

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 *Cybergeo* 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://cybergeo.revues.org/22492>

⁵ <http://cybergeo.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'Évaluation 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.

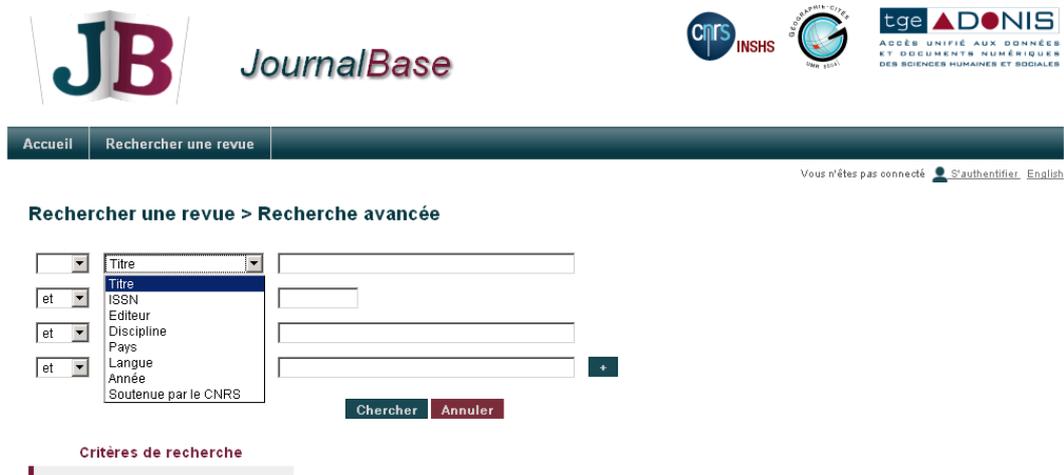


Figure 1 : Recherche avancée

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



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).

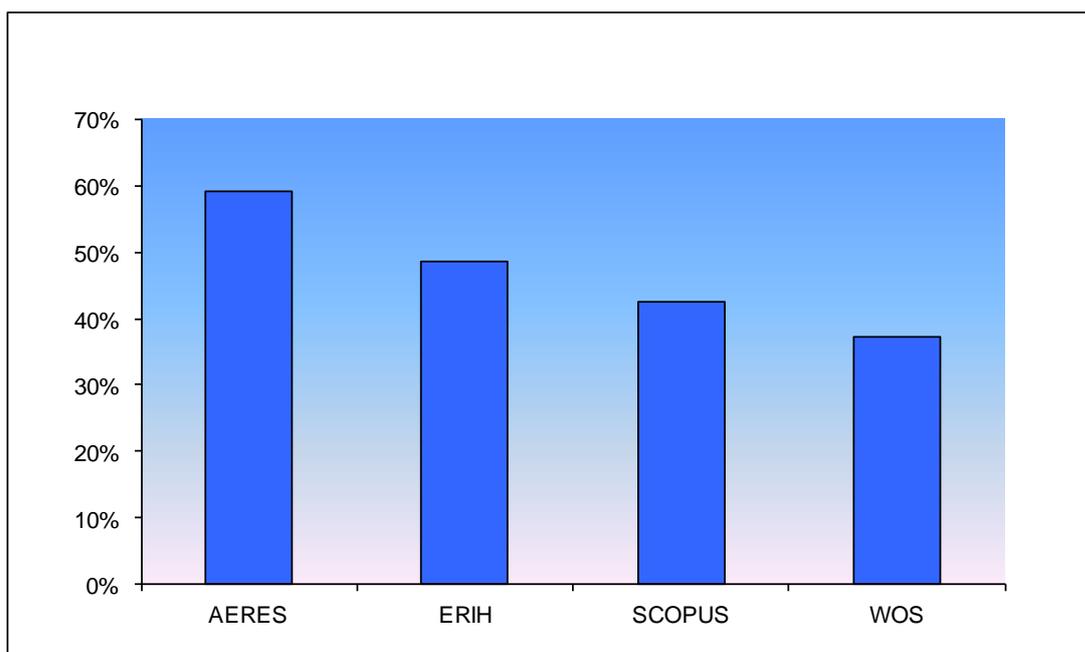


Figure 4 : % de revues sur le nombre total de revues par source étudiée

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://cybergeog.revues.org/22864>

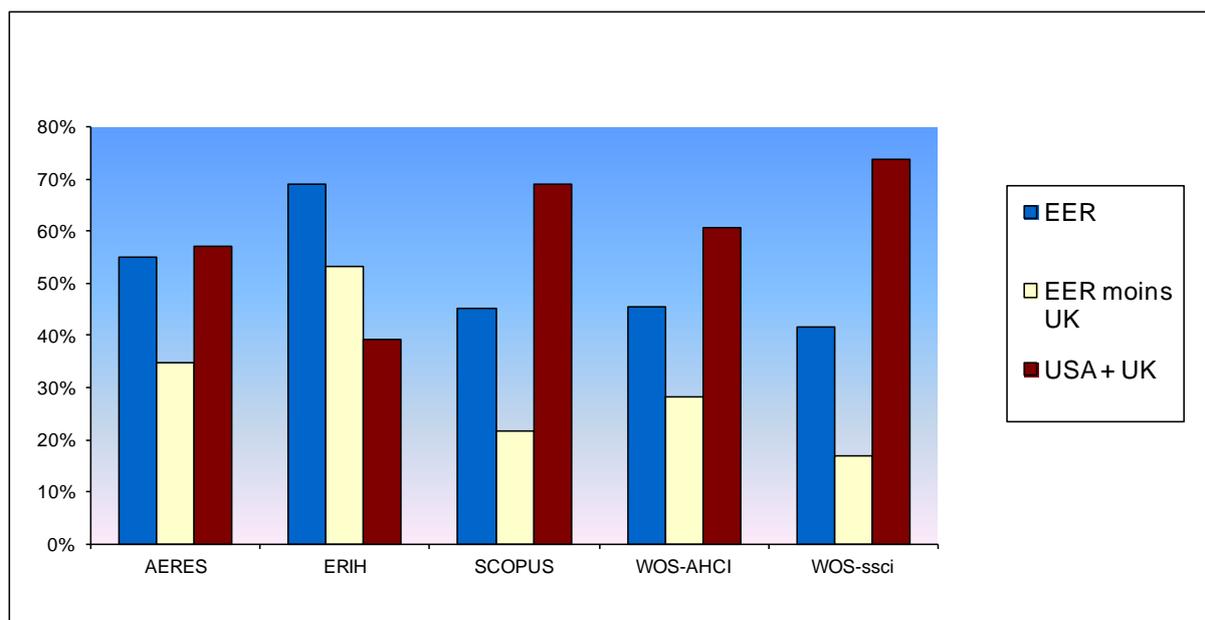


Figure 5 : Répartition des revues par groupe de pays d'édition

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 vue 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.

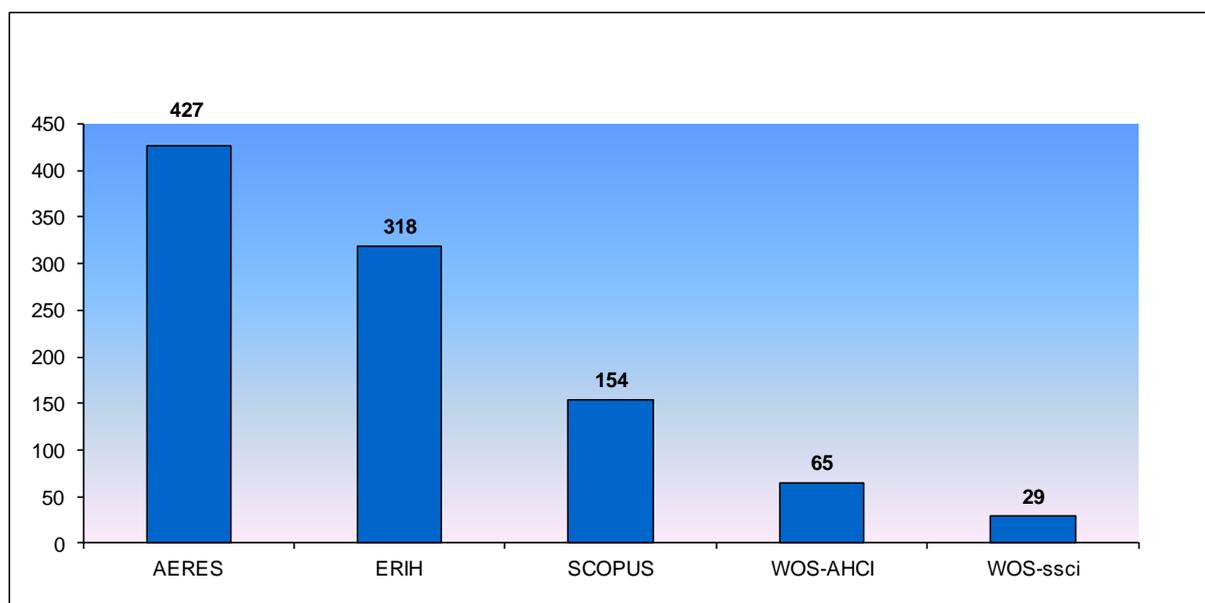


Figure 6 : nombre de revues françaises par source étudiée

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/>

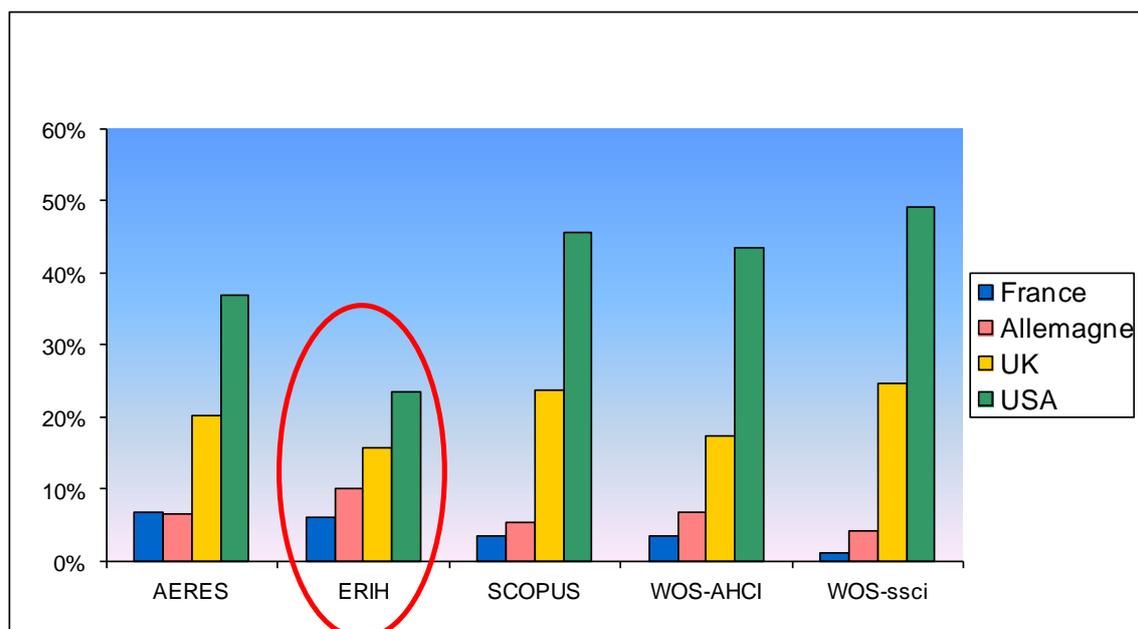


Figure 7 : Nombre de revues par pays et par source

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.

Revue 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

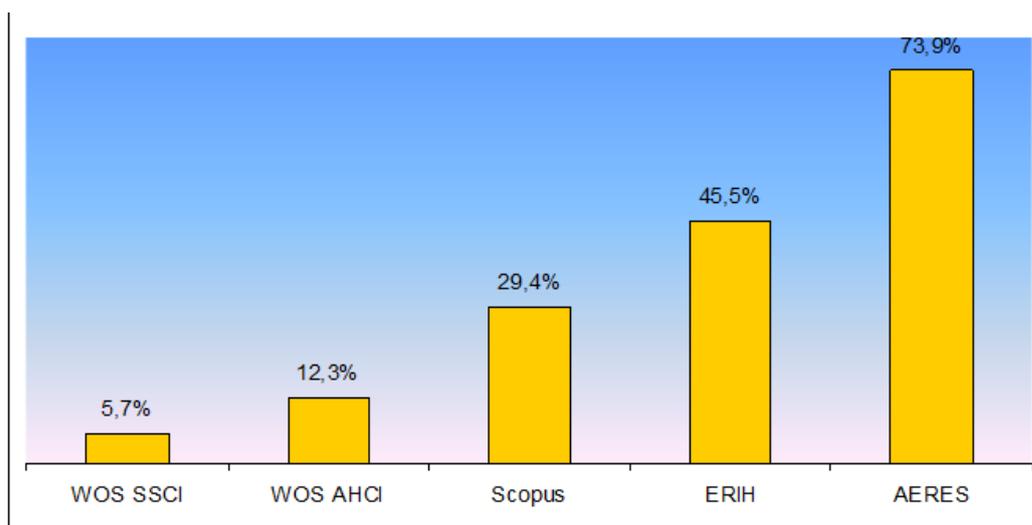


Figure 8 : Proportion de revues soutenues par le CNRS dans Journalbase

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.

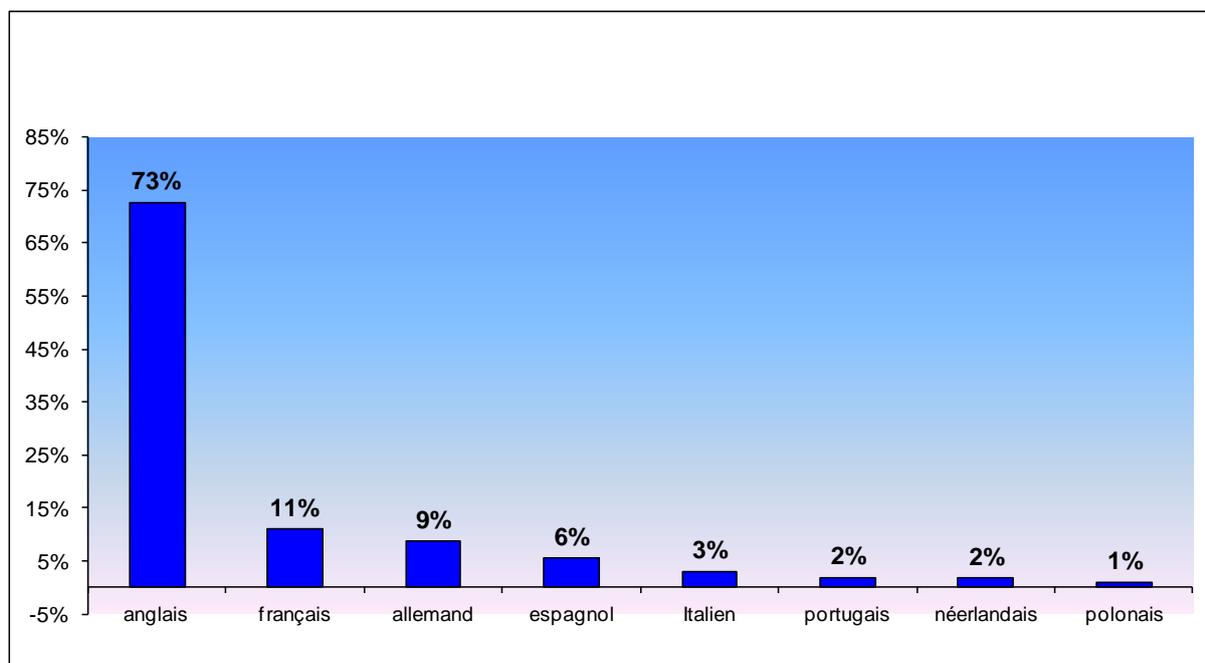


Figure 11 : Répartition des revues par langue

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.

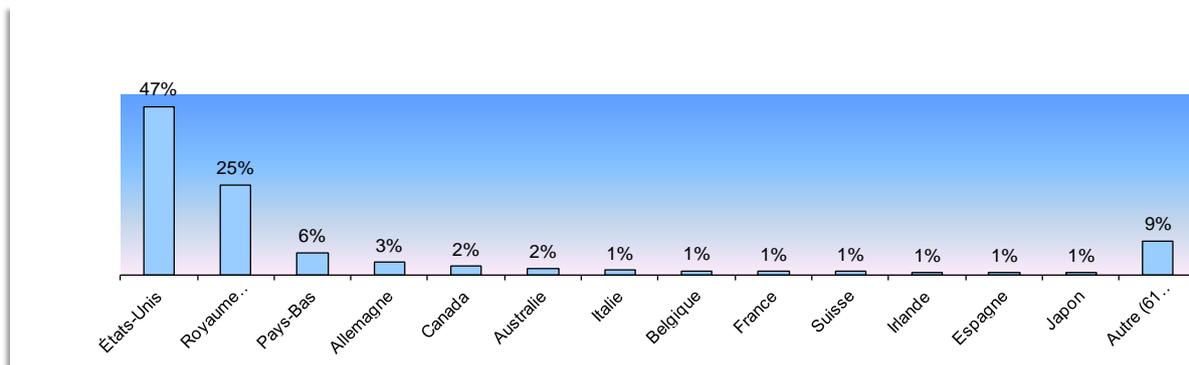


Figure 12 : Répartition des 7599 revues de langue anglaise par pays d'édition

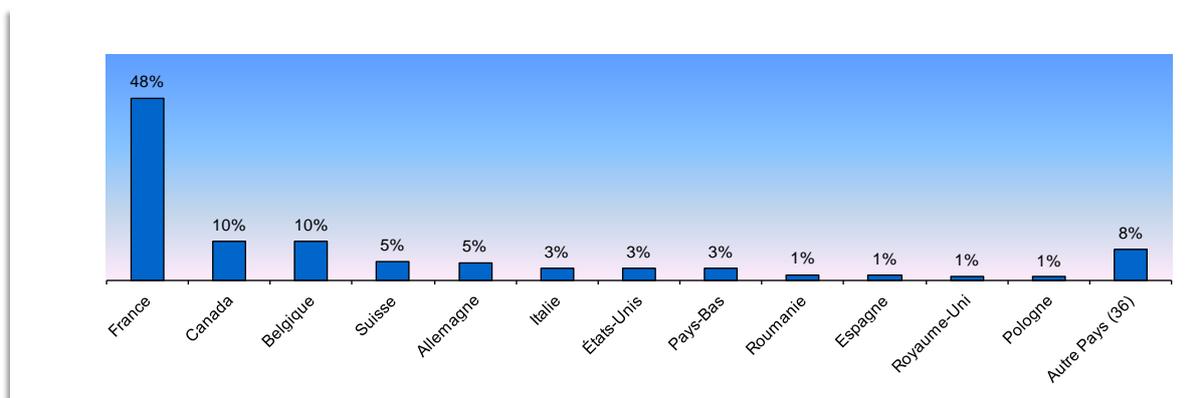


Figure 13 : Répartition des 1153 revues de langue française par pays d'édition

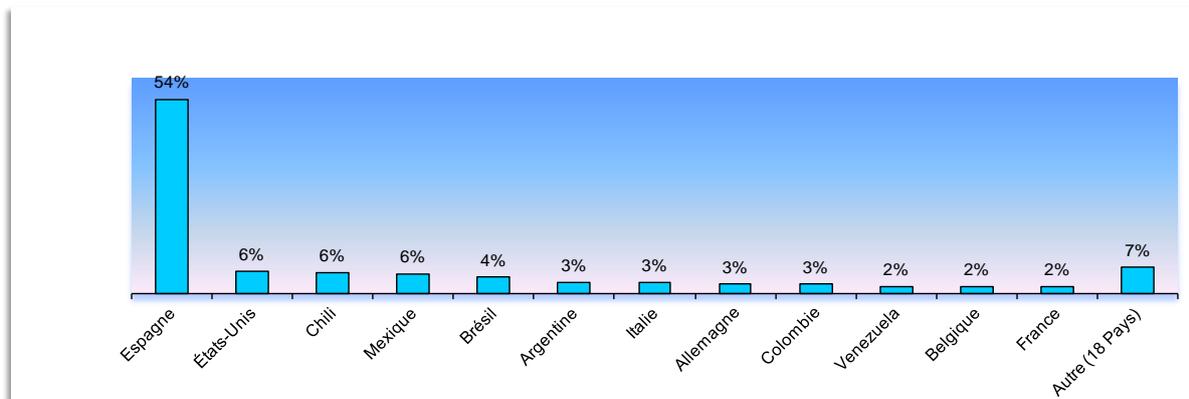


Figure 14 : Répartition des 578 revues de langue espagnole par pays d'édition

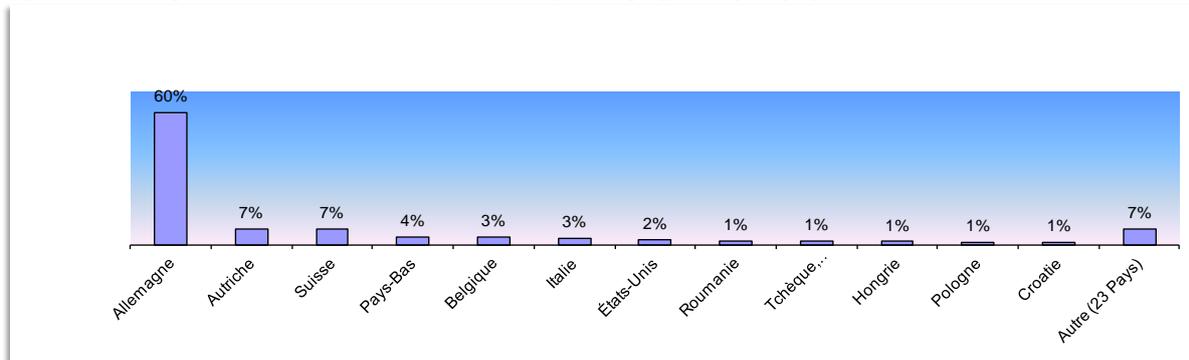


Figure 15 : Répartition des 910 revues de langue allemande par pays d'édition

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.