des indicateurs et des critères et introduire une pondération de ceux-ci ;
- garantir une stricte séparation des pouvoirs entre évaluateurs et décideurs ;
- assurer la pertinence et l'efficacité du processus.

Comment ces principes sont-ils mis en œuvre ?

Les régimes d'évaluation⁵

Au cours de la décennie écoulée de nouveaux dispositifs et de nouvelles procédures ont été mis en place dont les effets sur l'activité pédagogique et scientifique des enseignants-chercheurs ne sont pas sans incidence. Des analyses récentes parlent de « régimes d'évaluation » et distinguent une évaluation collégiale qui est le fait de jurys, de commissions, d'une part, et un nouveau régime qui a recours à des experts spécialisés dans l'évaluation, d'autre part. Les règles du premier régime, de tradition ancienne, sont bien connues des chercheurs qui soumettent leurs travaux à l'appréciation de jurys. Ce mode d'évaluation requiert la construction endogène de référents communs définis par les pairs siégeant dans l'instance d'évaluation (sections du CNU, du Comité national de la recherche scientifique, etc.). Le second mode, d'apparition plus récente, se présente comme un régime d'objectivité, fondé sur des outils d'évaluation (grilles, indicateurs bibliométriques, système de notation). L'exemple donné à l'appui de ce nouveau régime est l'Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES). Créée en 2007, l'AERES est un dispositif transversal d'évaluation prenant appui sur un corpus d'outils standardisés applicable à tous les secteurs de la recherche. L'agence évalue les unités de recherche, les formations et les établissements⁶. Si les principes énoncés par l'AERES et qui ont trait à l'impartialité, la fiabilité, la transparence des évaluations représentent une avancée, il est clair que chacun de ces régimes a ses qualités et ses défauts (ou plutôt ses limites).

Ceux qui caractérisent l'évaluation par les pairs sont souvent dénoncés. Si les pairs sont théoriquement bien placés pour apprécier l'originalité de la pensée scientifique, le caractère innovant du projet, le rayonnement national et international, l'impartialité de leur jugement peut être compromise par la fréquence excessive de ce type d'exercice, la subjectivité, les conflits d'intérêt ou encore les effets de groupe. Sous l'influence de quelques leaders certains champs scientifiques peuvent se réguler quasiment en vase clos.

L'évaluation fondée sur l'outil bibliométrie a aussi des limites. Les écueils procèdent du choix opportuniste des références, et du support de publication. Les contextes disciplinaires et les usages les spécifiant sont bien différenciés. Enfin, le choix des indicateurs et celui des bases de données

⁵ On doit le développement de cette notion à Nicolas Dodier, « Penser un régime d'évaluation de la recherche scientifique », 2009, <http://evaluation.hypotheses.org/452>

⁶ L'AERES met en place des comités de visite ad hoc des unités de recherche qui se livrent à un examen des activités à partir de grilles, de vérification d'indices, de critères de bonne gouvernance, etc.

pour les construire constituent un biais non négligeable. Le bon usage des indicateurs a besoin d'être défini, encadré et respecté.

Les indicateurs issus de la bibliométrie

Sur le plan méthodologique des avancées importantes ont été réalisées pour utiliser voire élaborer des indicateurs bibliométriques. Elles sont le fait de diverses instances nationales (comité national, AERES, ANR, etc.), européennes (ESF, ERC, etc.) et internationales (par exemple aux Etats-Unis, en Australie).

Beaucoup reste à faire pour améliorer les bases de données existantes, notamment afin qu'elles prennent mieux en compte les travaux en sciences humaines et sociales. A cet égard, il conviendrait de prendre en considération le spectre complet des recherches (fondamentales et appliquées), ainsi que les contextes régionaux, nationaux, et les publications dans les langues autres que l'anglais. La bibliométrie internationale reste fortement dépendante des indices de citations existants (Web of science, (Thomson Reuters), Scopus (Elsevier), Google Scholar, le logiciel Publish or Perish d'Ann-Wil Harzing qui fonctionne à partir de Google Scholar.

Que mesurent les indicateurs ? Peuvent-ils être considérés comme des guides de la qualité des travaux de recherche ? Peuvent-ils être considérés comme robustes ? On observera que le nombre de citations dont bénéficie un article ne garantit pas *ipso facto* sa qualité. Il peut être lié à un effet de notoriété ou de mode concernant un objet d'étude. Il dépend aussi de la taille de la communauté scientifique concernée.

Les outils connexes du type classement des revues représentent une autre source de catégorisation des publications. Les caractéristiques des diverses listes de classement font l'objet d'autres contributions dans le cadre des Actes de ce colloque, auxquelles le lecteur se rapportera. Bornons-nous ici à proposer de pondérer la logique bibliographique par une logique scientifique, à conseiller de ne pas surestimer les textes banals publiés dans de bonnes revues aux dépens de textes intéressants publiés dans des supports moins conformes. La qualité plus scientifique d'un article n'est pas fonction de la qualité du support et un *impact factor* élevé d'une revue ne devrait pas avoir d'effet sur tel ou tel article.

Quels usages fait-on de ces outils bibliométriques ? Remarquons qu'ils sont habituellement utilisés pour établir une catégorisation hiérarchisée de la qualité scientifique, mais aussi comme outil de calcul en vue d'établir des classements, enfin comme instrument de comptage, par exemple pour établir le taux des « producteurs » d'une institution.

Quand et comment utiliser la bibliométrie ? L'outil peut être recommandé pour évaluer l'impact de la production scientifique d'un chercheur confirmé (un senior ayant une relativement longue carrière), mais il n'est pas approprié dans le cas d'un chercheur junior. L'outil bibliométrique peut être utile pour effectuer un tour d'horizon de candidats à l'échelle d'un jury mono disciplinaire mais dans le cas d'un jury pluridisciplinaire il ne doit l'être que si les experts sont avertis des différences entre les disciplines. Dans tous les cas de figure, la règle doit être de ne comparer que ce qui est comparable (le chercheur doit pouvoir être positionné par rapport à la distribution statistique des chercheurs de sa communauté).

Enfin, l'évaluation bibliométrique ne doit pas conduire à négliger l'originalité et l'innovation qui restent plus difficilement mesurables.

Comment évaluer l'excellence ?

Élément de labellisation ou de reconnaissance, l'excellence scientifique ne se décrète pas, elle doit être identifiée. La question de l'excellence académique s'inscrit dans la problématique plus générale de l'évaluation : comment juger de l'intérêt et de la qualité d'une recherche scientifique et selon quels critères en rendre compte ⁷ ? Définie comme une appréciation de la qualité de la recherche, l'excellence scientifique ne peut être tenue pour une valeur absolue en soi mais comme une notion relative et donc l'expression d'un classement (*ranking*). Pour revêtir du sens, le classement des « meilleurs » (chercheurs ou projets) doit être établi à l'issue d'une procédure d'évaluation dont les fondements méthodologiques et les règles sont fixés à l'avance, et respectés par les experts. La reconnaissance de l'excellence académique est l'aboutissement d'une compétition inscrite dans un cadre normatif. Les performances d'un chercheur ou d'une équipe ne peuvent être mesurées exclusivement à partir d'indicateurs bibliométriques, même si ceux-ci peuvent apporter des éléments d'appréciation. L'identification de l'excellence relève d'une construction nécessairement collective et relationnelle dans le cadre d'un mode d'évaluation par les pairs.

Pour illustrer notre propos, la procédure de sélection-évaluation de l'IUF qui prend appui sur une vingtaine d'années d'expérience et autant de promotions de lauréats peut être analysée. La mission d'évaluer, de comparer et de traduire en classement les candidats est confiée à deux jurys (junior et senior) indépendants et souverains, nommés par arrêté ministériel sur proposition d'une commission de désignation. Les lauréats sont sélectionnés au terme d'une évaluation exigeante par ces jurys qui comportent au moins deux cinquièmes d'universitaires étrangers. La pertinence et la faisabilité du projet présenté par le candidat interviennent dans l'appréciation émise par les jurys selon des critères comparables à ceux utilisés par les instances d'évaluation telles que les comités de l'ERC. Les dossiers des candidats sont évalués par deux rapporteurs, spécialistes de la discipline, les avis font ensuite l'objet d'une discussion en sous-jurys et d'une délibération en séance plénière pour établir le classement final. Les jurys sont pluridisciplinaires, il n'y a pas de fonctionnement par sous-groupe étroitement disciplinaire, mais seulement deux sous-jurys (sciences de la matière, de la nature et de la vie, d'une part, sciences humaines et sociales, d'autre part). La difficulté réside dans la diversité des modes d'évaluation de la production scientifique selon les disciplines. La grille d'évaluation, inspirée de celle de l'ERC, est commune mais il est recommandé aux jurés de tenir compte des standards de la discipline. Il est très ambitieux de vouloir juger de l'excellence indépendamment des cadres disciplinaires. Le croisement des regards disciplinaires pratiqué au sein des jurys présente l'avantage de permettre de mieux évaluer l'intérêt des recherches interdisciplinaires et de celles qui ouvrent de nouveaux fronts de la connaissance. Un tel mode de sélection repose d'abord sur la qualité des pairs formant le jury. Le choix se fonde sur un certain nombre de critères : parcours académique de haut niveau, indépendance d'esprit, aptitude à évaluer l'originalité et l'innovation, sens de l'équité... Le bon fonctionnement du jury requiert de chacun de ses membres le respect des règles informelles d'évaluation, de la souveraineté de chaque discipline, ainsi que d'une capacité à délibérer.

⁷ Voir l'ouvrage de MICHÈLE LAMONT, *How Professors Think: Inside the Curious World of Academic Judgment*, Harvard University Press, 2009, ainsi que les actes du Symposium/Thinking Academic Evaluation after Michèle Lamont's *How Professors Think* - Edited by Anna Carola Freschi and Marco Santoro, *SOCIOLOGICA*, 3/2010

Conclusion : Comment améliorer l'évaluation de la recherche en sciences humaines et sociales ?

Soulignons la complexité du processus d'évaluation, sa lourdeur (la fréquence et le caractère répétitif des évaluations). L'évaluation se combine, en France, à une rationalité bureaucratique centralisée risquant de donner à l'évaluation une emprise excessive qui finira d'ailleurs par la vider de son contenu. Il n'est pas de problème qui échappe à l'évaluation et l'autonomie des universités peut renforcer ce mécanisme par le développement d'évaluations internes à chacune d'elles : comités d'évaluation pour les promotions, comités « qualités » qui se mettent en place, audits internes formés par les universités et par les PRES, appels d'offre des investissements d'avenir.

L'adhésion de principe à l'évaluation ne devrait pas conduire à ignorer cette inflation évaluative dont on sait qu'elle conduira nécessairement à une routinisation des procédures et à une augmentation continue des « coûts de transaction ». On risque de répondre à chaque problème par une nouvelle évaluation.

L'agence d'évaluation repose sur le principe d'une séparation du jugement et de la décision, de l'expertise et de la politique. C'est excellent. Mais dans ce cas, il faut souhaiter que les décideurs ne se défaussent pas en créant de nouvelles évaluations en aval ou de manière parallèle. Que les experts expertisent et que les décideurs décident sans se cacher derrière de nouvelles évaluations par les universités, les PRES, les régions, les comités de programmes... Réfléchissons sur la dérive évaluative, non pas pour nous opposer à l'évaluation, mais pour en défendre le principe et l'efficacité.

Pour ce qui est de l'évaluation elle-même, le grand risque reste celui d'un usage mécanique et non raisonné de la bibliométrie : trop serrée, elle exclut, trop large, elle se vide de sens, et surtout, elle formate les modes d'expression des chercheurs les plus jeunes. Les travaux de longue durée (l'écriture d'un livre), la réflexion conceptuelle, l'imagination créatrice, sont autant de risques que les jeunes chercheurs sont réticents à courir en raison des pratiques de l'évaluation. Au final, il convient de plaider en faveur d'un mode d'évaluation hybride intégrant les bénéfices du jugement par les pairs et l'usage réfléchi et critique des outils d'évaluation permettant d'établir des notations objectives.

ERIH: Dix ans après

(European Index for the Humanities)

Colloque sur l'évaluation en sciences humaines et
sociales

Gretty M. Mirdal
CNRS, Paris
Vendredi 10 juin 2011

Plan de ma présentation:

- Le début des années 2000: Pourquoi une nouvelle classification?
- Réflexions sur la nécessité de créer un nouvel outil
- Choix d'une méthode apparemment valable: l'évaluation par les pairs
- Les listes initiales
- Critiques positives et négatives après la première publications des listes
- Révisions des listes initiales
- La nouvelle classification: Les « NATS » et les « INTS »
- Bons et mauvais usages de l'ERIH

Pourquoi une nouvelle classification?

Les sciences humaines – et surtout les revues européennes de ces domaines - étaient très mal représentées dans les classifications internationales.

Même dans les bases de données plus spécifiques comme par exemple PsychINFO, seulement un tiers des revues francophones y figuraient. La situation était pire pour d'autres domaines et d'autres langues moins répandues.

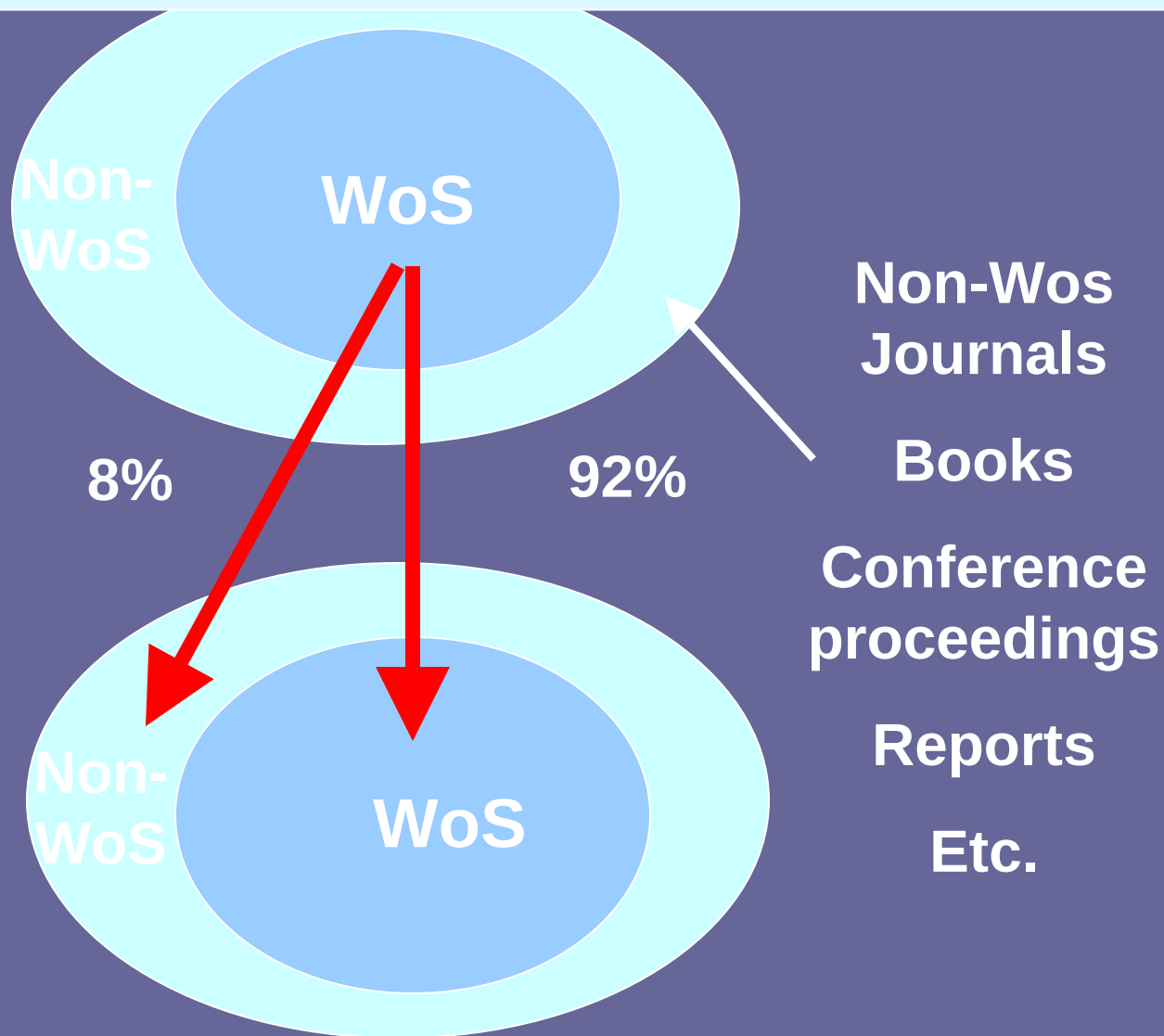
Measurement of internal WoS Coverage for Molecular Biology

(Moed, 2005)

Citing/Source



Cited/Target

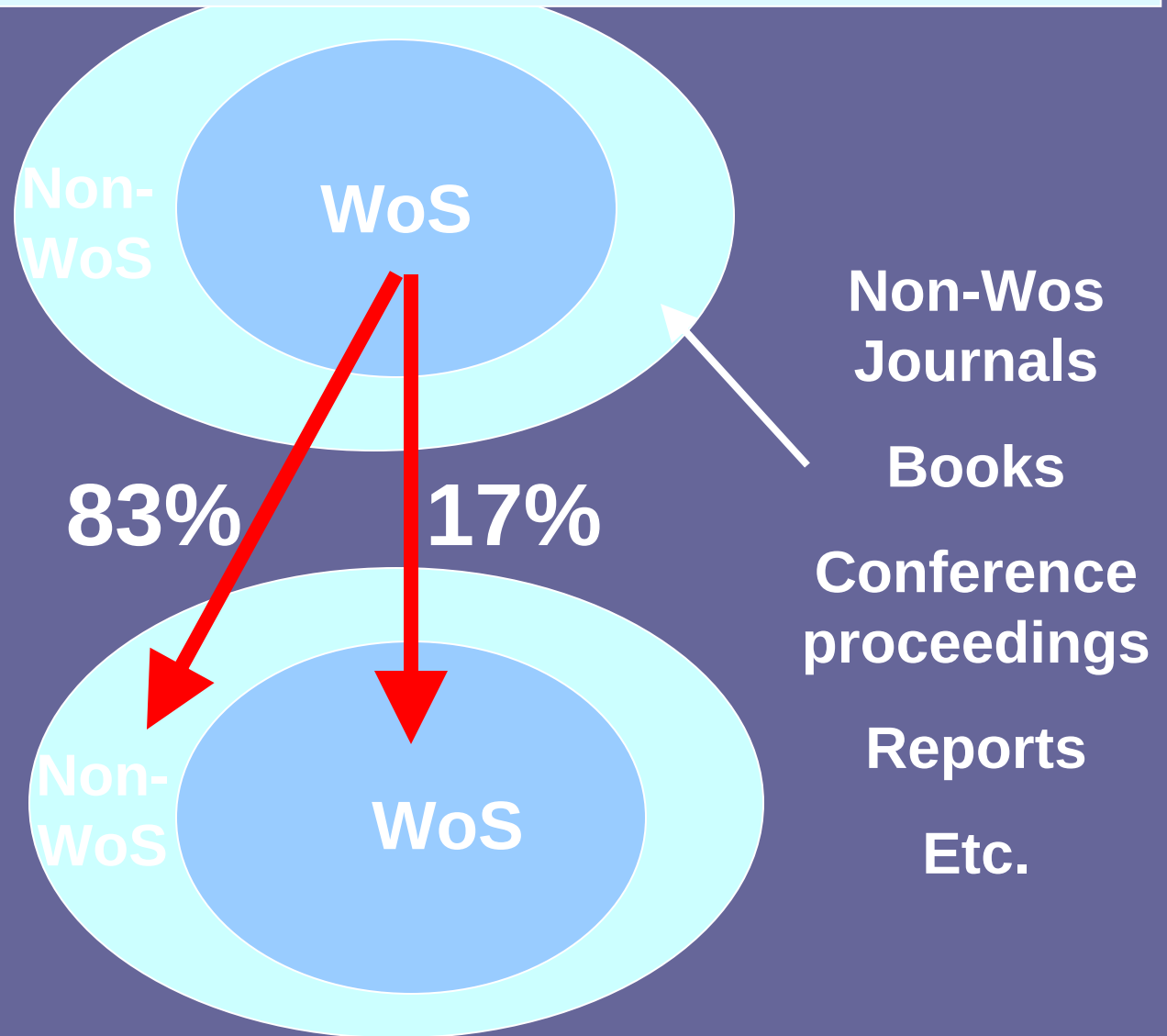


Measurement of internal WoS Coverage for the Humanities

Citing/Source



Cited/Target



Partout la même préoccupation:

"Comment assurer de solides fondements à une évaluation de la recherche européenne en SHS?"

Réunion de l'ESF à Budapest, 2001

CS du CNRS, juillet 2002:

Appréhension et méfiance

Malgré une grande méfiance envers la bibliométrie, il y avait consensus dans la plupart des organisations de recherches dans toute l'Europe, qu'il était nécessaire de développer un "Reference Index" qui pourrait présenter, mettre en vue et faciliter l'accès aux publications scientifiques européennes de haute qualité.

L'ESF Standing Committee for the Humanities prit donc la responsabilité de développer cet instrument qu'on décida de nommer **European Reference Index for the Humanities (ERIH)**.

Alain Peyraube

Je tiens à souligner le rôle primordial qu' **Alain Peyraube**, directeur de recherches au CNRS, a joué dans la création de l'ERIH depuis le début des années 2000 et tout au long du processus qui a enfin abouti à la diffusion des listes révisées en Mai 2011. C'est Alain Peyraube qui a introduit l'idée à l'ESF et lui qui a présidé et continue à présider son comité de pilotage.

Description des méthodes

Les agences de recherche, agences de moyens, associations disciplinaires et centres de recherches spécialisées et certaines académies furent contactés.

On leur demanda d'identifier les revues de bonne qualité dans tous les domaines des sciences humaines et dans tous les pays représentés à la Fondation Européenne de la Science.

Rétrospectivement , quand on considère l'envergure de l'exercice, et les moyens extrêmement limités de l'ESF, on se rend compte de la témérité de l'entreprise, mais aussi de l'engagement qui l'a soutenue.

Dans une première phase, il était nécessaire d'identifier les disciplines principales et de former des comités d'experts (expert panels) dans chacune de celles-ci, afin d'évaluer les 14,000 revues proposées par les organisations membres de l'ESF, c'est à dire les organisations de recherche qui constituent l'ESF, soit 78 "Member Organisations" dans 30 pays différents.

.

Nous savons aujourd'hui que plusieurs agences et plusieurs associations qui auraient dû être consultées,
-n'ont pas reçu la demande de l'ESF,
-n'ont pas désiré y répondre ou
-n'ont pas réalisé l'importance de la requête.

L'ERIH est donc basé sur les noms de revues qui ont été proposées à l'ESF, la Fondation n'ayant aucune possibilité de découvrir des données qui ne lui ont pas été soumises.

Disciplines / Expert Panels under ERIH

- Anthropology
- Archaeology
- Art and Art history
- Classical Studies
- Gender Studies
- History
- History and Philosophy of Science
- Linguistics
- Literary Studies
- Music & Musicology
- Oriental & African Studies
- Pedagogical & Educational Research
- Philosophy
- Psychology
- Religious Studies & Theology

ERIH: the process

ESF/SCH Member Organisations

list journals according to disciplines

15 ERIH Expert Panels

analyse, complete, harmonise and finalise lists
grade journals in categories A,B,C

ERIH Steering Committee

supervises and unifies methodology (“bottom up”)
reports to SCH
reports through ESF/SCH to HERA/European Commission

ESF Standing Committee for the Humanities

approves final reference lists and ensures acceptance
develops financial continuity plan
makes recommendations on future development of ERIH

Consignes données aux panels

- 1) In order to be included a journal should fulfil normal international academic standards of quality, i.e.
 - it must have a quality control policy that governs the selection of articles.
 - peer review of submitted articles, (but it is also accepted that some high-quality journals have other systems of ensuring quality control).

2) All journals must also fulfil basic publishing standards:

- ISSN,
- timeliness of publication,
- active operations of an editorial board,
- complete bibliographic information for all cited references, and
- full address information for every author.

Journal categories in the 2007-version

Category A (target: 5-20% of all titles):

- High-ranking, international level publication;
- Very strong reputation among researchers of the field;
- Regularly cited all over the world

Category B:

- Standard, international level publication;
- Good reputation among researchers of the field in different countries

Category C:

- Only European publications to be considered (ESF Member Organisations);
- Important local / regional level publication;
- Mainly domestic readership, but occasionally cited outside the publishing country

Les **14 ERIH Initial Lists** préparées par les *expert panels* furent aussi contrôlées et validées par *le ESF Standing Committee for the Humanities* et par *HERA* (Humanities in the European Research Area) et publiées en 2007-2008.

Le terme 'initial lists' était supposé signaler qu'il s'agissait justement d'une première phase de cet exercice.

C'était prévu dès le début, qu'il y aurait une deuxième phase basée sur les réactions reçues à la publication des listes initiales, et nous nous attendions donc à recevoir des suggestions, commentaires et remarques critiques.

Problèmes liés à la catégorisation

A l'origine la différence entre les catégories A, B, et C, était conçue comme étant une différence de catégorie et pas de qualité. L'ERIH n'est pas une échelle ordinale. Un C ne devrait donc pas être perçu comme étant trois fois moins bon qu'un A, les catégories A et B étant réservées pour les revues internationales, le C pour les revues nationales.

Mais cela n'a pas été le cas. Les catégories ont été perçues comme étant hiérarchiques, ce qui nous a beaucoup inquiétés et qui nous a valu beaucoup de critiques, certaines exagérées et agressives, d'autres absolument justifiées

Une faille dans la logique de la version 2007?

Category A

- High-ranking, international level publication;

Category B:

- Standard, international level publication;

Category C:

- Only European publications to be considered (ESF Member Organisations);
- Important local / regional level publication;
- Mainly domestic readership, but occasionally cited outside the publishing country

ERIH – Criticism

- **115 negative reactions sent to the ESF** between November 2008 – June 2009:

- **Statistics:**

Anthropology	1	History and Philosophy of Science	62
Archaeology	2	Linguistics	4
Art and Art History	1	Literature	1
Classical Studies	3	Musicology	3
History	9	Religious Studies and Theology	4
		Miscellaneous	25

The ERIH-Steering Committee's reply to critical editors (2009)

To the Editors of Journals in History of Science, Technology and Medicine (1)

- *Humanities scholars are being presently evaluated in many different fora on the basis of inadequate instruments.*
- *The lists are to be continually updated by taking into account comments and feedback from editors, publishing houses, scientific associations, member organizations and other stakeholders.*
- *Staying away from these developments and placing an embargo on the ERIH will not lead to better working conditions and greater visibility of the Humanities in Europe.*
- *The ERIH is in the process of rethinking this categorization in A, B, and C, and of remodeling it according to a division into international, national or regional.*

To the Editors of Journals in History of Science, Technology and Medicine (2)

ERIH is led by scholars for scholars and the feedback mechanism enables scholars to communicate their views directly to the Expert Panels, and to the Steering Committee. The feedback that is being submitted by individual scholars and scientists as well as by expert communities will have an important role in the updating of the “Initial Lists”. In full recognition of the early stage of the process, and the shortcomings of the present version of the ERIH, we would like to ask the editors of journals in the field of History of Science, Technology and Medicine to reconsider their decision not to contribute to the process, and to let the project benefit from their critique and comments.

Gretty M. Mirdal (Chair, Standing Committee for the Humanities), University of Copenhagen

Professor Alain Peyraube (Chair, ERIH), CNRS and ERC Scientific Council,

Prof. Ferenc Kiefer, Hungarian Academy of Sciences,

Prof. Arto Mustajoki, University of Helsinki, Academy of Finland,

Prof. Michael Worton, Vice-Provost, University College London.

La deuxième phase

La deuxième (seconde?) phase de l'ERIH s'étend de 2008 à 2011. Le but de cette étape était de permettre aux chercheurs, éditeurs, et autres intéressés de pouvoir envoyer de nouvelles informations sur leurs pratiques de suscription, des pourcentages d'acceptation et rejet des articles reçus, etc. afin de **corriger des erreurs de classification**.

Les Expert Panels qui se réunirent entre Novembre 2008 et Janvier 2011, utilisèrent l'information ainsi obtenue, (**3540 fichiers**) ainsi que des données internationalement accessibles (journal websites, Ulrich's database, European Library Portal, etc.).

La nouvelle classification

NATional (NAT) European publications with a recognised scholarly significance among researchers in the respective research domains in a particular (mostly linguistically circumscribed) readership group in Europe; occasionally cited outside the publishing country, though their main target group is the domestic academic community.

INTernational (INT): both European and non-European publications with an internationally recognised scholarly significance among researchers in the respective research domains, and which are regularly cited worldwide.

International journals are themselves classified into two sub-categories based on a combination of two criteria: influence and scope:

INT1 Sub-Category: international publications with high visibility and influence among researchers in the various research domains in different countries, regularly cited all over the world.

INT2 Sub-Category: international publications with significant visibility and influence in the various research domains in different countries.

W Category Journals: journals which published their first issue three years or less before the closing date for feedbacks for a given panel”. Closing dates list is available [here](#).

Listes parues:

- Anthropology
- Gender Studies
- History
- Philosophy of science
- Linguistics
- Musicology
- Pedagogical and educational research
- Philosophy
- Psychology

Listes à paraître dans le courant de 2011:

- Archaeology
- Art and Art history
- Literature
- Classical Studies
- Religious Studies

Pour accéder aux nouvelles listes:

www.esf.org → Humanities → ERIH

ou bien

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

ERIH revised lists -
Search Criteria

ISSN

Journal Title

Discipline

Category 2007

Category 2011

Exemple d'une recherche

Search Results:

ISSN	Journal Title	Discipline	Category2007	Category 2011
0003-5033	L'Année Psychologique	Psychology	NAT	INT2

Total: 1 record(s)

Plus de 140 chercheurs de 28 pays différents ont participé à titre bénévole aux deux phases de l'ERIH. Ils ont préparé et planifié le travail, ils ont participé aux réunions des panels et aux discussions et débats électroniques. C'est grâce à leurs efforts que le projet ERIH a pu être accompli. Des remerciements particuliers sont adressés aux présidents des panels qui ont été responsables de piloter le travail et de maintenir la motivation des membres du panels à travers ce processus difficile et complexe.

Pour et contre
.....
Forces et faiblesses

1) Nouveaux standards d'excellence

Les revues évaluées par l'ERIH doivent répondre à un certain nombre de critères:

- évaluation par les pairs,
- comité de rédaction international,
- régularité des parutions,
- ouverture aux auteurs, et
- information bibliographique internationale.

Il est déjà possible d'observer que plusieurs revues ont commencé à adhérer à ces critères afin de pouvoir figurer dans l'ERIH

2) Plus grande visibilité

Plusieurs bases de données ont incorporé l'ERIH dans leurs titres, par exemple:

AMSTERDAM, June 10, 2009 – Elsevier announced today that its flagship product Scopus®, the world's largest abstract and citation database of peer-reviewed literature, has almost doubled its current Arts & Humanities (A&H) titles to 3,500, by adding top global journals using the European Science Foundation's **European Reference Index for Humanities (ERIH)**. This marks the first time that many of the ERIH titles will be available through a global citation database at no additional cost. (...)



3) Meilleure accessibilité

“ With the addition of the ERIH’s titles, Scopus can provide researchers with better access to national and international A&H journals. Subjects with the most journals include literature and literary theory (30% of new titles), general arts and humanities (22%), history (17%) and visual/performing arts (16%). More than half (60%) are published from the EMEA (Europe, Middle East and Africa) region, 38% from the Americas and 2% are from Asia-Pacific”

Par contre:

Le risque d'un mauvais emploi de l'ERIH

Malgré plusieurs mises en garde, il y a des preuves que l'ERIH est utilisé à des fins pour lesquelles il n'a pas été conçu:

« As they stand, the lists are not a bibliometric tool for the evaluation of individual researchers. The distinction between the categories A, B and C is to be understood as being not primarily qualitative and the categorisation is determined by issues such as scope and audience as explained in the ERIH Guidelines. Thus, such categorizations of journals do not prejudice the scientific quality of individual articles that appear in those journals ».

Représentation insuffisante de certaines disciplines

Certaines disciplines comme l'histoire, la linguistique, et la littérature étant très vastes, il a été difficile de couvrir les différentes sous-disciplines dans les panels.

D'autre part, des disciplines comme les sciences de la communication, films et médias n'ont pas eu leurs propres panel, ce qui a posé des problèmes d'expertise.

Couverture linguistique insuffisante

Le but de l'ERIH était de couvrir les revues scientifiques dans les sciences humaines dans la plupart des langues européennes. Ce but a naturellement été plus facile à atteindre dans les langues les plus répandues, tandis que dans certaines langues, la représentation des revues est faible dû au manque de communication entre l'ERIH et les « Member Organisations » de l'ESF.

Effets pervers de la bibliométrie?

On prétend que quand les paléontologues chinois commencèrent à trouver des os de dinosaures, il promirent une récompense pour chaque os qui leur serait remis. Le résultat fut que les fermiers qui trouvèrent de pareils os, les cassèrent en plusieurs morceaux afin d'obtenir plus de récompenses.

La politique de récompense par publication soutenue par plusieurs universités européennes ne favorisent pas nécessairement l'efficacité, tout comme les mécanismes de marketing dans la recherche ne favorisent pas la créativité.

Il serait pertinent de se demander comment la manie actuelle de mesurer l'impact des publications impactera elle-même l'avenir des sciences humaines.

Le but de l'ERIH était de prévenir certains de ces effets négatifs et surtout d'empêcher la disparition de bonnes revues de langues européennes autres que l'anglais.

L'importance du maintien de la biodiversité est unanimement acceptée. La diversité culturelle et linguistique devrait jouir de la même compréhension et du même soutien. Pourtant cela n'est pas le cas. La politique de publication qui incite les chercheurs à publier leur travaux principalement en anglais, risque de mener à une uniformité et à un déclin de la diversité pas seulement linguistique, mais aussi culturelle.

Si l'ERIH peut contribuer à arrêter ce mouvement destructeur, un des ses buts principaux aura été atteint.

Open access et évaluation des productions scientifiques dans l'espace européen de la recherche

Celina Ramjoué

Chargée du dossier "open access"

Commission européenne, Direction Générale de la Recherche et de l'innovation

Introduction

Bonjour à tous. Merci beaucoup d'avoir invité la Commission européenne à participer à ce colloque important. Je m'excuse de ne pas pouvoir être présente aujourd'hui physiquement et je remercie vivement Jean-Claude Guédon qui a accepté de vous présenter quelques réflexions de ma part sur « Open access et évaluation des productions scientifiques dans l'Espace Européen de la recherche ».

Comme vous l'aurez vu dans le programme, je travaille pour la Commission européenne, et plus spécifiquement pour la Direction Générale de la Recherche et de l'innovation. Je m'y occupe d'« open access » (ou libre accès) depuis 2006.

Les activités de la Commission européenne sur open access sont ancrées dans le contexte politique plus vaste de "l'Espace européen de la recherche" (EER) et "l'Union de l'Innovation".

L'"Espace européen de la recherche" réunit les activités, politiques et programmes européens de recherche et de développement menés dans une perspective transnationale.

L'Union de l'innovation est une des initiatives phare "d'Europe 2020" visant à stimuler et à accélérer l'innovation en Europe.

Je vous parlerai aujourd'hui de l'évaluation des productions scientifiques et d'Open Access en trois parties :

- Premièrement, je vous décrirai ce que fait la Commission européenne en matière d'évaluation des productions scientifiques en insistant aussi sur ce qu'elle ne fait pas ... ou pas encore.
- Deuxièmement, je vous parlerai de ce que nous faisons dans le domaine d'open access et comment cela nous mène vers le thème plus vaste de l'évaluation des productions scientifiques.
- Troisièmement, je dirai quelques mots sur les potentiels développements futurs en ce qui concerne une réflexion sur l'évaluation des productions scientifiques.

Mes réflexions portent sur la recherche en général, et pas seulement sur les sciences humaines et sociales qui sont le thème de ce colloque.

Ceci dit, comme je viens moi-même du monde de la recherche en sciences sociales, j'ai une compréhension et une sympathie particulières pour les SHS.

La Commission européenne et l'évaluation des productions scientifiques

a) La Commission s'occupe depuis des décennies d'évaluer - ou plus précisément: de faire évaluer par des experts externes - des contenus scientifiques sous forme de projets soumis aux programmes cadres pour la recherche et le développement.

Beaucoup d'entre vous auront déjà eu l'expérience soit de soumettre ou d'évaluer de tels projets de recherche européens. Ces projets à évaluer sont une sorte de "pré-production" scientifique, car

il ne s'agit justement que d'un *projet* de recherche. Mais très souvent cette pré-production est basée sur bien d'autres recherches déjà effectuées.

Ce domaine de l'évaluation de projets est très important pour la Commission, car il s'agit d'assurer la qualité de la recherche qu'elle finance grâce à un processus d'évaluation robuste, fiable et performant. Pour la Commission et pour d'autres organismes finançant la recherche, il y a lieu de réfléchir aux détails de l'évaluation de la recherche financée, notamment en ce qui concerne les critères d'évaluation.

Par exemple, quel doit être le poids du critère "impact potentiel" par rapport au critère "qualité scientifique"? Doit-on penser à inclure de nouveaux critères, par exemple le fait de tenir compte des pratiques de recherche « nouvelles » ou « éclairées », dont l'utilisation de l'open access ?

b) La Commission européenne se penche aussi sur l'évaluation des productions scientifiques issues des projets qu'elle finance. Elle cherche bien sûr à évaluer l'impact des programmes de recherche européens, par exemple en générant des statistiques sur le nombre de projets financés par domaine, le nombre de brevets liés à un projet européen, etc. Cette évaluation est déjà plus difficile, car elle se concentre surtout sur ce qui se passe *après la fin* d'un projet formel de recherche.

Nous travaillons continuellement et essayons d'améliorer cette partie de l'évaluation des productions scientifiques, ceci bien entendu dans l'objectif de pouvoir démontrer l'utilité de la recherche européenne et de rendre attrayant nos programmes de recherche.

c) Enfin, la Commission européenne est relativement peu avancée dans ses réflexions en matière d'évaluation des productions scientifiques de manière générale (non limité aux programmes cadres).

Elle ne fait que commencer de se poser les questions de ce colloque et de participer au débat plus large qui remet en question les pratiques communes de l'évaluation des productions scientifiques. Dans les deux prochaines parties, j'exposerai ce début de questionnement, qui, j'espère, se développera rapidement.

Open access et évaluation des productions scientifiques

Me voici donc arrivée au thème de l'*open access* (ou libre accès).

La Commission européenne a commencé à s'occuper de l'*open access* en 2006, à travers la publication d'une étude économique sur les marchés des publications scientifiques en Europe. Ont suivi en 2007 une Communication et des Conclusions du Conseil Compétitivité sur l'information scientifique à l'ère numérique: accès, dissémination et préservation. En 2008, nous avons lancé un pilote *open access* dans le 7^{ème} Programme Cadre. En 2009, la Commission a entamé un processus invitant les Etats Membres à devenir plus actifs concernant les informations scientifiques numériques dans le cadre plus vaste de l'Espace européen de la recherche. En 2010, Neelie Kroes devient Commissaire pour la stratégie numérique et proclame l'importance et le potentiel d'*open access*, non seulement aux publications, mais aussi aux données issues de la recherche. Toujours en 2010, les initiatives phare "Union de l'Innovation" et la "stratégie numérique pour l'Europe" viennent affirmer l'importance du thème d'*open access* dans une perspective de croissance et d'innovation. 2011 est l'année de préparation pour une nouvelle série de documents politiques de la Commission. J'y reviendrai.

En quoi ce parcours "open access" est-il pertinent pour l'évaluation des productions scientifiques ?

Il est pertinent dans la mesure où, au fil des années parcourues avec ce thème passionnant d'*open access*, nous nous sommes rendus compte que c'était un thème parmi beaucoup d'autres remettant en question les pratiques de recherche communes. Je veux dire par là qu'il n'est plus suffisant ni accepté de faire de la recherche comme elle était souvent faite dans le passé, c'est-à-dire plus ou moins en vase clos, sans échange avec le monde hors de la recherche.

Au contraire, il faut maintenant que la recherche, et je parle avant tout de recherche appliquée, soit responsable, qu'elle réponde à la politique et à la société. Cette recherche doit tenir compte des besoins et points de vue des utilisateurs potentiels et des groupes sociétaux concernés. Elle doit être transparente, ouverte et correctement diffusée, d'où le besoin d'*open access*.

C'est en passant par cet objectif de recherche et d'innovation responsable, contenant entre autres la vision 'ouverte' du libre accès, que l'on arrive entre autres au thème plus large de l'évaluation des productions scientifiques.

Bien sûr, la question de l'accès est liée de manière très intime à la bibliométrie :

- Par exemple, les revues *open access* peuvent souffrir du monopole du Facteur d'Impact (Impact Factor).
- Par exemple, les bases de données "classiques" de revues scientifiques ne reprennent pas toutes les revues en *open access*.

L'avancée de l'*open access* se heurte donc très rapidement aux outils et aux pratiques dominants.

En même temps, le progrès scientifique et le potentiel innovateur ne peuvent pas se développer de façon optimale sans *open access*.

Il faut donc repenser le système d'évaluation des productions scientifiques. D'où la très grande utilité de ce colloque et de la série d'études *JournalBase* qui y a mené.

Vers une réflexion sur l'évaluation des productions scientifiques

Open access et, de manière plus large, 'recherche et innovation responsable': voilà donc le contexte dans lequel la Commission européenne commence à s'intéresser de plus près aux questions d'évaluation des productions scientifiques. Par conséquent, en 2009, nous avons publié un appel à propositions visant à évaluer les systèmes d'évaluation des productions scientifiques au niveau du chercheur individuel¹.

Étaient invités des projets comparatifs se penchant de manière critique sur la façon courante d'évaluer le succès, la qualité, l'excellence et l'impact de la production scientifique. Les thèmes à aborder pouvaient inclure, par exemple, le système de relecture par les pairs (« peer review »), les indicateurs bibliométriques, la structure des carrières scientifiques, et la question des obstacles dans les carrières scientifiques liés spécifiquement au genre.

Deux projets ont été retenus pour financement et ont débuté cette année, chacun pour une durée de trois ans.

- **Le premier projet s'intitule « Academic Careers Understood through Measurement and Norms », acronyme : ACUMEN.**

ACUMEN constate :

- que l'on attend de plus en plus que la recherche ait un impact socio-économique,
- mais que l'on n'adapte pas le système d'évaluation des productions scientifiques à cette attente.
- Le projet se propose d'analyser cette situation, notamment
 - les problèmes que comportent les critères d'évaluation quantitatifs et le manque de critères qualitatifs,
 - les problèmes des systèmes de « peer review »
 - le fonctionnement (et le dysfonctionnement) des carrières scientifiques

Sur cette base, ACUMEN vise à préparer un **portefeuille de critères et lignes directrices d'évaluation** pour remédier à cette situation.

¹ SIS-2010-1.3.3-1 "Assessing how research outputs at individual researcher level are evaluated and measured" (Science in Society Work Programme 2010).

➤ **Le deuxième projet retenu est SISOB, « An Observatory for Science in Society based in Social Models ».**

L'objectif de SISOB est de développer un outil à travers lequel il est possible de mesurer l'appropriation et la diffusion des connaissances scientifiques.

Il s'agit donc d'un outil d'évaluation allant au-delà des outils classiques et se concentrant sur la question de l'impact sociétal de la recherche.

Nous avons donc deux projets très prometteurs et complémentaires.

ACUMEN développe une façon plus « sociétale » d'évaluer les chercheurs et SISOB développe un outil pour mesurer l'impact sociétal de la recherche.

Nous voyons ces projets comme un début de réflexion au niveau européen, à continuer, à alimenter et à construire dans les prochaines années.

Conclusion

En conclusion de cette brève intervention, je voudrais vous signaler que la Commission européenne adoptera, à la fin de cette année, une nouvelle Communication ainsi qu'une première Recommandation aux Etats-membres, portant sur les informations scientifiques numériques.

Plus précisément, elle se concentrera sur les thèmes d'accès, d'utilisation et de préservation des informations scientifiques, mais il est aussi probable qu'elle aborde, de manière probablement assez générale, la question des systèmes d'évaluation de recherche. L'objectif serait de signaler qu'il existe un lien bien réel entre ces thèmes, et qu'ils ne peuvent être traités de manière complète les uns sans les autres.

Nous avons eu, il y a quinze jours, une audition publique préparant ce processus politique menant à la Communication et la Recommandation prévues. Lors de cette audition, nous avons écouté plusieurs interventions passionnantes, puis un vif débat sur la question des systèmes d'évaluation de recherche. Le message est clair : l'Europe doit continuer à s'intéresser à ce domaine.

Introduction au libre accès dans la recherche

Le libre accès aux productions scientifiques : un nouveau modèle pour les sciences humaines et sociales (SHS)

[Christine Kosmopoulos](#), CNRS-UMR Géographie-cités

Les nouvelles technologies de l'information et de la communication (NTIC) ont radicalement modifié la nature de la communication scientifique, en favorisant le développement de nouveaux supports de diffusion d'échange. Dans cet esprit, l'émergence du modèle du libre accès en association avec le web 2.0 vient bousculer les pratiques en termes de publication, de diffusion, de veille et on peut le supposer en termes d'évaluation.

Qu'est ce le libre accès ?

Pour faire court, je reprendrai ici la définition donnée par Peter Suber¹, un des principaux initiateurs du libre accès : "le libre accès est numérique, en ligne, sans coût pour le lecteur et dispensé de la plupart des restrictions sur le copyright and les licences."

Autrement dit, lorsque les auteurs consentent à mettre leur document en libre accès, ils consentent par avance à la lecture, au téléchargement, à la copie, au partage, au stockage, à l'impression, à la recherche, à l'association du texte intégral de leur production et cela sans aucune restriction. La plupart des auteurs conservent cependant le copyright (s'il n'est pas rétrocédé à la revue) et bien entendu leur droit de stopper la commercialisation de leur travail. Toutes les sources utilisées issues du libre accès doivent être citées en bonne et due forme selon les pratiques courantes de toute publication scientifique.

Impulsé par différentes initiatives depuis les années 90, le mouvement international de l'*Open access* (libre accès) a été officialisé lors du *Budapest Open access Initiative* en décembre 2001. Cet appel regroupait des représentants de différentes institutions de recherche et universitaires mondiales². Jean-Claude Guéron et Francis André, ici présents, y ont participé. Depuis, chaque année, une rencontre internationale a lieu qui réunit les différents acteurs de la recherche et de la diffusion scientifique en libre accès afin de débattre des avancées du modèle, de travailler à l'interopérabilité des données, ainsi qu'aux services associés (citations, impact, data mining etc.). La dernière rencontre, riche de plus 300 participants, et d'une vingtaine de pays représentés, a eu lieu à Pékin en octobre 2010.

Les différents modèles du libre accès scientifique

Afin d'éviter toute confusion, il convient de rappeler que le libre accès scientifique peut prendre différentes formes que je classe en 3 catégories :

- L'édition en libre accès comprend toutes les revues scientifiques en libre accès. Ces revues répondent à tous les critères scientifiques et d'impartialité des revues papier à comité de lecture, mais elles ont la particularité d'être strictement numériques. Le DOAJ³ hébergé par la bibliothèque nationale de Suède recense plus de 6000 revues en libre accès, toutes disciplines confondues. Mais on trouve aussi de plus en plus d'ouvrages scientifiques en libre accès. Je reviendrai dessus à la fin de cette présentation.
- La seconde catégorie concerne les archives ouvertes. Les premières plates formes de dépôt en archives ouvertes ont été développées dans les années 90 sous l'impulsion de la communauté

¹ Suber, P., 2007, Open Access Overview: Focusing on open access to peer-reviewed research articles and their preprints

² <http://www.soros.org/openaccess/read>

³ <http://www.doaj.org/>

des physiciens de Los Alamos, en particulier Paul Ginsparg⁴. Il s'agit de permettre l'accès aux textes intégraux des articles, publiés ou non. Dans ce cas, ce sont les auteurs eux-mêmes qui déposent sur le serveur de leur institution leur pré ou post-print (si nécessaire y a des accords passés avec les éditeurs) ou leurs documents de travail. Les archives ouvertes peuvent également accueillir des thèses, des mémoires, des cours, des bases de données, des fonds de photos, vidéos etc. Au CNRS, c'est le *Centre de la Communication Scientifique Directe* (CCSD) qui a mis en place plusieurs plate-formes d'archives ouvertes⁵. Tous les sites d'archives sont recensés dans le *Directory of Open access Repositories* (OpenDOAR)⁶, produit par l'université de Nottingham. A ce jour, le site recense plus de 900 sites d'archives ouvertes en Europe, 464 en Amérique du nord et 1830 en tout. Ce chiffre est en constante évolution.

- La troisième catégorie concerne un nouveau modèle collaboratif basé sur le web 2.0. Il s'agit des wikis et des blogs qui se diffusent très rapidement dans la communauté scientifique. Il s'agit là de lieux d'échanges dont la forme interpelle le système traditionnel de communication et de publication scientifiques.

Comment fonctionne le libre accès ?

Quel contrôle de la qualité scientifique est exercé dans le libre accès ?

Les critères pour publier dans une revue en libre accès sont exactement les mêmes que ceux utilisés dans le modèle traditionnel payant. Autrement dit, tous les sujets universitaires et de recherche peuvent y être traités et chaque revue se doit d'appliquer le système d'évaluation par les pairs.

En ce qui concerne les archives ouvertes, il y a différentes pratiques. Chaque auteur, reconnu comme membre d'une institution scientifique ou universitaire peut décider de déposer, c'est-à-dire "auto-archiver" son document. Si c'est un article, il peut déposer un texte soumis à une revue, et là il s'agit d'une pré-print, mais cela peut être aussi un texte accepté pour publication, et là il s'agit d'un post-print. Le 7^{ème} programme cadre de l'Union Européenne préconise très fortement le dépôt en archives ouvertes. Dans ce cas, il s'agit pour l'auteur de déposer son article évalué et accepté dans un délai en accord avec l'éditeur⁷. Comme je l'ai dit précédemment, beaucoup d'autres documents comme les bases de données, vidéos, rapports, photos, thèses etc. peuvent aussi être déposés sur les sites d'archives ouvertes. Une fois le dépôt validé par l'équipe responsable de l'archive ouverte, le document est proposé en libre accès.

Pour ce qui est des wikis et des blogs, la qualité est auto-contrôlée par la communauté scientifique elle-même.

Comment l'information est-elle diffusée ?

Afin de permettre une diffusion large des documents en libre accès, le modèle du libre accès travaille à développer des systèmes d'interopérabilité de plus en plus performants entre les serveurs et les moteurs de recherche⁸. Ce qui signifie que lorsque l'on dépose un document en archives ouvertes, un article dans une revue en accès en libre, ces documents ou tous les textes qui touchent à la même thématique peuvent être rediffusés sur d'autres sites de manière regroupée. Des milliers de sites fonctionnent actuellement selon certaines normes d'interopérabilité.

Un exemple intéressant de ce type de développement est celui de *Driver: the Networking European*

⁴ <http://arxiv.org/>

⁵ <http://www.ccsd.cnrs.fr/>

⁶ <http://www.opendoar.org/>

⁷ Cette initiative représente à peu près 20% du budget du 7^{ème} PCRD et s'applique à 7 domaines scientifiques différents dont les sciences humaines et sociales.

⁸ L'université de Bielefeld en Allemagne moissonne dans plus 1800 entrepôts et recense désormais plus de 25 millions de documents.

Scientific Repository, un super portail européen lancé en 2008 sous l'impulsion du 6^{ème} PCRD, qui combine 249 répertoires d'archives ouvertes de 33 pays.

Je ne m'étendrai pas ici sur les services associés au libre accès qui sont développés en association avec le web 2.0. Ce qui est à retenir, c'est que d'ores et déjà un certain nombre d'outils sont proposés tel que les plates-formes collaboratives, la réception d'une information instantanée et ciblée avec les flux RSS, des outils d'indexation qui mettent en relation les publications sur les questions similaires ("related articles" dans Google Scholar), le data mining etc. qui permettent d'extraire et de ré-utiliser l'information permettant des analyses statistiques et autres.

Quels sont les avantages du libre accès en science ?

Une diffusion élargie et en libre accès des documents via des systèmes d'interopérabilité accroît la visibilité des productions scientifiques. Elle permet également un accès inédit à des données jusqu'ici peu valorisées, comme les traductions, les productions audiovisuelles, les cartes, les bases de données, les catalogues, les rapports etc. via les archives ouvertes

Par l'intermédiaire du moissonnage des métadonnées, on peut collecter et croiser des données de sources différentes et ainsi se créer ses propres répertoires. Il va sans dire que le coût est bien inférieur pour la communauté scientifique en comparaison des prix d'abonnement des éditeurs scientifiques commerciaux, si toutefois on ne répercute pas le poids financier sur l'auteur comme c'est le cas dans le modèle libre accès/auteur-payeur qui essaie de s'imposer actuellement notamment dans les sciences dures.

Ce qui est également intéressant, c'est que globalement le libre accès implique un accroissement des taux de citations comme cela a été montré récemment dans différentes études⁹.

Qu'en est-il des sciences humaines et sociales et quel est le défi à relever pour l'avenir ?

Pour ce qui est des archives ouvertes, le rapport en 2008 de la commission européenne sur le sujet souligne un très net retard dans le dépôt sur les sites institutionnels qui serait lié à un manque d'information et de connaissance sur ce nouveau moyen de diffusion. Les sciences sociales et plus particulièrement les sciences humaines doivent accélérer leur dépôt en libre accès et adopter de nouvelles pratiques. Concrètement, il s'agirait que les politiques publiques de la recherche prévoient des formations de sensibilisation au sein de la communauté scientifique des avantages du libre accès pour la science. C'est ce qui est en train de se passer au CNRS, puisque la direction de l'INSHS a demandé à chaque unité de nommer un correspondant IST qui a, entre autres, pour mission de mobiliser les chercheurs au dépôt sur HALSHS du CCSD.

Concernant les revues en libre accès, 25% des revues actuellement recensées dans le DOAJ sont des revues de sciences sociales et 21% des revues de sciences humaines, ce n'est pas un mauvais score mais elles ne représentent que 5% (571 revues) des revues recensées dans *JournalBase* à partir des sources à visée bibliométrique (AERES, ERIH, WOS, SCOPUS)¹⁰.

On peut souligner des initiatives européennes très intéressantes comme OAPEN (*Open access Publishing in European Networks*) qui réunit 11 presses universitaires européennes et dont le but est de créer la plus importante collection de livres scientifiques en libre accès dans les différentes langues européennes issus de différents domaines des sciences humaines et sociales. Ce projet commencé en septembre 2009 propose déjà plusieurs centaines de livres en accès libre.

Pour finir, je dirais qu'en raison de leur spécificité, les sciences humaines et sociales (SHS) ont un rôle de pionnier à jouer dans le développement de ces modèles. Intégrer la recherche en SHS au réseau du libre accès, c'est permettre à de nouvelles formes d'évaluation scientifique d'émerger,

⁹ Gargouri Y., Hajjem C., Larivière V., Gingras Y., Carr L., Brody T., et al., 2010, « Self-Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research ». *PLoS ONE*, vol.5, num 10, e13636.

¹⁰ <http://journalbase.cnrs.fr/>

contrôlées par la communauté scientifiques, des formes plus harmonisées, et dans lesquelles les instruments bibliométriques ne sont qu'un aspect.

JournalBase, Une étude comparative sur les bases de données scientifiques à vocation bibliométrique pour les SHS

Michèle Dassa¹ et Christine Kosmopoulos²

¹CNRS-Institut des Sciences Humaines et Sociales

²CNRS-UMR Géographie-cités

Présentation du projet

Le contexte

En 1964, Eugène Garfield, fondateur de l'*Institute for Scientific Information (ISI)*, créait la célèbre base de données *Science Citation Index (SCI)*. A l'origine cette base de données d'indexation des citations n'était pas destinée à évaluer la production scientifique. Il s'agissait d'un outil documentaire hautement spécialisé pour les sciences dures et très utilisé en médecine, outil qui permettait par l'intermédiaire des citations produites dans les articles d'associer les publications traitant du même sujet tant en recherche fondamentale, qu'en recherche appliquée. Ce nouveau système d'indexation se proposait de remplacer, dans ces disciplines, le système d'indexation classique. En effet, la méthode d'indexation par thésaurus utilisé par les bibliothécaires offrait des entrées beaucoup trop généralistes pour des sujets très spécialisés en particulier dans le domaine de la recherche scientifique.

Cet instrument révolutionnaire a connu un très grand succès, la plupart des laboratoires en sciences dures étaient abonnés aux fameux fascicules d'indexation. Ce n'est que plus tard que cet outil au départ strictement destiné à l'analyse et l'information documentaire des articles scientifiques a évolué ces dernières années vers un outil d'analyse quantitative des citations.

Cette évolution s'est accélérée avec le développement des nouvelles technologies. En 1992, l'entreprise ISI a été rachetée par le groupe Thomson qui a développé le célèbre *Web of Science (WOS)* en ligne intégrant les trois grands index de citations créées par E. Garfield : *SCI*, *Art and Humanities Citation Index (AHCI)*, *Social Science Citation Index (SSCI)*, index auxquels sont désormais associés toute une série d'indicateurs bibliométriques¹.

A côté de l'évaluation par les pairs, ces outils quantitatifs destinés à mesurer la qualité du travail des chercheurs, en permettant la comparaison au moyen d'indicateurs chiffrés, sont utilisés dans un certain nombre de sciences dures et tentent de s'imposer en *Sciences Humaines et Sociales (SHS)*. Pour le SHS, ce débat n'est pas nouveau, puisque dès 1989, le *Comité National d'évaluation de la Recherche (CNER)* se prononçait en faveur de la prise en compte d'indicateurs bibliométriques pour l'évaluation des travaux des chercheurs en SHS. De 2007 à 2009, la *Direction des Partenariats (DPA)* au CNRS s'est efforcée d'appliquer ces méthodes en demandant aux laboratoires, à côté de leur rapport quadriennal, de remplir des fichiers excel comprenant des indicateurs bibliométriques. Ces indicateurs devaient provenir du *Web of Science (WOS)* ou de *Publish or Perish (PoP)*, outil bibliométrique en libre accès qui utilise les données de *Google Scholar (GS)*.

La tendance est donc forte de voir s'appliquer, à côté de la revue des pairs, ce modèle en SHS, mais les outils proposés (index, listes d'autorité etc.) sont-ils adaptés ? A ce stade, il nous a semblé indispensable de faire un recensement exhaustif du contenu des outils actuellement disponibles pour ce type d'analyse, afin de connaître de manière objective leur représentativité et le crédit que l'on peut accorder à leurs résultats statistiques.

¹ Par bibliométrie, on entend en général l'application des mathématiques et des méthodes statistiques aux livres, articles et autres moyens de communication

L'historique

Nous avons saisi l'occasion de l'appel à projets lancé en 2007 par le *Très Grand Équipement Adonis* (TGE) pour les SHS du CNRS², sur les *Outils innovants de traitement numérique pour la valorisation et la diffusion des données* pour soumettre notre projet.

Le projet a été accepté et a démarré en janvier 2008. L'idée était de faire une comparaison systématique des bases de données et listes d'autorité bibliographiques nationales et internationales pour les SHS (*Web of Science*, *Scopus*, *ERIH*, *AERES*) afin d'évaluer la représentativité de ces sources pour les revues selon les disciplines mais aussi selon les pays. Il s'agissait de créer une base de données avec toutes ces informations et de développer une plateforme de consultation et d'interrogation.

Un premier état de l'avancée des travaux a été présenté en septembre 2008 à l'université d'été du TGE, suivis de plusieurs autres présentations notamment en avril 2009 à l'INS2I à Grenoble³. A ces différentes occasions, nous avons fait le point sur les nombreux problèmes rencontrés dans la comparaison de ces sources et la nécessité devant laquelle nous nous trouvions de devoir concevoir et développer une méthodologie adaptée. Un an et demi après le lancement de notre projet, nous avons publié en juin 2009 les premiers résultats dans un dossier de *Cybergeogeo* qui expose notre méthode, le choix de nos rubriques disciplinaires et dresse un tableau comparatif des différentes sources bibliométriques citées. En fusionnant toutes les données sur les revues indexées, nous sommes parvenues à une liste exhaustive d'un peu plus de 9500 titres de revues. Le rapport final de 720 p., au format PDF recensait alors 9500 revues dans 27 champs disciplinaires, accompagnées de la source dans laquelle elles se trouvaient recensées⁴. En janvier 2010, nous avons publié une analyse détaillée des résultats pour l'année 2009⁵. Ces résultats sont très proches de ceux de l'année 2010 que nous allons vous présenter aujourd'hui.

Les informations contenues dans les bases de données de *JournalBase* pour 2009 et 2010 sont en ligne et sont désormais interrogeables. Nous profitons de l'occasion de ce colloque, pour lancer officiellement le site bilingue et en libre accès de la base de données *JournalBase* qui a servi à notre étude comparative⁶.

JournalBase : un recensement exhaustif

La question de l'application du modèle bibliométrique d'évaluation de la production scientifique se pose de façon récurrente dans l'ensemble de la communauté SHS, nationale ou internationale et cela dans une certaine confusion et angoisse tout à fait légitimes. En effet, une simple consultation du WOS montre que la source est très loin d'être exhaustive pour les SHS, bien plus, l'intégration de revues non anglo-saxonnes est extrêmement difficile et cela bien qu'elles répondent à tous les critères exigés par *Thomson Reuters*. Par conséquent utiliser le WOS pour évaluer la qualité du travail d'un laboratoire ou d'un chercheur en SHS est d'évidence tout à fait inadapté. Mais existe-t-il d'autres outils qui viendraient combler les lacunes du WOS ?

En effet, d'autres bases de données que le WOS ont vu le jour, notamment *Scopus* d'Elsevier, voilà déjà quelques années. En 2007, la *European Science Foundation* (ESF) en collaboration avec le consortium *Humanities in the European Research Area* (HERA)⁷ s'est également lancée dans le recensement des revues européennes en SHS et a publié une première liste de revues pour les

² <http://www.tge-adonis.fr/>

³ <http://mi2s.imag.fr/pm/indicateurs-bibliometriques-production-scientifique-et-evaluation-du-chercheur>

⁴ <http://cybergeogeo.revues.org/22492>

⁵ <http://cybergeogeo.revues.org/22864>

⁶ <http://journalbase.cnrs.fr/>

⁷ ESF : *European Science Foundation* est une association non gouvernementale qui comprend 80 organisations membres dédiée à la recherche scientifique dans 30 pays européens. Elle a été fondée en 1974. Pour en savoir plus : <http://www.esf.org/>

HERA : <http://www.heranet.info/>

sciences humaines ERIH comprenant environ 5000 revues. Une autre liste pour les sciences sociales est en cours de réalisation.

Du côté national, en juillet 2008 c'est au tour de l'Agence d'Evaluation de la Recherche et de l'Enseignement Supérieur(AERES) de publier des listes de revues SHS.

Et puis il y a bien entendu *Publish or Perish* (PoP) et *Google Scholar* (GS) qui sont souvent utilisés, alors que leur fiabilité n'a jamais été démontrée. Il s'agit de systèmes instables sur lesquels nous n'avons pas d'informations rigoureuses, de listes ou de bases de données sur lesquelles nous aurions pu nous appuyer pour les analyser. Nous n'avons par conséquent pas intégré GS et PoP dans notre étude comparative *JournalBase*⁸.

Aucune étude sérieuse des sources utilisées n'ayant été faite jusqu'à là. En procédant à un inventaire exhaustif et comparatif, *JournalBase* constitue la première source de données objectives sur les bases et listes de données à vocation bibliométriques. Nous pouvons désormais répondre par des données concrètes aux questions que se pose la communauté SHS et contribuer à l'avancement du débat sur un terrain vérifiable, ce qui n'a jamais été le cas auparavant. Mais au-delà de cette étude et des résultats qui vont vous être présentés maintenant, l'objectif de notre colloque est également d'élargir la réflexion en faveur de propositions innovantes, associées précisément au développement des nouvelles technologies de l'information et de la communication.

Un outil inédit en libre accès au service des SHS

JournalBase 2010, offre un service inédit en SHS, puisque non seulement il recense toutes les revues du WOS, de *Scopus*, des listes ERIH, des listes de l'AERES, mais de plus nous l'avons enrichi des informations suivantes qui ne figuraient pas dans les sources⁹:

- ISSN
- e-ISSN
- la langue de publication
- le pays de l'éditeur
- l'éditeur
- le site web de la revue ou de l'éditeur
- la nature de leur accès (libre ou payant)
- la mention de son appartenance au DOAJ¹⁰
- le soutien par le CNRS

L'objectif est de donner accès à une liste la plus exhaustive possible des revues recensées dans une ou plusieurs disciplines et à des informations parfois méconnues. L'accès instantané au site web de la revue participe de cette volonté de valorisation. Il va de soi que toutes les revues issues de ces quatre sources sont à comité de lecture. Si ce n'est pas le cas, la responsabilité en revient au producteur de ces listes qui n'auraient pas examiné de près les critères d'intégration des revues à leur liste.

⁸ Lancé fin 2004, GS n'est en effet pas une base de données bibliographique à proprement parler; il s'agit d'un moteur de recherche spécialisé dans la littérature scientifique. La base qui en est issue est la propriété du groupe qui a lancé le moteur de recherches Google en 1998. Les documents indexés proviennent des éditeurs scientifiques, des sociétés savantes, des dépôts d'archives ouvertes, des universités et d'autres organismes de recherche. *Google Scholar* utilise comme Google un outil robot qui vient explorer, classer et indexer le contenu d'un site qui lui a été signalé, mais il n'y a à ce jour aucune information précise sur les sources de GS. Des études ont été faites qui montrent que même lorsque certaines sources sont identifiées (HAL, revues.org par exemple), les références ne sortent pas de façon systématique. C'est la raison pour laquelle nous avons renoncé à intégrer GS à notre étude comparative. GS et PoP nécessitent une méthodologie d'étude spécifique.

⁹ Nous n'avons pas retenu les éventuels classements (A, B, C, etc) AERES et ERIH

¹⁰ DOAJ...

Sans se substituer à *Thomson Reuters*, *Elsevier*, l'*AERES* ou l'*ESF* qui seuls éditent les listes intégrales qu'ils produisent, il est possible sur *JournalBase* d'interroger par discipline, par pays d'éditeur, par langue, par ISSN etc. Un large choix de requêtes est proposé par la recherche avancée. On peut faire des requêtes croisées sur une langue, une discipline, une année etc. comme le montre la figure 1.

Figure 1 : Recherche avancée

Les résultats affichent les sources dans lesquelles est recensée la revue consultée (figure 2).

Les informations indiquées ci-dessous ont fait l'objet de nombreuses vérifications croisées. Si, malgré notre vigilance, vous constatez une erreur, nous vous remercions de nous la signaler. A noter qu'il faut être inscrit pour signaler une erreur.

Revue : Acta Ethnographica Hungarica

Titre :	Acta Ethnographica Hungarica
Editeur :	Akademiai Kiado Rt.
Pays de l'éditeur :	HONGRIE
Langue(s) de publication :	allemand, anglais
ISSN :	1216-9803
ISSN électronique :	1588-2586
Site web :	http://www.akademiai.com/publications
Référencement 2010 :	ERIH, Scopus

Référencement : ▲ 2010 ● 2009

	AERES	ERIH	WOS-AHCI	WOS-SSCI	Scopus
Anthropologie, Ethnologie et Folklore	●	▲ ●			
Archéologie					

Figure 2 : Signaler des informations

Nous avons opté pour le libre accès afin que l'ensemble de la communauté scientifique, - laboratoires, chercheurs, organismes scientifiques européens et internationaux ¹¹, mais aussi plus largement les producteurs de l'information scientifique (IST), éditeurs commerciaux, sociétés savantes, éditeurs de revues, experts en IST, sans oublier les instances d'évaluation nationales (Comité National du CNRS (CN), Comité National des Universités (CNU), l'AERES, Agence Nationale de la Recherche (ANR) etc.- puissent accéder et utiliser ces données sans restriction.

¹¹ ESF, Communauté européenne, l'Observatoire des Sciences et des Technologies (OST) du Canada etc.

Les échanges avec tous ces acteurs et l'interactivité sont très attendus, puisque que l'on peut intervenir sur le site bilingue et faire des propositions de modifications ou d'insertions de revues qui par erreur auraient été oubliées, comme cela est montré sur la figure 2

L'élaboration de JournalBase

Comment a-t-on pu réunir dans une même base ces différentes sources ? Rappel de la méthode pionnière en 27 disciplines

Avant d'entrer de manière plus fine dans les résultats de cette étude, et pour qu'il n'y ait pas de confusion sur le découpage des champs disciplinaires que nous allons vous présenter, nous nous proposons de vous rappeler ici très rapidement quelle a été notre méthode et pourquoi nous nous retrouvons avec 27 rubriques disciplinaires.

Le premier constat qui ressortissait de la collecte de l'information de chacune des sources, c'était leur forte hétérogénéité et l'extrême difficulté à traiter de manière comparative leurs données. Tout d'abord, nous avons été confrontées à la très forte diversité des sources, tant d'un point de vue qualitatif, dans les niveaux et la qualité de l'indexation des bases¹², la précision et l'exactitude des données, le type et le nombre de champs informés, que d'un point de vue quantitatif notamment dans les divers formats des fichiers (PDF, html, excel etc.). Nous avons aussi observé un grand nombre d'erreurs dans certaines bases, des doublons, des triplons, des erreurs sur les titres, sur les ISSN, des ISSN manquants, des confusions entre les ISSN papier et électronique, des erreurs dans les indexations etc. Après un très long travail de vérification et de nettoyage de ces sources réalisé par une jeune stagiaire de l'INTD¹³, nous avons développé une méthodologie.

Pour comparer les quatre sources, nous avons décidé de passer par une nomenclature disciplinaire commune qu'il nous a fallu établir. Pour cela, nous avons été chercher dans l'existant. Une vingtaine de listes disciplinaires s'est d'emblée présentée à nous : tout d'abord celles des fournisseurs de données que nous étudions : WOS, Scopus, ERIH et AERES mais aussi celles de l'ESF (différente d'ERIH), du CNU, CN du CNRS, HAL, INIST, FRANCIS etc.

Finalement, notre choix s'est porté sur la classification européenne en deux listes de l'ESF. Celle-ci avait l'avantage de dépasser le niveau national tout en étant assez proche des spécialisations françaises. Nous sommes donc parties en 2008 des deux listes initiales de l'ESF, que nous avons traduites, puis modifiées pour finalement les réunir en une seule liste originale de 27 catégories (Figure 3).

¹² A l'AERES, par exemple, il y avait des listes qui ne comprenaient que le nom de la revue, sans aucune autre donnée. Dans le cas de l'AERES, les listes sont classées par discipline, mais d'autres sources comme *Scopus* n'étaient pas classées par discipline seulement indexées par des mots clés (à partir de certains articles qu'elles contenaient) et dont les niveaux d'indexation pouvaient atteindre 20 ! Suite aux différents échanges que nous avons eus avec *Scopus* qui soulignaient l'inanité de leur indexation (bien que nous en ayons tenu compte dans notre recensement de 2009), nos interlocuteurs ont décidé de supprimer près de 99% des mots clés de leur fichier excel de recensement des revues dans Scopus. Ceci a pour résultat qu'il est à présent tout aussi difficile de dispatcher des revues de Scopus dans les rubriques de *JournalBase*.

¹³ Sonia Christon



Les 27 rubriques adoptées

Anthropologie, Ethnologie et Folklore	Littérature
Archéologie	Musique et Musicologie
Art et Histoire de l'Art	Pédagogie et Sciences de l'Education
Démographie	Philosophie
Droit	Psychologie et sciences cognitives
Economie	Relations Internationales
Etudes classiques	Religion et Théologie
Etudes des Femmes	Science Politique
Etudes du Genre	Sciences commerciales et administratives
Etudes Orientales et Africaines	Sciences de la communication
Géographie	Sciences de l'environnement
Histoire	Sociologie
Histoire et Philosophie des Sciences	Statistiques et Informatique
Linguistique	

Figure 3 : les 27 rubriques adoptées

Comme nous l'avons publié en janvier 2010, cette nomenclature n'est cependant pas exempte de faiblesses. Les catégories qu'elle utilise ne se situent pas au même niveau d'objet scientifique ou d'institution (par exemple la catégorie « relations internationales » pourrait être intégrée dans la rubrique « science politique »). Il n'est pas évident que la distinction entre la catégorie « étude des femmes » et « étude du genre » soit absolument pertinente et nécessaire, etc. (Dassa, Kosmopoulos, Pumain, 2010)¹⁴.

Au nombreux problèmes rencontrés, est venu s'ajouter la difficulté de suivre l'évolution des listes, aucune information n'étant donnée en ligne sur la mise à jour des pages ou des listes et cela malgré les différents messages que nous avons adressés. Cela a été particulièrement compliqué en 2008 et 2009 pour les listes de l'AERES (depuis un an les dates de mises à jour sont désormais signalées).

Analyse des résultats pour 2010

Le périmètre de l'étude

Notre étude porte :

- sur les 2 index des sciences humaines et sociales de la plateforme WOS (Web of Science) : l'AHCI (*Art and Humanities Citation Index*) et le SSCI (*Social Sciences Citation index*) qui comportent respectivement 1885 et 2642 revues. Les revues du WOS-SHS, une fois dédoublonnées, représentent 3936 titres sur un total d'environ 11 000 revues.
- sur *Scopus*, base développée en 2004 par le groupe d'édition Elsevier pour concurrencer Thomson Scientific. Le recensement de 2010 fait apparaître un nombre total 4503 revues en SHS pour un total de plus 15000 revues actives.
- sur les listes de revues ERIH : produite en 2007 par l'ESF et le consortium HERA : 5131 revues ont été recensées pour les sciences humaines.

¹⁴ Sur la nomenclature, voir notre article : <http://cybergeog.revues.org/22864>

- sur les listes de l'AERES : Le nombre de revues recensées dernièrement dans ces listes est au total de 6245 sans le droit et les sciences de la communication.

Au total, JournalBase2010 comprend 10 756 revues différentes contre 9551 recensées en 2009¹⁵.

Un noyau dur de 1033 revues sur 10756 revues que compte JournalBase2010 soit seulement 9.6%

On constate qu'il existe un spectre très large des revues en SHS, certaines revues étant d'ailleurs présentes que dans une seule base ou liste. Cela a pour conséquence que selon la source utilisée, on obtiendra des résultats bibliométriques, comme par exemple le nombre de publications, forcément différents. De ce point de vue, aucune des bases ou listes ne peut donc apporter des résultats conformes à la réalité.

Sur la figure 4, on peut voir de manière plus détaillée le % des sources dans Journalbase 2010 : 37.3 % des revues sont présentes dans le WOS, 42.6 % dans Scopus, 48.6 % dans ERIH (qui ne comprend pour l'instant que les sciences humaines) et 59.1 % dans les listes de l'AERES (sans le droit).

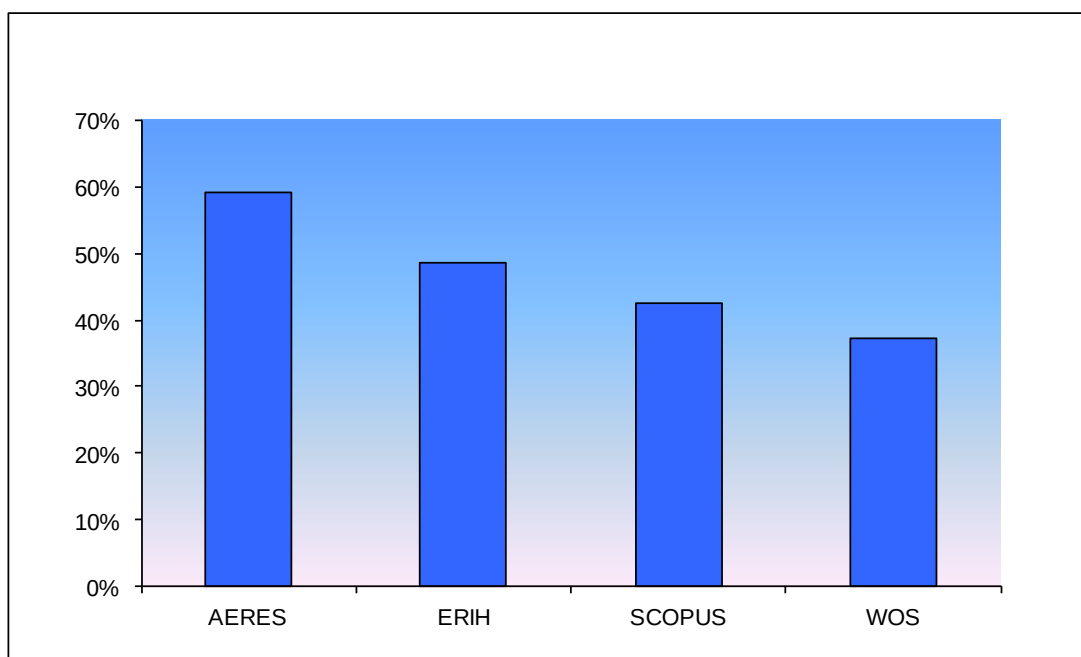


Figure 4 : % de revues sur le nombre total de revues par source étudiée

Répartition des revues par pays d'édition

La figure 5, montre la répartition des revues par groupe de pays d'édition.

¹⁵ voir notre article : <http://cybergegeo.revues.org/22864>

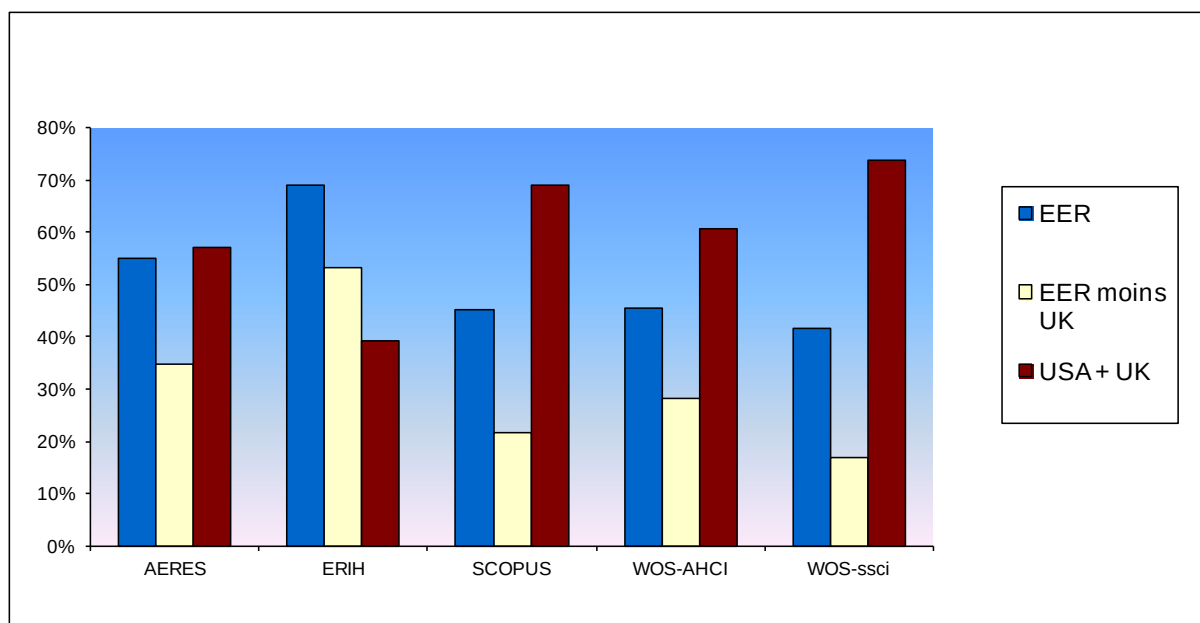


Figure 5 : Répartition des revues par groupe de pays d'édition

Nous avons différencié l'espace européen de la recherche¹⁶ (EER), l'espace européen de la recherche sans le Royaume-Uni (EER moins UK) et le groupe Etats-Unis et Royaume-Uni (USA + UK).

On observe (figure 5) comme lors de l'étude 2009, une prédominance de la littérature anglo-saxonne dans les bases commerciales (WOS et SCOPUS), ce qui n'est pas le cas d'ERIH et de l'AERES. On peut cependant remarquer que les listes AERES comportent une proportion plus importantes de revues anglo-saxonnes que les listes ERIH. ERIH apparaît donc comme la source la plus complète et finalement le plus homogène du point de vu de la répartition géographique.

Si on considère l'ensemble « Etats-Unis et Royaume-Uni » (USA + UK, figure 5) on vérifie que les bases commerciales *Wos* et *Scopus* répertorient une très grande majorité de revues anglophones, dans tous les cas plus de 2/3 des revues et jusqu'à près de 75 % pour la base WOS-SSCI.

ERIH qui se positionne comme un outil européen ne présente effectivement que 39,3 % de revues éditées au Royaume-Uni et aux Etats-Unis. Inversement, il répertorie plus de 50 % (53,2 %) de revues européennes hors du Royaume-Uni, alors que l'on n'en trouve qu'environ 17 % dans le WOS-SSCI. Les listes de AERES comportent près de 60 % (57%) de revues anglo-saxonnes et environ 35 % de revues européennes hors Royaume-Uni.

Il est intéressant de remarquer que si l'on compare les listes ERIH et le WOS-AHCI qui regroupent uniquement les Humanités, on constate que le profil de répartition n'est pas du tout le même.

La présence française

La répartition du nombre de revues françaises selon les sources est présentée dans la figure 6. On constate que la présence française est relativement faible quel que soit les sources étudiées. Ce sont les listes de l'AERES qui présentent le plus grand nombre de revues françaises avec 427 revues soit seulement environ 4 % de JournalBase et 6,8 % de l'ensemble des revues de l'AERES.

¹⁶ L'espace Européen de la Recherche est composé des 27 pays de l'Union européenne plus la Suisse, la Norvège, le Lichtenstein, Israël et l'Islande.

Sur les 2010 revues recensées en SHS, environ 600 à 1200 revues françaises en SHS sont considérées comme actives¹⁷.

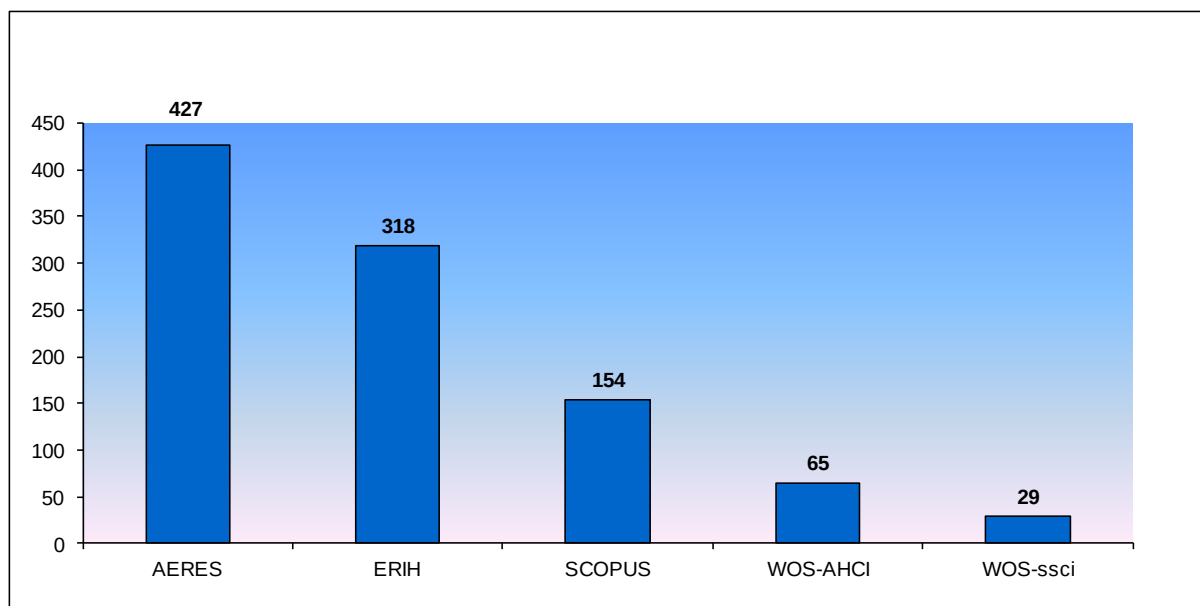


Figure 6 : nombre de revues françaises par source étudiée

La part des revues éditées en France incluse dans JournalBase 2010 est d'un peu moins de 3,5 % dans la base AHCI avec 65 revues, elle descend à 1,1 % des revues du SSCI et 3,4 % de celles de *Scopus* avec respectivement 29 et 154 revues recensées. La part française dans *Scopus* reste faible même en 2010, puisque sur 1100 nouvelles revues intégrées, moins de 10% sont françaises. On atteint un total pour 2010 d'un peu plus de 150 revues françaises dans *Scopus*.

La France sous représentée

Les revues françaises sont très clairement sous-représentées dans les bases de données et les listes d'autorité. On a voulu dans la figure 7 reprendre une comparaison souvent faite entre les USA, le Royaume-Uni, l'Allemagne et la France, comparaison toujours défavorable à la France. On observe là encore une prédominance de la littérature anglo-saxonne (USA et UK) dans les bases commerciales.

¹⁷ Rapport du GF2i sur l'édition scientifique française en Sciences Humaines et Sociales. 2009. Document de synthèse. <http://www.gfii.asso.fr/>

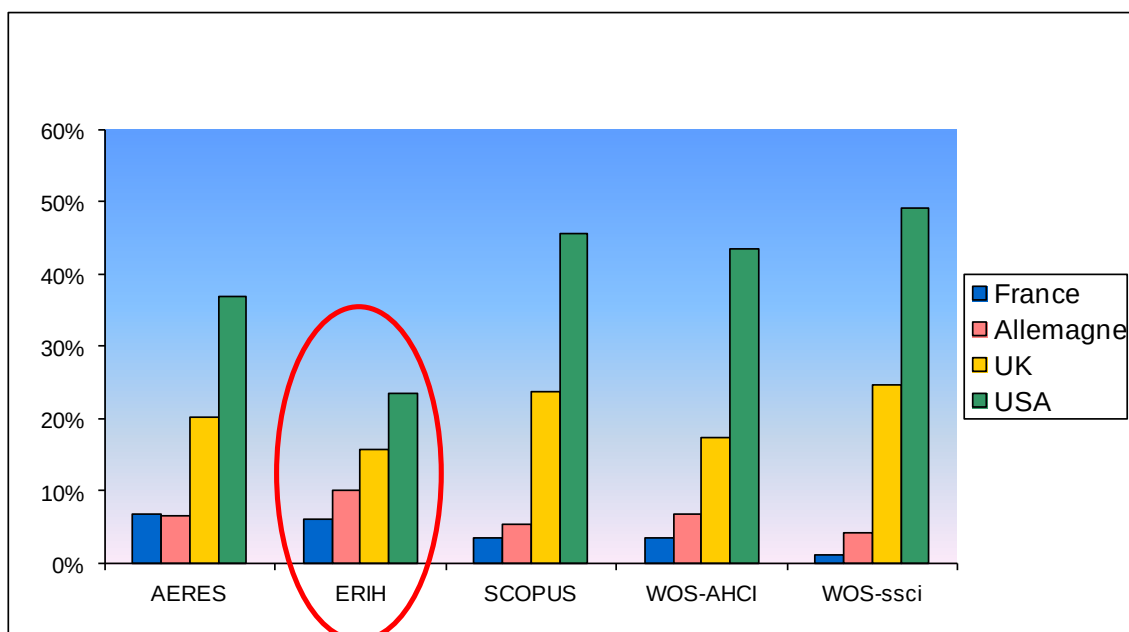


Figure 7 : Nombre de revues par pays et par source

En réalité, la France se trouve ainsi 3ème en Europe, après le Royaume-Uni, et l'Allemagne.

Autant on peut comprendre qu'il y ait beaucoup plus de revues recensées pour les USA, autant on ne saurait expliquer une telle différence entre le Royaume-Uni, la France et l'Allemagne, qui sont tous les trois des pays de taille comparable. La seule différence en nombre de chercheurs ne peut expliquer ces disproportions.

L'Allemagne et la France comportent approximativement le même nombre de chercheurs publics (autour de 100 000). Par contre le Royaume Uni comporte environ 50 % de plus de chercheurs publics, ils sont plus de 150 000 (OST 2010, p.181)¹⁸.

L'Allemagne avec des % assez faibles est tout de même mieux positionnée que la France. Les listes ERIH qui sont plus homogènes et plus européennes, comprennent une plus grosse proportion de revues éditées en l'Allemagne et en France que le WOS et Scopus. Cette figure 7 est importante, parce qu'elle met en évidence le biais des sources qui fournissent les indicateurs bibliométriques. Et donc quand il s'agit de comparer les pays à partir de ces indicateurs comme ça a pu être fait, il faut être prudent puisque entre 60 et 85% des revues françaises en SHS ne sont pas comprises dans ces bases et donc invisibles du point de vue bibliométrique.

Revues soutenues par le CNRS

Sur les 211 revues soutenues par le CNRS entre 2007 à 2010, on en trouve 174 dans JournalBase, soit près de 83 %. Sur la figure 8, on constate qu'une grande proportion de ces revues se trouve dans les AERES (156, soit près de 74 %).

On observe à nouveau une faible proportion de ces revues dans le WOS et Scopus et pratiquement la moitié dans ERIH.

¹⁸ Rapport de l'OST – Indicateurs de sciences et technologie, 2010, Edition Economica

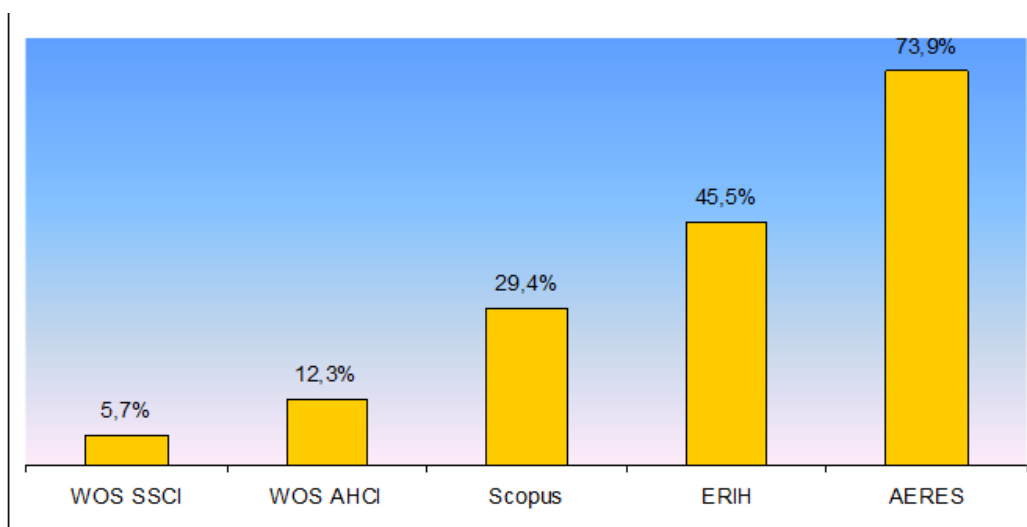


Figure 8 : Proportion de revues soutenues par le CNRS dans Journalbase

Répartition des revues de JournalBase par discipline

Ce sont les disciplines « Psychologie et sciences cognitives », « art et histoire de l'art », « histoire » et « archéologie » qui comportent le plus grand nombre de revues (cf figure 9).

	JB	revues françaises	% de revues françaises	39 revues communes		JB	revues françaises	% de revues françaises	39 revues communes
Anthropologie, Ethnologie et Folklore	425	51	12,0%	4	Littérature	977	68	7,00%	3
Archéologie	1457	137	9,4%	13	Musique et Musicologie	188	7	3,70%	0
Art et Histoire de l'Art	2810	172	6,1%	23	Pédagogie et Sciences de l'Education	1110	85	7,70%	6
Démographie	235	63	26,8%	6	Philosophie	420	27	6,4%	5
Droit	240	1	0,4%	0	Psychologie et sciences cognitives	3646	53	1,50%	16
Economie	997	55	5,5%	3	Relations Internationales	183	3	1,60%	1
Etudes classiques	1372	126	9,2%	12	Religion et Théologie	416	23	5,50%	4
Etudes des Femmes	123	4	3,3%	1	Science Politique	793	47	5,90%	8
Etudes du Genre	135	3	2,2%	0	Sciences commerciales et administratives	1076	56	5,20%	4
Etudes Orientales et Africaines	46	0	0,0%	0	Sciences de la communication	286	3	1,00%	0
Géographie	741	52	7,0%	4	Sciences de l'environnement	929	55	5,90%	7
Histoire	1746	135	7,7%	12	Sociologie	904	66	7,30%	5
Histoire et Philosophie des Sciences	248	14	5,6%	5	Statistiques et Informatique	374	2	0,50%	2
Linguistique	671	48	7,2%	5	Total	10366	572	5,50%	39

Tableau 1 : Nombre total de revues par discipline dans JournalBase et % de revues françaises

Par contre ce sont les disciplines « démographie », « anthropologie » puis « archéologie » et « études classiques » qui ont la plus forte portion de revues françaises.

La proportion moyenne de revues françaises dans JournalBase est de 5,5 % et correspond à 572 revues différentes (cf tableau 1)

24 revues françaises communes aux 4 bases pour les sciences humaines

Il y a 24 revues françaises que l'on retrouve dans les 4 bases (WOS, SCOPUS, ERIH et AERES). On constate que les revues « cœur » françaises SHS ne sont pas uniquement des revues généralistes (cf figure 9). Elles couvrent principalement les sciences humaines puisque ERIH ne répertorie pas les sciences sociales.

Annales de Bretagne et des pays de l'Ouest
Annales historiques de la Révolution Française
Archives de Philosophie
Bulletin hispanique
Bulletin Monumental
Cahiers de Civilisation médiévale
Caravelle: Cahiers du Monde hispanique et luso-brésilien
Etudes Philosophiques
Etudes théologiques et religieuses
Français Moderne
L'Année Psychologique
Le Travail Humain
L'Homme. Revue française d'anthropologie
Linguistique
Mouvement social
Pratiques Psychologiques
Revue biblique
Revue de l'Histoire des Religions
Revue de Linguistique Romane
Revue de Métaphysique et de Morale
Revue des Langues Romanes
Revue des Sciences philosophiques et théologiques
Revue historique
Revue philosophique de la France et de l'Etranger

Figure 9: 24 revues françaises commune aux 4 bases

A noter qu'il y a seulement 24 revues françaises en libre accès sur 586 revues du DOAJ recensées dans JournalBase.

15 revues françaises communes aux 3 bases (sauf ERIH) pour les sciences sociales

Pour prendre en compte également les sciences sociales, nous avons enlevé les revues des listes ERIH, et considéré uniquement les revues communes aux 3 bases (WOS, SCOPUS et AERES). On constate également que les revues « cœur » ne sont pas uniquement des revues généralistes (cf figure 10)

Actes de la Recherche en Sciences Sociales
Cahiers du Monde Russe
Revue Européenne de Psychologie Appliquée
L'Evolution Psychiatrique
Population
Psychologie du travail et des organisations
Revue de Géographie Alpine
Revue Française de Sociologie
Revue Internationale de Psychologie Sociale
Sciences Sociales et Santé
Sociologie du Travail
Travail, genre et sociétés
Urban Morphology

Figure 10 : 15 revues françaises communes aux 3 bases pour les sciences sociales

Répartition des revues par langue de publication

La figure 11 montre la répartition des revues par langue de publication. On constate, sans suspense, que l'anglais est la langue majoritaire de publication des revues de JournalBase (avec 73 % des revues qui sont publiées en anglais). Par contre, ce qui est une bonne surprise, c'est que le français, loin derrière cependant, avec seulement 11 % des revues, est la 2^{ème} langue de publication puis vient l'allemand. L'espagnol, comme langue de publication, ne concerne que 6% des revues, ce qui est assez surprenant.

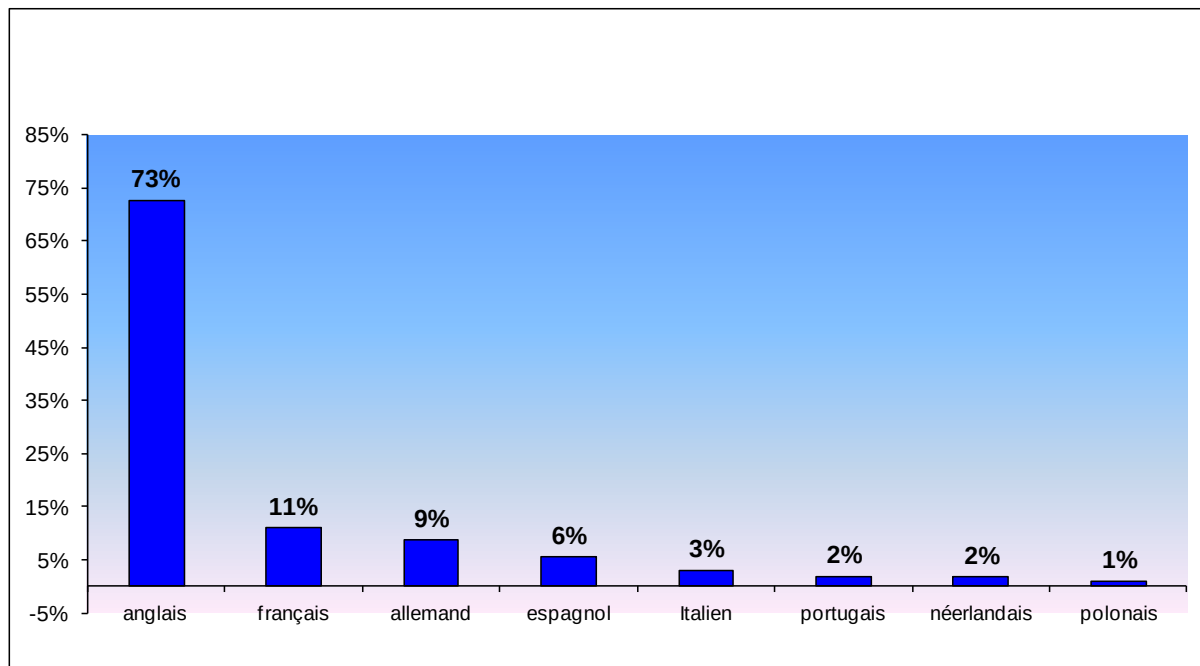


Figure 11 : Répartition des revues par langue

Répartition des revues en français et en anglais par pays d'édition et Répartition des revues en espagnol et en allemand par pays d'édition

Nous avons voulu ensuite regarder, dans les figures 12 à 15, la répartition des revues de langue française, anglaise, espagnole et allemande par pays d'édition.

Là encore, sans surprise, on constate que les revues de langue anglaise (figure 12) sont publiées dans une proportion de 72 % par les pays dont la langue maternelle est l'anglais : Etats-Unis et le Royaume Uni. Cependant, on retrouve ce même phénomène dans les 3 autres pays étudiés :

- Pour les revues en langue française (figure 13), 73 % sont publiées par les pays de langue maternelle français : la France, le Canada, la Belgique et la Suisse.

- Pour les revues en langue espagnole (figure 14), ce phénomène est également confirmé dans les mêmes proportions : 72 % des revues sont publiées par les pays de langue maternelle « espagnol » : Espagne, Chili, Mexique, Argentine, Colombie, Vénézuéla.

- Pour l'Allemagne (figure 15), on observe le même phénomène dans les mêmes proportions : 74 % de revues de langue allemande sont publiées par l'Allemagne, l'Autriche et la Suisse.

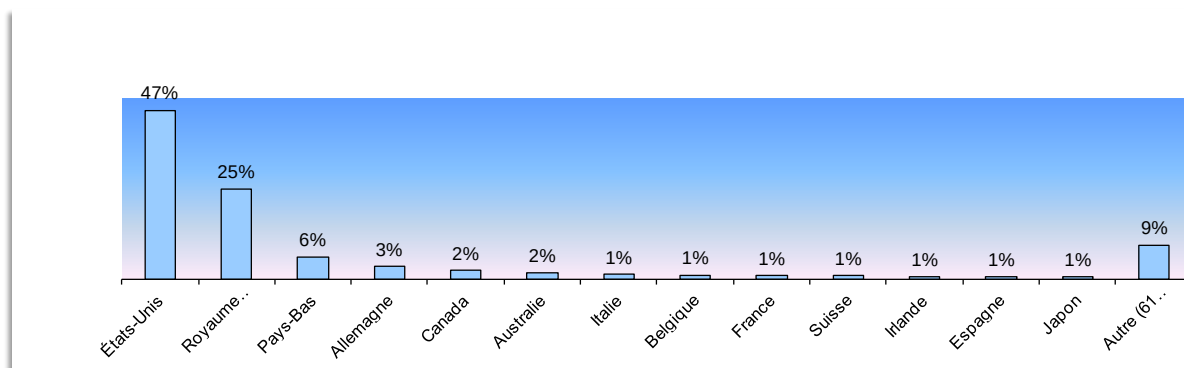


Figure 12 : Répartition des 7599 revues de langue anglaise par pays d'édition

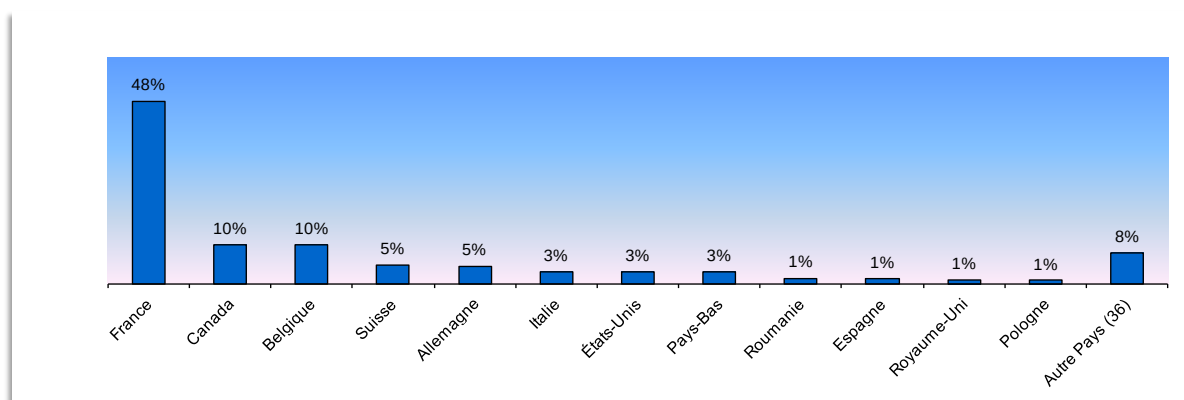


Figure 13 : Répartition des 1153 revues de langue française par pays d'édition

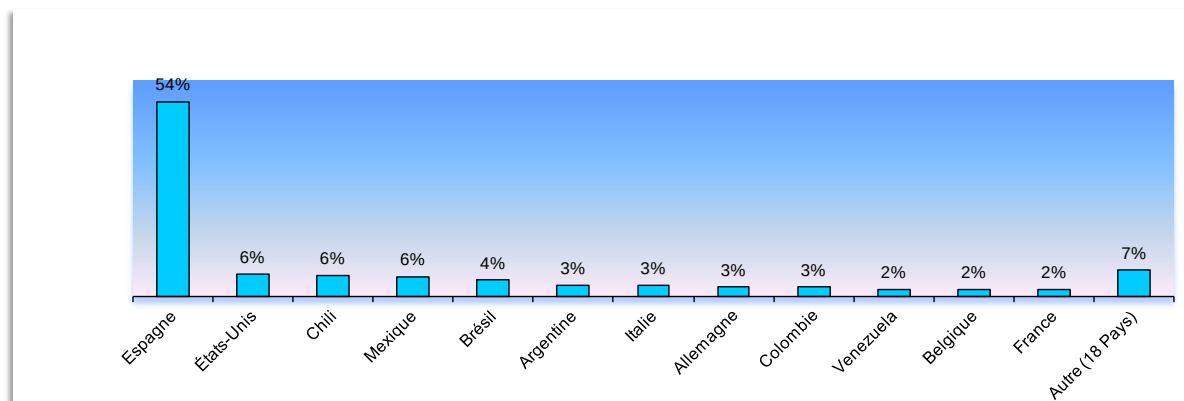


Figure 14 : Répartition des 578 revues de langue espagnole par pays d'édition

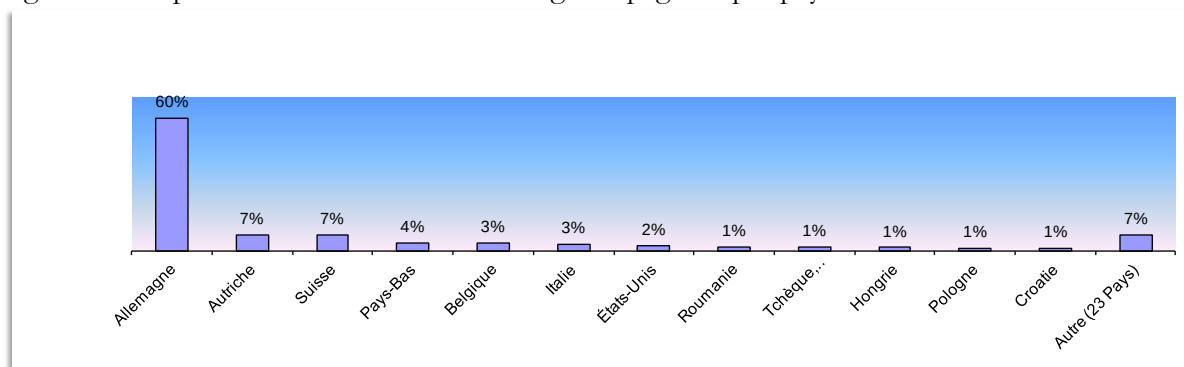


Figure 15 : Répartition des 910 revues de langue allemande par pays d'édition

Rapport entre pays d'édition et langue de publication

En faisant la démarche inverse, c'est-à-dire en étudiant les rapports entre le pays d'édition et la langue, on constate que près de 100 % des revues publiées dans les pays anglo-saxons sont publiées en anglais. Par contre, une proportion importante (environ 37 %) des revues publiées en Allemagne le sont en anglais, contre seulement 11,4 % pour les revues publiées en France et 14,7 % des revues publiées en Espagne (cf figure 17)

	USA	UK	Allemagne	France	Espagne
Anglais	3536	1883	264	67	50
Français	35	10	54	558	15
Allemand	18	5	550	6	4
Espagnol	35	4	15	11	311
Total revues	3552	1888	715	586	340
	USA	UK	Allemagne	France	Espagne
Anglais	99,5%	99,7%	36,9%	11,4%	14,7%
Français	1,0%	0,5%	7,6%	95,2%	4,4%
Allemand	0,5%	0,3%	76,9%	1,0%	1,2%
Espagnol	1,0%	0,2%	2,1%	1,9%	91,5%

Figure 16 : Rapport entre pays d'édition et langue de publication

Conclusion

Outre la confirmation de l'hégémonie de l'anglais et des pays anglo-saxons dans ces sources, nous retirons de nos analyses des points importants jusqu'à là méconnus :

- Les bases ne se recoupent qu'à environ 10%, ce qui signifie que selon la source utilisée, la base statistique à partir de laquelle on mesure les publications des chercheurs est très différente. Cette estimation de 10 % sera toutefois modifiée par l'introduction de données complémentaires, notamment en sciences sociales pour les listes de l'ESF en complément d'ERIH.
- Seulement 572 revues françaises sont présentes dans ces sources alors que le TGE Adonis en avait recensé 2010 dans une étude sur l'édition française dont environ 800 peuvent être considérées comme actives.
- Le français est la deuxième langue de publication après l'anglais (fig.11), de loin certes, mais elle est devant l'espagnol et l'allemand.
- Alors que les revues françaises ne représentent que 5.5% de JournalBase, 11% des revues de JB sont publiées en langue française.
- Les revues françaises sont très peu publiées en langue étrangère, contrairement à l'Allemagne qui approche les 40%.
- ERIH est la base la plus homogène et la plus européenne.

Les données concrètes fournies par JournalBase nous incitent donc à la plus grande prudence dans l'utilisation de ces outils pour l'évaluation des SHS.

En effet, nous constatons que les bases qui produisent ou qui sont destinées à produire de manière automatisée des indicateurs pour l'évaluation sont en l'état actuel lacunaires et ne peuvent qu'entraîner des résultats biaisés.

Il est prévu que JournalBase soit mis à jour tous les ans. Nous pourrions alors évaluer l'évolution de la représentativité de ces outils.

Toutefois, outre l'étude comparative qui a nourri ce projet, le site web de JournalBase peut être également considéré comme un outil de valorisation de l'édition scientifique en SHS. Grâce à sa très riche base de données, les chercheurs pourront y trouver des revues de leur discipline ou au carrefour de leur thématique, qu'ils ne connaissent peut-être pas... Ils peuvent également prendre connaissance des revues d'autres pays européens dans leur domaine.

Les indicateurs de la recherche en SHS

Jacques Dubucs

Ministère de l'Enseignement Supérieur et de la Recherche

L'évaluation des productions scientifiques : des innovations en SHS ?

CNRS

9-10 juin 2011

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

1 Avons-nous besoin d'indicateurs ?

2 De quels indicateurs avons-nous besoin ?

- Contexte et contraintes
- Les bases nouvelles
- Le travail en cours
- Questions particulières

L'état de l'opinion

- 1 Pleinement satisfait
- 2 D'accord sur le principe, mécontent des indicateurs existants
- 3 Hostile au principe

L'état de l'opinion

- 1 Pleinement satisfait 0%
- 2 D'accord sur le principe, mécontent des indicateurs existants
- 3 Hostile au principe

L'état de l'opinion

- 1 Pleinement satisfait 0%
- 2 D'accord sur le principe, mécontent des indicateurs existants 85%
- 3 Hostile au principe 15%

L'état de l'opinion

- 1 Pleinement satisfait 0%
- 2 D'accord sur le principe, mécontent des indicateurs existants 85%
- 3 Hostile au principe 15%

Deux questions

- 1 Avons-nous besoin d'indicateurs ? (pour les 15%)
- 2 De quels indicateurs avons-nous besoin ? (pour les autres)

En SHS, pas besoin d'indicateurs ?

Hyper-Spécifisme

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

■ Credo

- Les SHS sont spécifiquement spécifiques
- Les SHS diffèrent des autres sciences plus que les autres sciences ne diffèrent entre elles
- Les SHS ne sont pas une espèce du genre “science”

En SHS, pas besoin d'indicateurs ?

Hyper-Spécifisme

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

■ Credo

- Les SHS sont spécifiquement spécifiques
- Les SHS diffèrent des autres sciences plus que les autres sciences ne diffèrent entre elles
- Les SHS ne sont pas une espèce du genre “science”
- “Science” figure dans “Sciences de l’Homme et de la Société” comme “chat” dans “achat” : en vertu d’un simple accident orthographique

En SHS, pas besoin d'indicateurs ?

Hyper-Spécifisme

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

- Credo
 - Les SHS sont spécifiquement spécifiques
 - Les SHS diffèrent des autres sciences plus que les autres sciences ne diffèrent entre elles
 - Les SHS ne sont pas une espèce du genre “science”
 - “Science” figure dans “Sciences de l’Homme et de la Société” comme “chat” dans “achat” : en vertu d’un simple accident orthographique
- Conséquence ; la recherche d’indicateurs SHS comparables aux indicateurs qui ont cours dans les autres secteurs scientifiques est futile

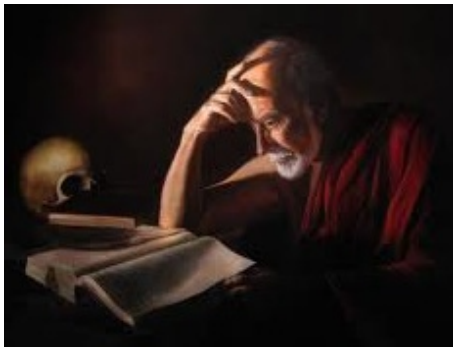
Hyper-spécifisme

Assumer ses choix

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?



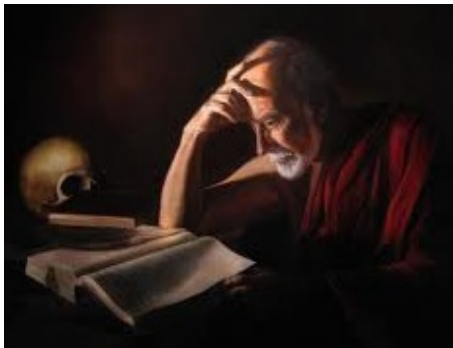
Hyper-spécifisme

Assumer ses choix

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?



- Saint-Jérôme : pas besoin d'infrastructures

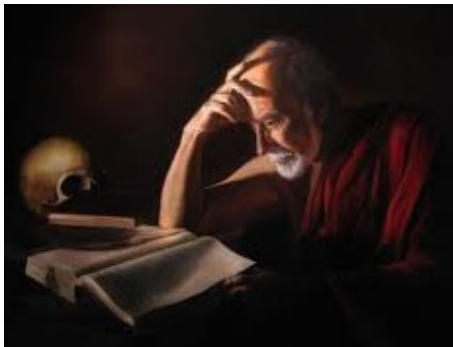
Hyper-spécifisme

Assumer ses choix

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?



- Saint-Jérôme : pas besoin d'infrastructures
- Difficile d'être à la fois hypo-spécifiste dans l'obtention des ressources et hyper-spécifiste dans leur utilisation

De quels indicateurs avons-nous besoin ?

Contexte et contraintes

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

1 Centralité des SHS

- Pour des maillons de plus en plus larges des chaînes causales étudiées par les sciences de la nature, les déterminants sont désormais humains



Extraction de sables bitumineux de Syncrude (Alberta) : chaque année plus de $30 \cdot 10^{12}$ tonnes de matériaux y seront déplacés (2 fois la masse annuelle des sédiments fluviaux du monde entier)

116/148

Contexte et contraintes

1. Centralité des SHS (suite)

- Puisque les hommes changent la manière dont se comporte la nature, il faut changer la manière dont nous analysons la nature

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

Contexte et contraintes

1. Centralité des SHS (suite)

- Puisque les hommes changent la manière dont se comporte la nature, il faut changer la manière dont nous analysons la nature
- SHS-isation des sciences de la nature : non pas mélangisme sans principe, mais nécessité de transactions scientifiques d'un type inédit

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

Contexte et contraintes

1. Centralité des SHS (suite)

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

- Puisque les hommes changent la manière dont se comporte la nature, il faut changer la manière dont nous analysons la nature
- SHS-isation des sciences de la nature : non pas mélangisme sans principe, mais nécessité de transactions scientifiques d'un type inédit
- Interdisciplinarité : non pas sociabilité académique diffuse, mais coordination organique de compétences et de régimes méthodologiques distincts autour d'objectifs scientifiques définis
- Ne nécessite aucunement le renoncement aux modes explicatifs usuels des SHS (représentations, préférences, mémoire, traditions) : le déterminant des actions demeure l'environnement perçu, non l'environnement distal
- Mais demande un effort sans précédent de **119/148** et d'exotéricité

De quels indicateurs avons-nous besoin ?

Contexte et contraintes

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

2 Internationalisation des SHS

- Accroissement de l'audience des SHS françaises
 - Problème quelque peu exagéré : la barrière linguistique
 - Sapir-Whorfisme : plus de choses en commun entre deux Allemands dont l'un dit "Es regnet" et l'autre "Es regnet nicht" qu'entre le premier et un Français qui dit : "il pleut"

De quels indicateurs avons-nous besoin ?

Contexte et contraintes

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

2 Internationalisation des SHS

- Accroissement de l'audience des SHS françaises
 - Problème quelque peu exagéré : la barrière linguistique
 - Sapir-Whorfisme : plus de choses en commun entre deux Allemands dont l'un dit "Es regnet" et l'autre "Es regnet nicht" qu'entre le premier et un Français qui dit : "il pleut"
- Accroissement de la coopération internationale en SHS
 - La convergence scientifique suppose la commensurabilité des descripteurs (cf Enquête européenne METRIS)

De quels indicateurs avons-nous besoin ?

Contexte et contraintes

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

3 Le décloisonnement du dispositif de recherche

- CNRS et Universités : vers un même régime d'évaluation (AERES)
- Nécessité d'une lisibilité mutuelle et d'indicateurs communs

De quels indicateurs avons-nous besoin ?

Contexte et contraintes

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

3 Le décloisonnement du dispositif de recherche

- CNRS et Universités : vers un même régime d'évaluation (AERES)
- Nécessité d'une lisibilité mutuelle et d'indicateurs communs

4 Les priorités nationales de la SNRI, l'Investissement d'Avenir

- Nécessité d'un repérage d'ensemble de l'activité SHS, précédemment illisible et indétectable (7 taxinomies distinctes ...)

De quels indicateurs avons-nous besoin ?

Les bases nouvelles

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

1 L'Alliance ATHENA

- Impératif de décloisonnement : tous les opérateurs SHS y sont représentés, y inclus ceux où domine la recherche applicative
- Impératif d'interdisciplinarité : la plupart des autres Alliances y sont représentées

De quels indicateurs avons-nous besoin ?

Les bases nouvelles

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

1 L'Alliance ATHENA

- Impératif de décloisonnement : tous les opérateurs SHS y sont représentés, y inclus ceux où domine la recherche applicative
- Impératif d'interdisciplinarité : la plupart des autres Alliances y sont représentées

2 Une nouvelle nomenclature

- Commune à tous les laboratoires et opérateurs SHS et à toutes les formations doctorales SHS
- Convergente avec celle de nos partenaires européens (ERC)
- Désormais utilisée dans les documents stratégiques de base (STRATER)

De quels indicateurs avons-nous besoin ?

Le travail en cours

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

Objectif du groupe de travail ATHENA/DGRI

- Caractéristiques des indicateurs SHS à définir
 - Corrélation adéquate avec le trait à indiquer
 - Aptitude à recueillir l'assentiment de la communauté SHS
 - Convergence entre les opérateurs de l'Alliance ATHENA
 - Commensurabilité avec les indicateurs utilisés par les partenaires européens du domaine
 - Capacité à être utilisés à parité avec les indicateurs en vigueur dans les autres domaines de la recherche scientifique

De quels indicateurs avons-nous besoin ?

Le travail en cours

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

Objectif du groupe de travail ATHENA/DGRI

- Caractéristiques des indicateurs SHS à définir
 - Corrélation adéquate avec le trait à indiquer
 - Aptitude à recueillir l'assentiment de la communauté SHS
 - Convergence entre les opérateurs de l'Alliance ATHENA
 - Commensurabilité avec les indicateurs utilisés par les partenaires européens du domaine
 - Capacité à être utilisés à parité avec les indicateurs en vigueur dans les autres domaines de la recherche scientifique
- Autres questions à examiner
 - Fiabilité du recueil des données pertinentes à l'échelle nationale
 - Dispositifs de nature à rendre ces données disponibles à l'ensemble des acteurs concernés
 - Dispositifs de critique et d'actualisation des indicateurs retenus
 - Utilisation de ces indicateurs dans les activités de prospective, de caractérisation et d'évaluation en SHS

127/148

De quels indicateurs avons-nous besoin ?

Questions particulières

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

- 1 Loi de Campbell : “Plus on utilise un indicateur social quantitatif à des fins de prise de décision, plus il est sujet à des pressions corruptives et plus il est capable de distordre et de corrompre le processus social qu’il est censé piloter” (Donald T. Campbell, “Assessing the Impact of Planned Social Change”, *Evaluation and Program Planning*, II-1, 1979, p 67-90)

De quels indicateurs avons-nous besoin ?

Questions particulières

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes
Les bases
nouvelles
Le travail en
cours

Questions
particulières

1 Loi de Campbell : “Plus on utilise un indicateur social quantitatif à des fins de prise de décision, plus il est sujet à des pressions corruptives et plus il est capable de distordre et de corrompre le processus social qu’il est censé piloter” (Donald T. Campbell, “Assessing the Impact of Planned Social Change”, *Evaluation and Program Planning*, II-1, 1979, p 67-90)

- Notion générale de “ I indique que φ ”
 - Fiabilité : si I se produit, φ se produit
 - Luminosité : l’occurrence de I doit être facilement détectable, et plus que ne le serait directement celle de φ
- Exemples : hauteur du mercure/température, démonstration/vérité mathématique
- Les indicateurs de performance scientifique tendent mécaniquement à perdre leur fiabilité : les acteurs tendent à les réaliser sans réaliser la performance qu’ils mesurent
- Remèdes possibles
 - 1 Actualisation régulière
 - 2 Non-détachabilité des attributions (exemple : faire toujours précéder les notations des motifs qui ont conduit à les attribuer)

De quels indicateurs avons-nous besoin ?

Questions particulières

Les
indicateurs de
la recherche
en SHS

Avons-nous
besoin
d'indicateurs ?

De quels
indicateurs
avons-nous
besoin ?

Contexte et
contraintes

Les bases
nouvelles

Le travail en
cours

Questions
particulières

2 Articulation verticale des indicateurs :

- Chercheurs individuels
 - Equipes
 - Laboratoires
 - Universités, sites
 - Ensemble national
-
- Rapport entre les indicateurs correspondants ?
 - Primes d'agrégation ?
 - Diversité des instances d'évaluation

L'évaluation scientifique en sciences humaines et sociales : questions méthodologiques et pistes pour des réponses

Ghislaine Filliatreau, OST

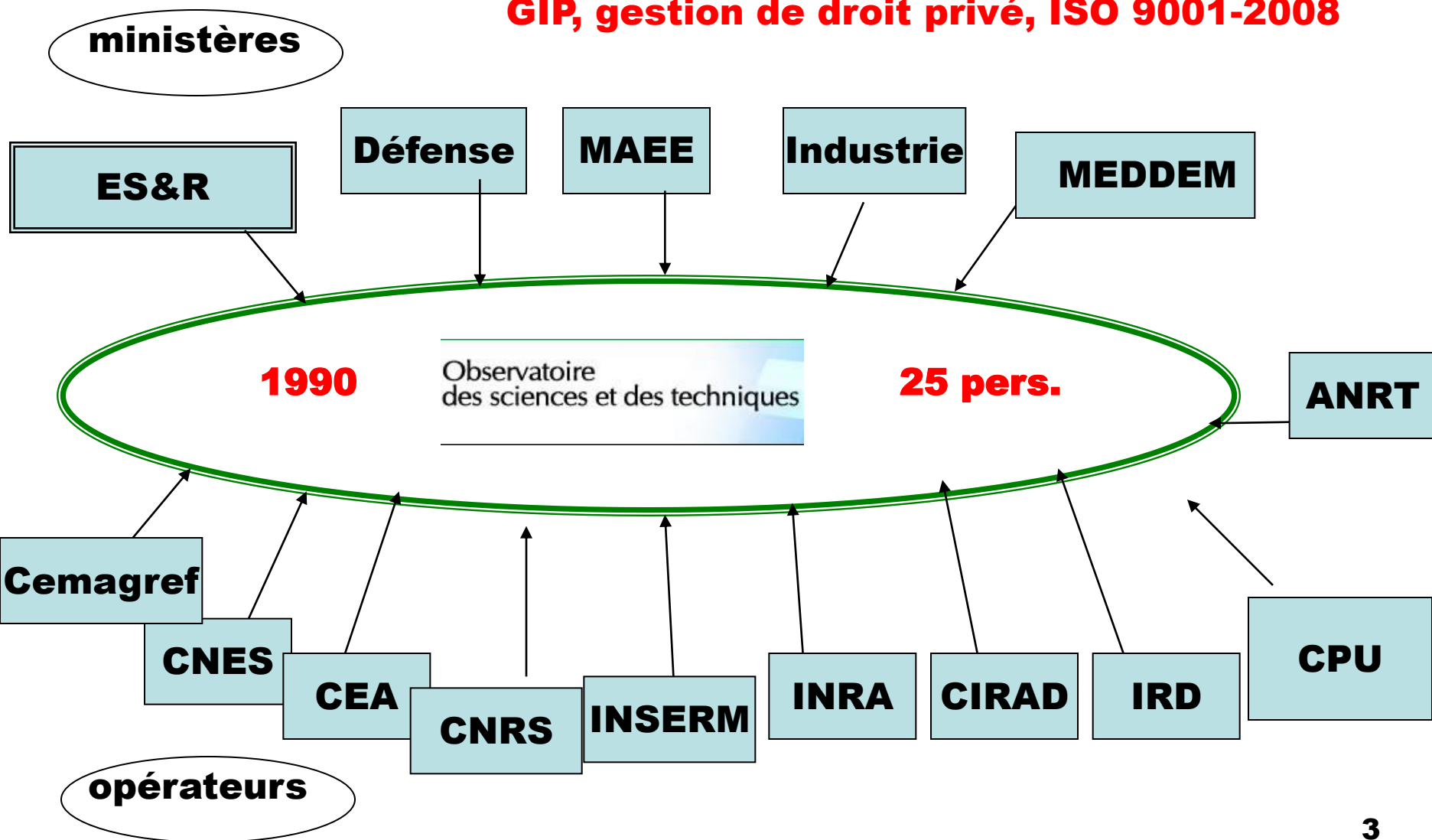
10 juin 2011, colloque CNRS Paris

Plan

- ▲ Brève présentation de l'OST
- ▲ Quelques rappels historique sur la bibliométrie
- ▲ Quelle(s) représentativités ?
- ▲ L'évaluation scientifique /l'évaluation de la science

L'OST, plate-forme « inter-niveaux » dédiée à la conception, la production et la diffusion d'indicateurs relatif à la science et la technologie

GIP, gestion de droit privé, ISO 9001-2008



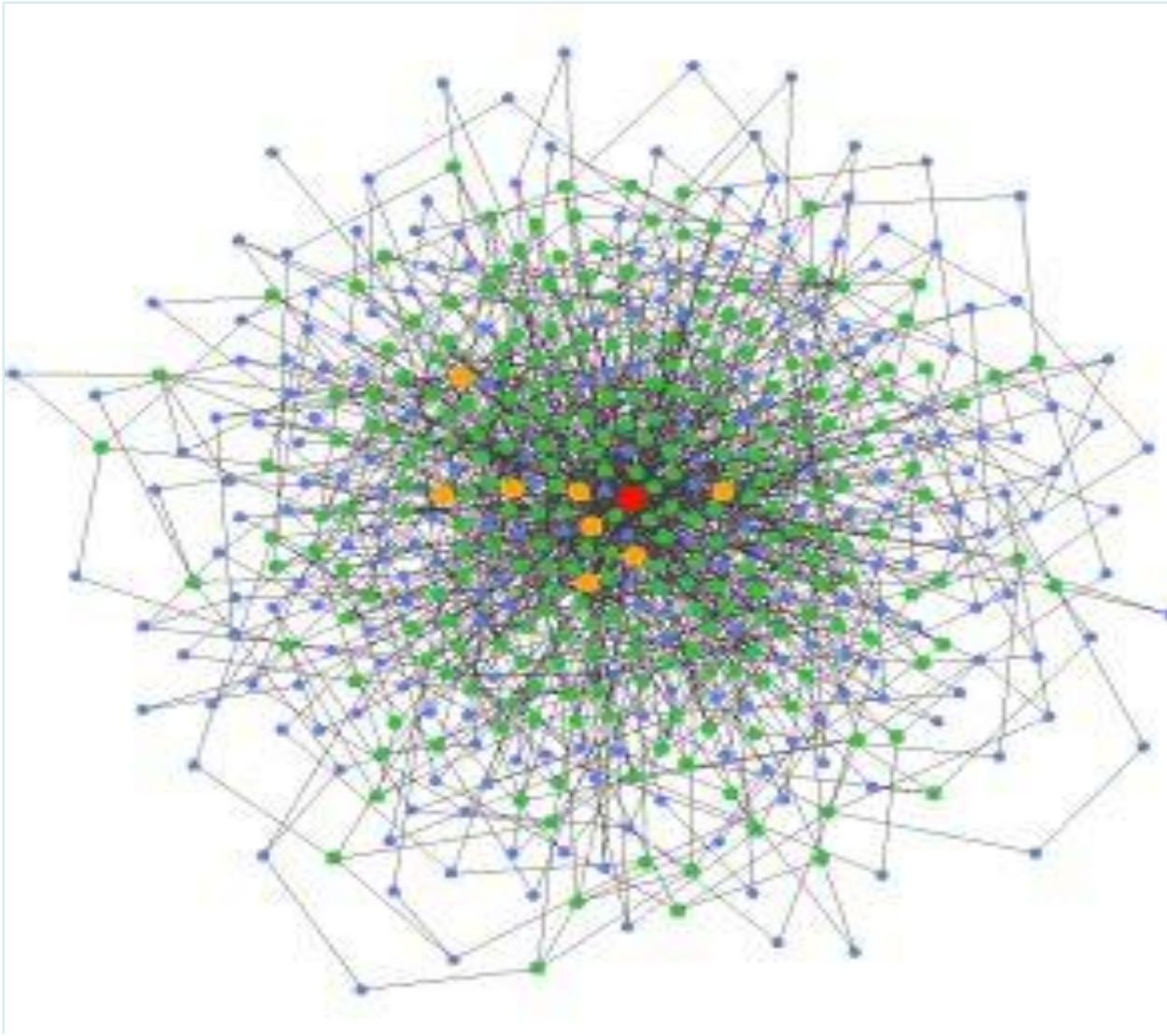
20 ans d'expérience

- **1) la notion d'Observatoire : dédié aux *indicateurs* quantitatifs pour la recherche, le développement et l'innovation**
 - en complément et à partir des statistiques R&DI (OCDE, ministères, OEB, Vos) : connaissance des données et capacité de concevoir des nomenclatures
 - grâce à une *infrastructure de production* : ingénierie de base de données, liberté de conception des indicateurs et des nomenclatures, chiffres robustes et comparatifs (ISO 9001:2008),
 - avec une mission nationale de surveillance permanente de la *position relative* R&DI des acteurs, par des indicateurs *comparatifs et leur contextualisation dans différents espaces de référence*
 - veille sur les politiques de recherche et des méthodes d'évaluation
 - capacité de collaborer avec la recherche (Ifris, ENID, Eurolio, le Beta, Paris 1, Paris 11, Bordeaux 4 etc...)
- **2) l'identité de GIP : notion de plate-forme partagée dédiée au travaux d'intérêt commun**
 - compréhension des besoins des acteurs et culture de service (nécessité de ressources sur contrats, position de plate-forme indépendante et de tiers de confiance)
 - capacité d'innover grâce aux mécanismes Coopératifs (acteur-réseau) et aux liens avec la recherche (appui sur des experts, collaborations recherche)
 - production organisée vers l'appui *opérationnel* aux décideurs
 - indicateurs conçus et utilisés pour des constats partagés : légitimité construite sur l'expertise professionnelle et l'exigence méthodologique (modèle de la statistique *scientifique* + pratique de la validation par des *experts*)
 - coopérations structurées par des outils communs («étude pilote SHS).

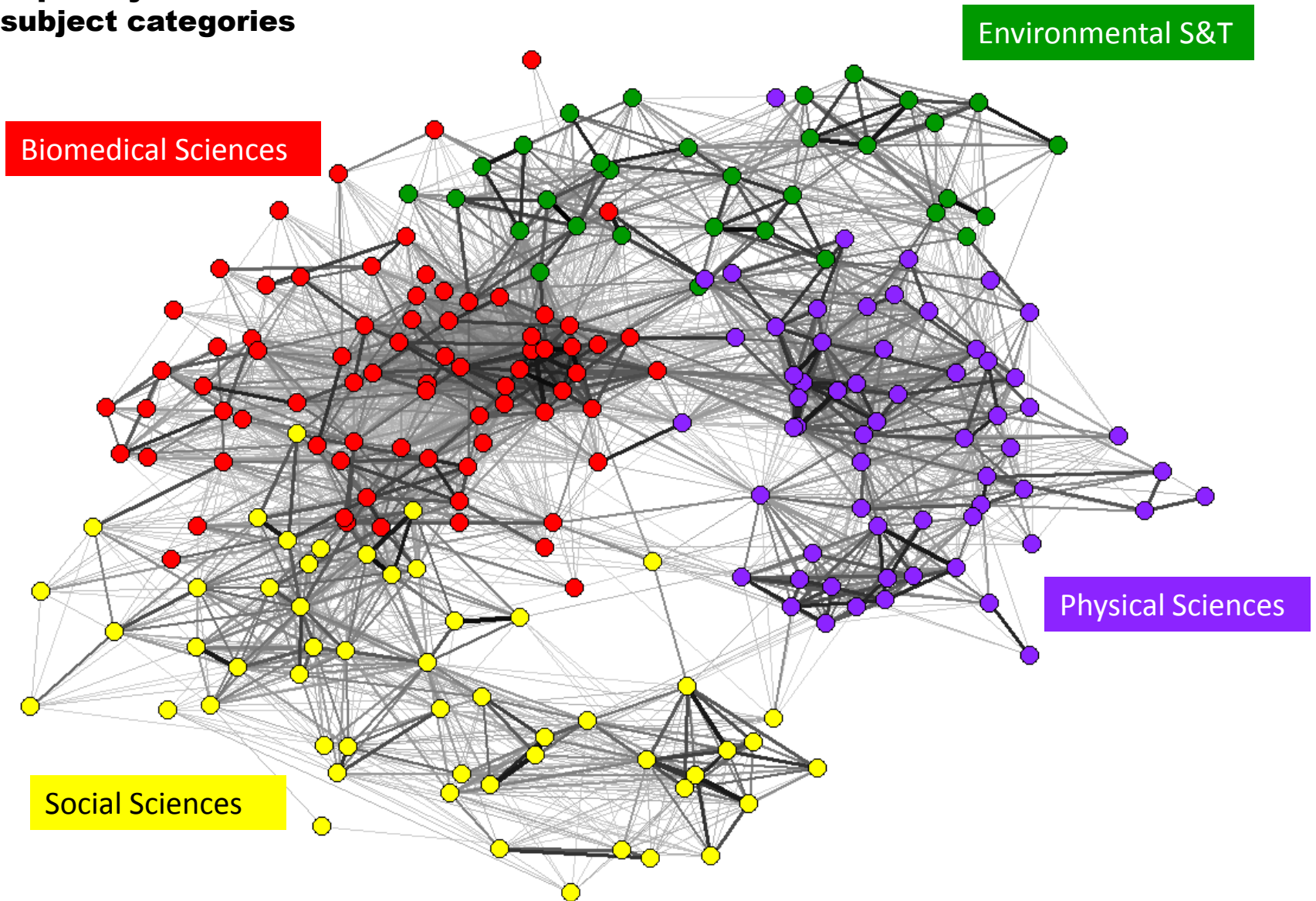
Exemples de travaux

- ☐ les indicateurs de la LOLF
 - ☐ les indicateurs pour les stratégies à l'international du MESR
 - ☐ les services tableau de bord et IPERU,
 - ☐ études spécifiques : institutionnelles et territoriales
 - ☐ l'étude méthodologique pour l'évaluation des pôles de compétitivité (DATAR)
 - ☐ les études thématiques (génomique, cancer, microbiologie, STIC, recherche spatiale),
 - ☐ les classements (U-Multirank, Shanghai)
 - ☐ les études cœur de métier (InformIn)
-
- ☐ dérouler : **<http://www.obs-ost.fr/fr/le-savoir-faire/etudes-en-ligne.htm>**



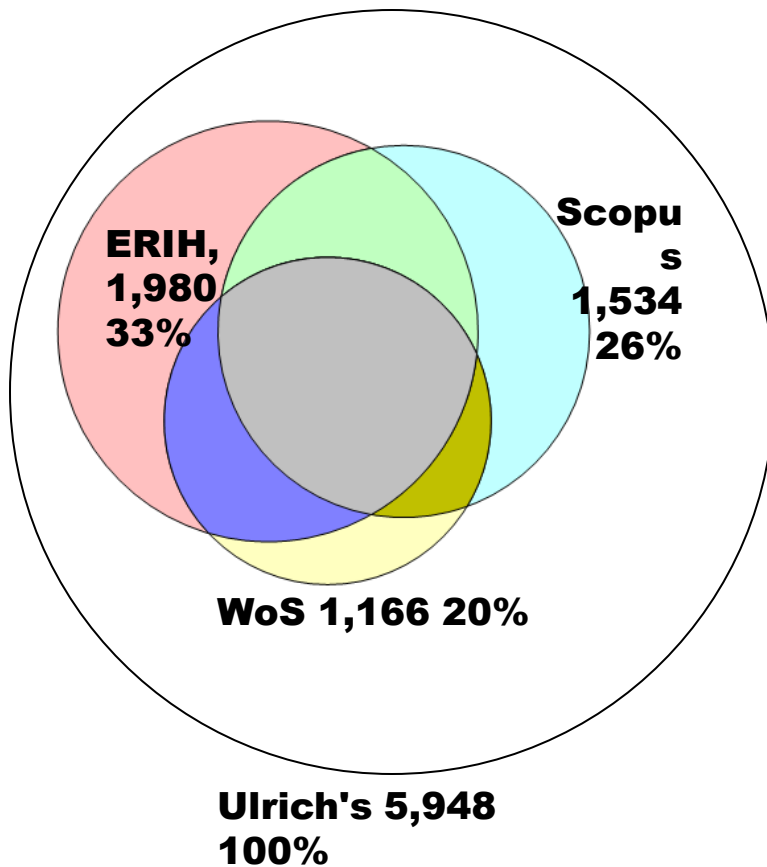


Global Map of Science 2009
4 disciplinary areas
222 subject categories

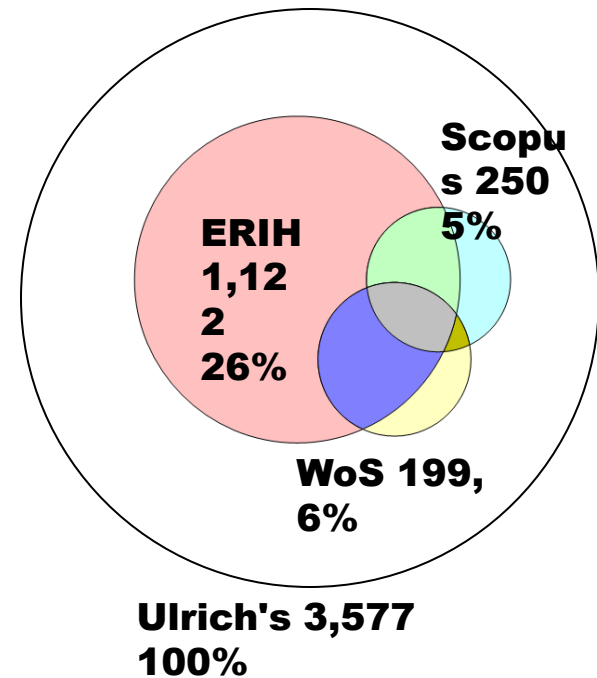


ERIH coverage

English language



European language, not English



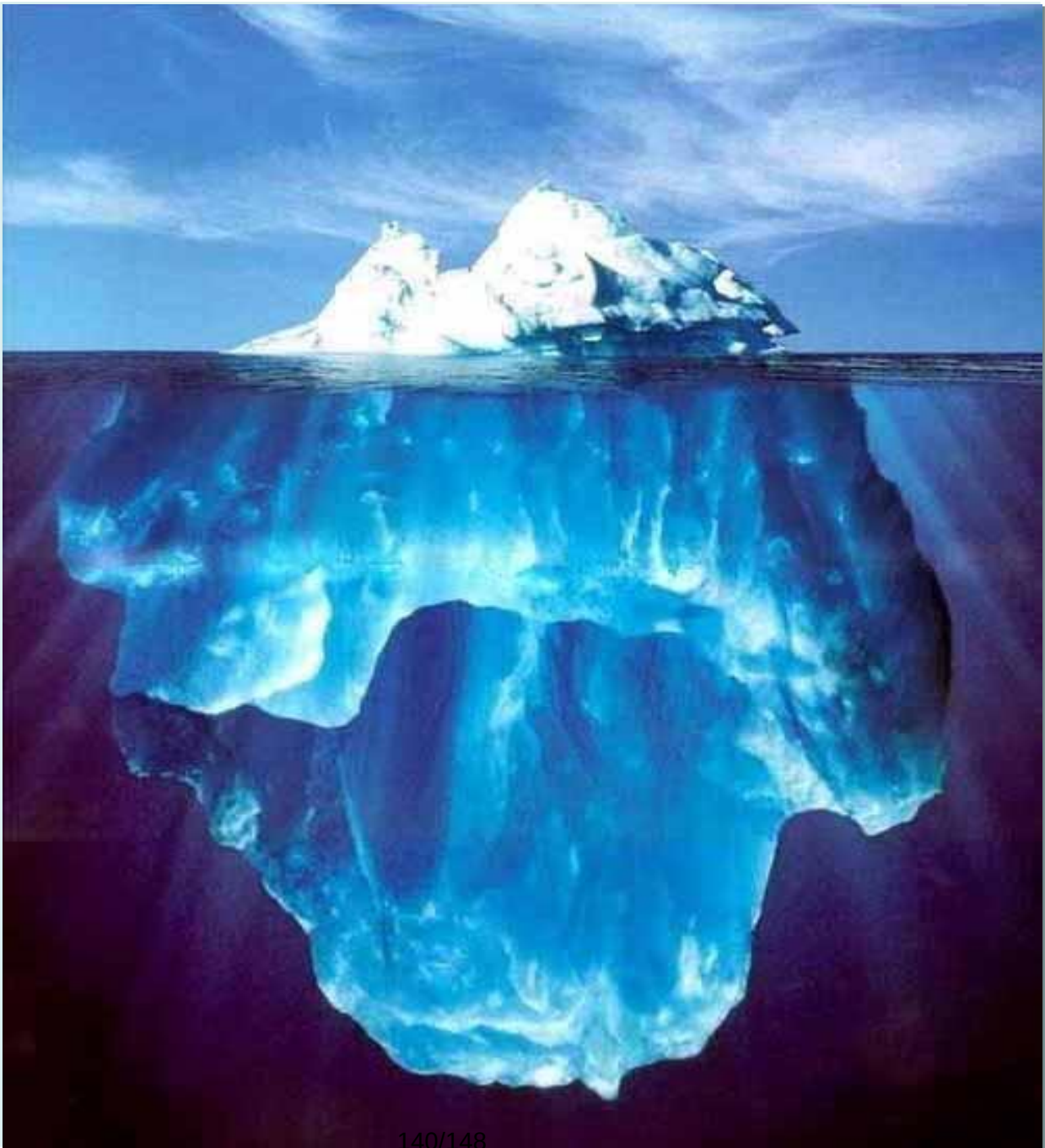
**Journals counted are:
ERIH fields**

**active (not ceased) & published regularly
identified as academic/scholarly by Ulrich's
published in a European country or U.S.**



La multiplication des modalités de circulation des connaissances scientifiques rend de plus en plus dangereuse l'idée d'une base exhaustive, et de plus en plus nécessaire le contrôle permanent de la représentativité (en tant que source de publication et en tant que source de citations) des bases utilisées comme proxies.

Ex. : bases institutionnelles et individuelles, bases de dépôt libre, bases de pre-prints, archives publiques, moteurs de recherche, listes de journaux, blogs et autres outils de partage d'information, bases factuelles, carnets de recherche, bases d'aggregation etc...



Trois pistes liées aux questions de représentativité :

- ☐ Représentativité « horizontale » (ie en terme de couverture) des bases utilisées :
 - ☐ Pour les journaux et autres documents : s'appuyer sur l'expertise des communautés de recherche
 - ☐ Pour vérifier factuellement l'acceptabilité des bases utilisées (en termes de principes de couverture et en termes d'indicateurs calculés) et pour les enrichir quand cela est possible
 - ☐ Pour créer des bases de revues d'ouvrages au fur et à mesure de leur mise en place par ces communautés (modèle : *mathematical reviews*)
- ☐ Représentativité verticale (c'est-à-dire proportionnalité avec les diverses activités liées aux missions de recherche) des comparaisons faites sur la base des indicateurs
 - ☐ études de cas en travaillant avec les communautés
- ☐ Représentativité pour les usages d'évaluation : évaluation scientifique/évaluation socio-économique
 - ☐ La bibliométrie est une évaluation scientifique « endogène », des études d'impact doivent être développées par ailleurs pour rendre compte de l'utilité socio-économique des recherches.

- L'OST, une équipe
 - www.obs-ost.fr

La production scientifique en sciences humaines et sociales au prisme de l'évaluation par l'AERES (France)

Pierre Glaudes

Délégué scientifique coordinateur des Sciences humaines et sociales
Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES)

Les principes

En guise de préliminaire, on voudrait indiquer quelques principes qui déterminent les pratiques d'évaluation à l'Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES)¹. Il convient en effet de rappeler que l'évaluation, telle que la comprennent les acteurs de l'Agence, doit être faite par les pairs et se fonder sur des processus raisonnés, transparents portant sur plusieurs critères : une évaluation pertinente doit être « multicritérisée » et mettre en jeu des batteries d'indicateurs. On considère en effet qu'il est vain de chercher l'indicateur universel, performant, discriminant². L'AERES utilise donc trois types d'indicateurs pour l'évaluation des unités de recherche : les indicateurs de *caractérisation* (typologie des unités, ventilation des effectifs, surfaces disponibles, etc.) ; les indicateurs d'*environnement* (type d'établissement ; contexte socio-économique ; taille du site urbain) ; les indicateurs de *performance*, qui concernent le plus le sujet de ce colloque.

Ces derniers indicateurs se distribuent en quatre sous-catégories : des indicateurs de *production scientifique*, qui doivent être à la fois qualitatifs et quantitatifs (c'est en ce point que s'active l'approche bibliométrique) ; des indicateurs d'*impact*, de couplage aux demandes et aux enjeux sociaux (ce qui suppose que l'on trouve le moyen de préciser qui est touché par les activités de recherche d'une unité – les milieux académiques, les étudiants, de plus larges groupes sociaux – et comment il l'est) ; des indicateurs de *gouvernance* ; des indicateurs de *stratégie* et de *projet*.

Ces grands types d'indicateurs se déclinent en indicateurs secondaires, qui mettent en jeu un ensemble d'items. Par exemple, la gouvernance, qui relève des indicateurs de performance, prend en compte *trois indicateurs secondaires* (la démocratie interne, la vie de l'unité, la qualité de la communication interne et externe), qui se démultiplient en une série de critères d'évaluation. Pour la démocratie interne : la configuration de l'équipe de direction, composition du conseil de laboratoire, la fréquence de ses réunions, la procédure d'affectation de crédit, etc. Pour la vie de l'unité : les types de séminaires, leur fréquence, leur périmètre, l'existence ou non d'un séminaire spécifique pour les doctorants, le mode d'organisation de celui-ci, etc. Pour la qualité de la communication interne et externe : l'existence de comptes rendus du conseil de laboratoire, d'une newsletter, d'un site web, etc.

Cet ensemble d'indicateurs pour être opératoire est assujéti à quelques règles. Les indicateurs de l'AERES sont des instruments de mesure portant sur des *structures collectives* (et non pas sur des individus), qui n'ont de valeur qu'en fonction d'un *objectif* qui doit toujours être clarifié au préalable (importance de la question « pour quoi faire ? »). Ils sont considérés comme des mesures dont la durée de validité est *limitée*, car ils sont fonction de l'évolution des milieux évalués, chaque indicateur, lorsqu'il est intégré par une communauté donnée, étant aussitôt perçu comme un objectif à atteindre et pouvant, de ce fait, devenir plus ou moins rapidement peu discriminant, donc obsolète (cf. la situation actuelle du taux de produisant). Ils sont par principe susceptibles d'*évoluer* et d'*être renouvelés* (exemple : dans un récent séminaire, on s'interrogeait sur l'opportunité d'intégrer parmi les critères de gouvernance une mesure du bonheur, de la

¹ <http://www.aeres-evaluation.fr/>

² On voit bien à cet égard les limites de la « note globale » dont le maintien et/ou la transformation sont en cours de discussion au sein du conseil de l'Agence.

convivialité, de la qualité du vivre ensemble au sein des unités de recherche). Ils combinent évaluation *quantitative* et *qualitative*, et, de ce fait, s'appuient sur des données bibliométriques.

Les listes de revues

Dans ce contexte, qu'en est-il de l'approche bibliométrique en Sciences humaines et sociales à l'Agence d'évaluation de la recherche et de l'enseignement supérieur ? Cette approche est à situer dans une réflexion globale sur la production scientifique. Deux facteurs doivent être pris en considération. Le premier est *l'objet de l'évaluation* par l'AERES. Celui-ci n'est pas les individus, on l'a dit, mais les groupes, parmi lesquels on distingue trois niveaux de « granularité », qui sont fonction de la taille : les petits groupes (5-15 personnes), soit le niveau des équipes réunies dans une unité de recherche ; les groupes moyens à grands (50 à 150 personnes), soit le niveau de l'unité de recherche ou du département au sein d'un établissement ; les très grands groupes (établissements, régions, pays). L'essentiel de l'évaluation faite par l'AERES porte sur les deux premiers niveaux, même si des rapports de vague et une synthèse quadriennale ont été tentés au niveau le plus large.

Le deuxième facteur à prendre en compte est *la visée de l'évaluation*. Dans l'évaluation des petits groupes et des groupes moyens à grands, l'AERES se fixe, en matière de production scientifique trois objectifs : produire des signaux d'alerte ; détecter l'émergence ; distinguer la consécration. C'est dans ce contexte que doit être située, parmi d'autres indicateurs de performance, l'approche bibliométrique, laquelle, en Sciences humaines et sociales, comme dans les autres domaines (Sciences et technologies, Sciences de la vie et de l'environnement) porte essentiellement sur les revues. Ce point soulève plusieurs problèmes.

Pour de nombreuses disciplines de Sciences humaines et sociales, notamment du côté des Humanités (lettres, langues, arts, philosophie), la production *d'ouvrages individuels* (ouvrages scientifiques, éditions de textes, traductions) ou *collectifs* (actes de colloque) reste la forme majeure de l'activité scientifique (à quoi s'ajoutent des spécificités liées à certaines disciplines : les productions artistiques théorisées pour les arts ; les rapports de fouille pour l'archéologie, etc.), dont il résulte une extrême diversité des productions scientifiques.

En l'absence de listes de revues internationales satisfaisantes pour l'ensemble des Sciences humaines et sociales en France – du fait notamment d'un déséquilibre, dans les listes existantes (*Web of Science*, *Scopus*...), jouant en faveur des revues anglo-saxonnes au détriment des revues de langue française –, l'AERES a considéré qu'il lui incombait d'intervenir dans ce domaine. Elle s'est fixé comme objectif, dès sa création en 2007, de rendre compte de la profusion internationale des périodiques, qui témoigne de la croissance de la communauté mondiale des chercheurs et de l'évolution dans la manière de publier les résultats de la recherche.

Elle a donc entrepris de dresser, par champ disciplinaire ou par domaine, des listes de revues scientifiques tenant une place importante dans la diffusion des résultats de la recherche à l'échelle nationale et surtout internationale, – listes qui ont été fondues dans une liste générale des revues en Sciences humaines et sociales publiée sur le site de l'Agence. Des commissions disciplinaires regroupant des représentants des sections du CoNRS et du CNU, des membres de l'IUF, des personnalités qualifiées étrangères et les délégués scientifiques de l'AERES ont été constituées en 2008 avec pour mission de tenter de parvenir à un consensus dans chaque communauté ; elles ont abouti à l'élaboration des premières listes, publiées sur le site de l'Agence. Ces listes sont actualisées chaque année en tenant compte du degré d'internationalisation, de l'organisation et des pratiques des milieux de chaque discipline : les commissions d'actualisation répondent aux sollicitations des périodiques, récemment créés ou non sélectionnés, et prennent en compte des revues en ligne.

La disparité entre les listes

Les conditions dans lesquelles ces listes particulières et cette liste globale ont été constituées ont abouti à un résultat mitigé, dont on ne peut pleinement se satisfaire. Dans le contexte de la politique volontariste d'une Agence récemment créée et soucieuse d'être opérationnelle dans les meilleurs délais, chaque domaine (ou chaque discipline) a été invité à établir des listes, sans qu'une méthodologie commune ait été fixée ni qu'une réflexion collective suffisante sur les critères de définition d'une « revue scientifique » ait été considérée comme un préalable. De là des résultats d'une extrême disparité.

Certaines disciplines, comme l'économie et la gestion, *ont simplement repris et combiné des listes établies ailleurs* (liste de la section 37 du CNRS, la liste *Econlit : Expanded Journal List* et la liste *Harzing : Journal Quality List*) : la pratique se limite ici à un simple référencement à des objets constitués à partir d'une méthodologie qui n'a pas été examinée par les délégués scientifiques de l'AERES et soumise à la démarche de recherche de la qualité qui est un des principes de l'évaluation de l'Agence. D'autres disciplines, comme le droit, ont constitué leur propre liste, à partir de *critères empiriques de notoriété*, fixés par les autorités représentatives des disciplines juridiques (présidents des sections 01, 02, 03 du CNU, président de la section 36 du CoNRS, DSA en charge du droit au sein du CNRS, président du conseil national du droit, président de la conférence des doyens des facultés de droit, présidents d'université à dominante juridique, personnalités qualifiées). D'autres disciplines encore ont établi des listes *en fonction d'indicateurs*, sans que ceux-ci soient unifiés pour l'ensemble du secteur Sciences humaines et sociales.

Pour la philosophie, quatre critères ont été utilisés : 1) publication comportant principalement des articles de recherche en philosophie ; 2) présence effective d'un comité scientifique et d'un comité de lecture distincts et comportant des chercheurs de la communauté internationale ; 3) lecture des articles proposés à la revue en « double aveugle » (y compris pour les « dossiers » avec contributions commandées) ; 4) sélectivité. Pour les sciences de l'information et de la communication, ce ne sont pas moins de sept critères qui ont permis d'établir la liste : 1) existence d'un comité de rédaction ouvert (représentativité des membres, non concentration sur un seul lieu de recherche, appel à des collègues étrangers) ; 2) sélection rigoureuse des articles (relecture croisée, en aveugle, commentaires critiques pouvant aboutir à des demandes de réécriture ou des refus) ; 3) ouverture internationale (présence d'auteurs étrangers, résumé en langue étrangère) ; 4) origine institutionnelle de la rédaction (association scientifique, laboratoires et équipes de recherche reconnues, soutien d'une maison d'édition reconnue) ; 5) exigences internes de publication (régularité annuelle des parutions, nombre minimal d'articles parus, part limitée d'autopublication des membres de la rédaction, taille moyenne des articles conséquente) ; 6) liens institutionnels et scientifiques avérés avec la discipline (origine disciplinaire des auteurs, problématiques soulevées, contribution aux débats de la discipline, apports à la diffusion des savoirs par l'existence de notes de lecture) ; 7) indexation dans les grandes banques de données documentaires internationales (Francis, ISI, etc.).

À ces disparités de méthode s'ajoutent des différences notables dans la nature des listes obtenues. Certains secteurs disciplinaires (droit, sciences de l'éducation, sociologie-démographie, philosophie, arts, histoire-histoire de l'art-archéologie, etc.) ont limité leur travail à l'établissement d'un périmètre scientifique. D'autres disciplines (Psychologie, STAPS par exemple) non seulement ont établi un périmètre scientifique, mais se sont engagés de surcroît dans la voie du « ranking », en procédant à un classement, sans qu'aient été fixées pour autant des normes de classement communes aux disciplines des Sciences humaines et sociales. En psychologie, on dispose d'une échelle à 5 niveaux : modéré, assez bon, bon, très bon, excellent. En STAPS, on est passé d'une échelle à deux niveaux (A+ et A) à une reconfiguration qui ne prend en compte que les revues dites « de rang A » (*i.e.* prise en compte dans le décompte des « producteurs »), parmi lesquelles on distingue désormais trois niveaux (1, 2, 3), en fonction de la qualité scientifique et du rayonnement. Enfin, un dernier groupe de disciplines (la langue et la littérature française, la littérature comparée, les langues et littératures anciennes, les langues, littératures et civilisations

étrangères) ont refusé d'établir des listes en vue d'une évaluation bibliométrique, considérant que cette approche n'étant pas pertinente. (Ces disciplines sont en train d'évoluer : les lettres ont par exemple accepté le principe d'un périmètre scientifique permettant d'établir une liste de revues).

Dans ce contexte contrasté surgissent inévitablement des problèmes, comme des disparités d'indexation ou de classement liées en particulier à la pluridisciplinarité. Il peut arriver qu'une revue du fait de son extension pluridisciplinaire soit indexée dans un domaine, mais pas dans un autre ; ou qu'elle soit diversement classée selon les domaines de spécialité où elle a été retenue. Cela n'est pas sans incidence sur la cohérence de la liste globale des revues en Sciences humaines et sociales produite par l'Agence.

Le premier bilan

Après cinq années d'existence, l'AERES est à l'heure d'un premier bilan et s'oriente donc vers une reprise méthodique de la question des indicateurs. Alors qu'elle vient de changer de président, l'Agence, profitant de l'allègement de la prochaine vague C (du fait au passage à un rythme d'évaluation quinquennal), a fixé comme l'une de ses priorités la question des indicateurs en Sciences humaines et sociales. Cette question sera abordée en se fondant une fois encore sur quelques principes. Un travail sur les indicateurs a pour préalable la définition des *éléments de culture partagée*, de langage commun, en ayant le souci de distinguer l'essentiel de l'accessoire. S'agissant en particulier des Sciences humaines et sociales, il est important d'abord de *relativiser leur exceptionnalité* en la rapportant à des facteurs qui sont peut-être moins constitutifs (*i.e.* relevant d'un fait de « nature » ou d'une « essence ») qu'épistémologiques, historiques et institutionnels. La coexistence des Sciences humaines et sociales au sein de l'AERES avec les Sciences de la vie et de l'environnement et les Sciences et technologies apprend à ceux qui en font l'expérience qu'il y a des clivages au sein de ces grands domaines qui clivent aussi les Sciences humaines et sociales : recherche finalisée *vs* recherche fondamentale, recherche collective *vs* recherche individuelle, grosses structures de recherche *vs* petites structures, etc. En conséquence, il y a des similitudes entre certains domaines des Sciences humaines et sociales et certains domaines des Sciences et technologies et des Sciences de la vie et de l'environnement : la recherche individuelle est de première importance en mathématique comme en philosophie ou en lettres, par exemple ; en biologie, la tendance est aux petites équipes, comme en droit, etc.

Il faut ensuite *définir ce qui constitue les Sciences humaines et sociales comme telles* (la question du sens ? celle de la valeur ? la combinaison, variable selon les disciplines, de la modélisation et de l'herméneutique ? la typologie des productions ? les modalités de la recherche ? etc.), non pas négativement (par exclusion, retranchement, parti pris d'insularité) mais positivement (par affirmation du caractère propre). Il convient enfin de poser en principe qu'elles sont *commensurables* aux autres disciplines du savoir, qu'il existe donc des indicateurs pertinents pour objectiver leurs performances (et rendre compte de l'utilisation des crédits dont elles bénéficient), tout en tenant compte de leurs spécificités.

Un autre préalable à la réflexion consistera à tenir compte des méthodes et pratiques concurrentes dans le domaine des Sciences humaines et sociales : travailler en concertation avec les autres parties prenantes : DGRI, OST, CoNRS, DSA de l'INSHS, sections concernées du CNU, alliance ATHENA considérer les démarches d'évaluation et les indicateurs retenus par les institutions d'évaluation étrangères équivalentes à l'AERES (ERIH, HEFCE en Grande-Bretagne, NWD aux Pays-Bas, etc.).

Les perspectives

C'est dans ce contexte général qu'il faut situer l'intérêt de l'Agence d'évaluation de la recherche et de l'enseignement supérieur pour la bibliométrie, pour les listes de revue et pour les

bases de données indexant ces listes. Dans ce domaine, somme toute limité mais crucial, le travail sur les indicateurs programmé à l'Agence consistera à se doter *a posteriori* de critères communs en vue de la constitution de périmètres scientifiques pour les revues des différents domaines ou disciplines.

Plus précisément, il s'agira d'abord de poser une définition de ce que l'on entend par « revue scientifique » (au-delà de la simple présence d'un numéro ISSN). Il faudra ensuite définir les indicateurs pertinents permettant d'évaluer ce type de revues en Sciences humaines et sociales et, le cas échéant, de les classer. En l'état, embryonnaire, de la réflexion, voici un premier repérage d'indicateurs secondaires et les items afférents, qui restent à valider :

- *Mode de publication et de diffusion* :
 - Existence d'un diffuseur ;
 - Nombre d'exemplaires diffusés par numéro ;
 - Existence d'une publication en ligne ;
 - Modalités de cette publication (*open access*, péage) ;
- *Présentation matérielle de la publication* :
 - Présence de notes de bas de page dans les articles ;
 - Présence d'une bibliographie par article ;
 - Existence d'abstracts en langue anglaise ;
- *Politique éditoriale*
 - Existence d'une ligne éditoriale identifiable ;
 - Adéquation des contenus aux objectifs scientifiques de la publication ;
 - Diversité des auteurs d'articles (extension locale, nationale, internationale) ;
 - Diversité des éditeurs scientifiques/numéro (extension locale, nationale, internationale) ;
 - Taux d'acceptation ;
 - Existence d'un site web (en français seul ; bilingue : anglais/autre langue) présentant la revue et ses procédures de sélection.
- *Notoriété*
 - Présence au catalogue des bibliothèques régionales, nationales, internationales
 - Niveau de citation des articles dans les bases disponibles ;
 - Niveau de citation des éditeurs scientifiques dans les bases disponibles ;
 - Évaluation par les pairs.

Ce n'est qu'au moment où l'AERES aura établi et testé la batterie d'indicateurs permettant de dresser un état des lieux pour les revues en Sciences humaines et sociales (mais aussi pour les ouvrages individuels et collectifs, notamment les actes de colloque) qu'elle pourra reprendre et affiner les périmètres scientifiques qu'elle a dessinés pour chaque domaine disciplinaire, qu'elle pourra aussi avancer dans la voie d'un classement des revues (en affectant un coefficient à chaque indicateur et en définissant des paliers permettant de passer du niveau inférieur au niveau supérieur de l'échelle ?), qui est son objectif à terme. Deux hypothèses sont envisagées.

La première est la constitution d'un périmètre scientifique large (recensement le plus complet possible des revues d'un champ scientifique, nationales et internationale) et, à l'intérieur de ce périmètre, la distinction des revues « à étoile » (i.e. des revues ayant atteint, au vu des indicateurs retenus, un certain niveau de qualité). Quatre indicateurs, à première vue, sont considérés comme

prioritaires dans la liste dressée plus haut : l'existence d'une publication en ligne ; l'existence d'abstracts en langue anglaise ; la diversité des auteurs d'articles ; le taux d'acceptation. La seconde hypothèse porte sur la constitution d'un périmètre restreint d'emblée à des revues remplissant un ensemble de conditions d'accès à la liste de l'AERES (conditions toujours fixées en fonction des indicateurs déjà mentionnés).

Il ne s'agit là que de pistes à peine ébauchées qui seront affinées de septembre à décembre 2010 dans un séminaire hebdomadaire qui réunira les délégués scientifiques de l'AERES et des spécialistes de l'évaluation. La priorité est à la constitution, au sein de ce groupe de travail, d'*un référentiel* : un corpus d'indicateurs communs à toutes les disciplines des Sciences humaines et sociales. Ceux-ci pourront être différemment pourvus de coefficients, voire, pour certains, activés ou désactivés selon les disciplines, pour tenir compte des spécificités de chaque domaine, la question d'indicateurs fondamentaux, qu'il est impossible de neutraliser, étant évidemment posée.

Une fois ce corpus établi, les délégués scientifiques en Sciences humaines et sociales mèneront auprès de leurs communautés respectives une consultation, en s'appuyant sur les représentants du CNU et du CoNRS concernés, sur les DSA du CNRS, ainsi que sur des personnalités qualifiées. Cette consultation aura pour but de légitimer la démarche et de s'assurer de son acceptabilité, l'objectif visé, dont on espère qu'il n'est pas trop chimérique, étant celui d'une adhésion largement majoritaire. L'AERES, on le voit, est engagée dans une démarche de rationalisation méthodique de ses indicateurs, qu'elle souhaite fonder sur une exigence de qualité tenant compte de la diversité des pratiques en Sciences humaines et sociales et permettant, en même temps, de caractériser avec cohérence, rigueur et finesse les productions issues de ces disciplines.

Index des auteurs

Bosquet Clément.....	3
Combes Pierre-philippe.....	3
Dassa Michèle.....	19, 87
Dubucs Jacques.....	104
Filliatreau Ghislaine.....	131
Glaudes Pierre.....	143
Kosmopoulos Christine.....	83, 87
Maurel Marie-claude.....	31
Mirdal Gretty.....	37
Pétard Jean-pierre.....	10
Ramjoué Celina.....	79
Ruggiu François-joseph.....	1
Sidéra Isabelle.....	19

