
SCOPUS ET WORLDCAT

Regards croisés sur deux outils bibliographiques

Philippe MOTTET

Documentaliste, Université de Liège (ULg) - Bibliothèque des Sciences et Techniques

Ninfa GRECO

Directrice, Université de Liège (ULg) - Bibliothèque des Sciences et Techniques

▪ Les outils de recherche bibliographique, imprimés puis électroniques, ont longtemps été le domaine réservé de sociétés commerciales qui proposaient des produits clé sur porte, multidisciplinaires ou spécialisés. Ils restent incontournables pour la plupart des institutions scientifiques par la qualité de leur contenu, la précision de leurs systèmes d'interrogation et la facilité de traiter et de sauvegarder les résultats des recherches réalisées. Ce monopole a été remis en question par l'accessibilité, toujours plus grande, et gratuite, de l'information sur Internet. Parmi les nombreux outils bibliographiques actuellement disponibles, nous avons choisi de comparer un système commercial traditionnel, utilisé dans les universités de la Communauté française de Belgique : *Scopus*, et une solution issue du monde des bibliothèques, consultable gratuitement, du moins partiellement, sur Internet : *WorldCat*. La comparaison des deux produits a donné lieu à des résultats parfois inattendus.

▪ Of het nu in hun eerdere gedrukte vorm was of in de latere elektronische versies, bibliografische zoekmiddelen waren lang het exclusieve terrein van privébedrijven die kant-en-klare, multidisciplinaire of gespecialiseerde producten afleverden. Deze instrumenten blijven onontbeerlijk voor de meeste wetenschappelijke instellingen, vanwege de inhoudelijke kwaliteit, de precisie van de zoeksystemen en het gemak waarmee de zoekresultaten behandeld en bewaard kunnen worden. Dat monopolie komt op losse schroeven te staan door de steeds beter toegankelijke gratis informatie op het Internet. Te midden van de vele bibliografische instrumenten die momenteel beschikbaar zijn, hebben wij ervoor gekozen om een traditioneel commercieel systeem, *Scopus*, in gebruik in de Universiteiten van de Franse Gemeenschap in België, te vergelijken met een oplossing die afkomstig is uit de bibliotheekwereld en minstens deels gratis te raadplegen is op het Internet: *WorldCat*. De vergelijking van beide producten leverde soms onverwachte resultaten op.

La *Googlemania* est planétaire et atteint aussi le monde des bibliothèques, phénomène amplifié par le développement d'outils tels que *Google Scholar* et *Google Books*. Soucieux de rester en phase avec les usagers, les bibliothécaires choisissent de plus en plus de surfer sur la vague, notamment en optant pour des solutions intégrées d'accès aux ressources bibliographiques, de type "Web-Scale Discovery Tool" tels que *Primo*, *EDS* ou *Summon*¹. Ces solutions permettent aux usagers d'interroger globalement les contenus de bibliothèque sur un mode de recherche de type "Google". Les professionnels de l'information apprécient actuellement à des degrés divers cette tendance. Certains considèrent que les outils traditionnels, soit les bases de données bibliographiques et les catalogues de bibliothèques (OPAC), perdent de leur visibilité derrière ces interfaces intuitives et seraient sous-exploités.

Mais à l'heure actuelle, qu'en est-il justement du clivage entre les bases de données bibliographiques et les catalogues, les premières servant par définition à révéler l'existence des documents et les seconds à localiser ceux-ci ? Nous avons eu la curiosité d'effectuer en parallèle une recherche d'articles scientifiques contenant dans leur titre un terme très spécifique, "euplotes"², d'une part dans *Scopus*, la base de données bibliographique multidisciplinaire utilisée par les universités de la

Communauté française de Belgique³ et d'autre part dans *WorldCat*, réputé le plus grand catalogue OPAC du monde (son nom est la contraction de "World Catalog") et qui permet d'effectuer des recherches dans les collections de milliers de bibliothèques⁴. De façon inattendue, non seulement nous avons obtenu des résultats dans *WorldCat* dont la vocation en tant que catalogue n'est pas a priori de dépouiller des périodiques scientifiques, mais de plus le nombre de réponses dans *Worldcat* était nettement supérieur à celui de *Scopus*⁵.

Intrigués, nous avons décidé d'approfondir la comparaison entre *Scopus* et *WorldCat* quant à la fonction très spécifique d' "outil de recherche d'articles scientifiques".

Présentation des outils

*Scopus*⁶ est une base de données multidisciplinaire lancée par Elsevier en 2004. Concurrente de *Web of Science*, elle dépouille 19.500 publications dont 18.500 journaux "peer-reviewed", répertoriant à ce jour 49 millions de références⁷, dont les plus anciens datent du 19^e siècle.

Outre sa fonction de recherche documentaire, *Scopus*, tout comme le produit concurrent *Web of*

Science, accorde une large place aux outils bibliométriques qui donnent des indications sur la valeur des revues scientifiques ("Scimago" chez *Scopus*, "facteur d'impact" pour le *Web of Science*) et qui peuvent être exploités dans le cadre de l'évaluation des curriculum vitae des chercheurs (h-index,...). *Scopus* offre également diverses fonctions d'analyse, de sauvegarde et d'exportation des références, options appréciées par les usagers issus des institutions académiques et scientifiques.

WorldCat a été mis en œuvre en 1971 par l'OCLC⁸ sous le nom de *OLUC (OCLC Online Union Catalog)*. Rassemblant initialement les catalogues des bibliothèques de l'Ohio, le réseau s'est progressivement étendu aux États-Unis puis au reste du monde. Au départ, il s'agissait d'une association de bibliothèques coopérant sans but lucratif, notamment à la production d'un catalogue mondial. Depuis quelques années, avec pour but affiché un meilleur partage de la connaissance, l'OCLC déploie une politique plus commerciale. L'OCLC produit ainsi sa propre solution de type "Web-Scale Discovery Tool", *WorldCat Local*, dont les contenus ne cessent de croître au fur et à mesure des accords pris avec les fournisseurs d'information : agrégateurs, éditeurs, *Google Books*, *OAlster*...⁹. *WorldCat Local* annonce ainsi plus d'un milliard de notices de formats divers : 91.495 journaux, plus de 788 millions d'articles, plus de 227 millions de livres, mais également plus de 14 millions d'e-books, des thèses, des cartes, des enregistrements sonores et visuels, des comptes rendus de conférences, etc.

Dans cette étude, nous avons utilisé la plate-forme *WorldCat.org*¹⁰ qui permet d'interroger gratuitement la base de données *WorldCat*. Déterminer à quoi donne accès la plate-forme gratuite n'est pas évident. La page d'accueil *Worldcat.org* annonce une connexion à plus de 10.000 bibliothèques à travers le monde et l'accès à 1,5 milliards d'ouvrages variés (livres, DVD, CD, articles...). La consultation de "WorldCat.org frequently asked questions" (FAQ)¹¹ nous indique que *WorldCat.org* permet à l'utilisateur d'interroger l'entièreté du contenu de la base de données *WorldCat*, avec quelques restrictions cependant : les résultats de la recherche n'afficheront pas nécessairement tous les items

et l'utilisateur n'aura accès qu'aux collections des bibliothèques qui souscrivent à *FirstSearch*, un produit commercial de l'OCLC. En conséquence, les bibliothèques sont invitées à souscrire à *FirstSearch* dont un des objectifs est d'augmenter la visibilité des collections des bibliothèques sur le web¹². Par ailleurs, des références d'articles issues de *ArticleFirst*, *British Library Inside Serials*, *ERIC*, *MEDLINE*, *JSTOR*, *PapersFirst*, *Proceedings-First* ou *Elsevier*¹³ seraient aussi disponibles via *WorldCat.org*.

Dans cet article, nous avons donc sciemment choisi de comparer *Scopus*, une base de données commerciale "payante", avec la version de *WorldCat* accessible gratuitement sur Internet (*WorldCat.org*).

Comparaison

Plus précisément, notre comparaison entre *WorldCat.org* et *Scopus* porte donc sur la recherche d'articles scientifiques. Mais avant cela, une rapide analyse des écrans de recherche et de résultats, de leurs fonctionnalités annexes et de leurs facilités d'utilisation, nous renseignera sur la qualité des informations que nous pouvons espérer obtenir lors de recherches bibliographiques.

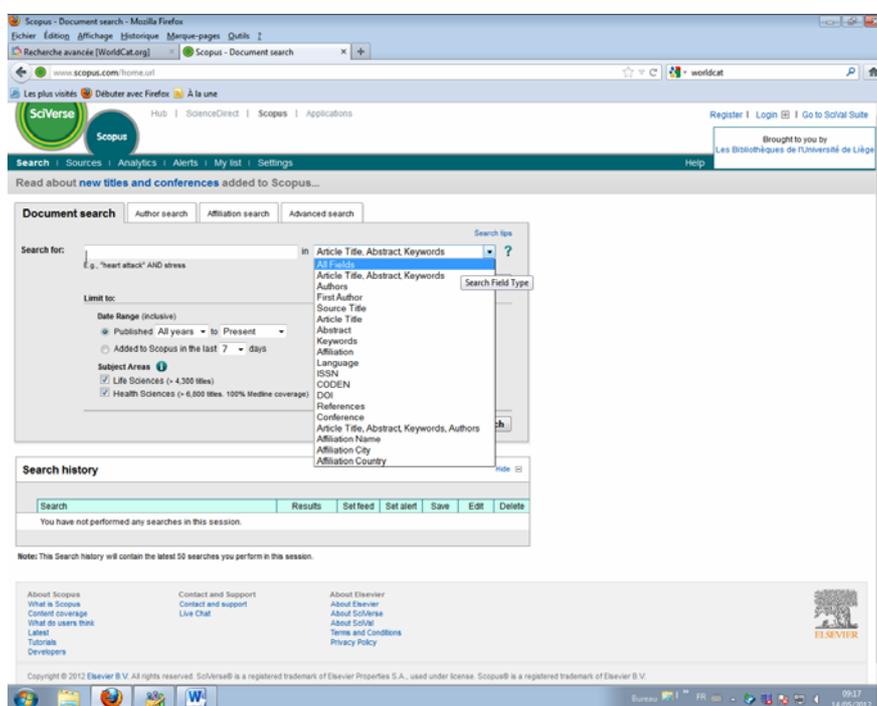


Fig. 1 : Page d'accueil de Scopus.

Fonctionnalités et facilités d'utilisation

La page d'accueil de *Scopus*¹⁴ (Fig. 1) propose différentes options de recherche au moyen d'onglets. En cas de recherche de documents (Document search), un menu déroulant liste les champs interrogeables, très spécifiques et adaptés à la recherche de documents (article, synthèse, conférence, éditorial...) parus dans des périodiques scientifiques. Par ailleurs, d'autres critères (date, type de document, disciplines) peuvent être préalablement sélectionnés pour limiter la requête. Nous sommes ici dans la configuration traditionnelle des bases de données bibliographiques en ligne, qui offre à

par date ou dans un vaste choix de types de documents (document d'archives, article, livre audio, livre, fichier d'ordinateur, périodique, carte géographique, musique, partition, site web, enregistrement sonore, jeu, matériel visuel...). Il est également possible de pré-sélectionner le contenu des documents (fiction, documentaire, ...) et le public visé (jeune ou adulte), fonctions évidemment inutiles dans *Scopus*.

La page d'accueil de *Scopus* propose des liens vers des tutoriels qui permettent d'accéder à une recherche fine, par exemple par l'usage des opérateurs booléens ou d'une syntaxe spécifique. L'équivalent ne se retrouve pas sur la page de recherche avancée de *WorldCat.org*.

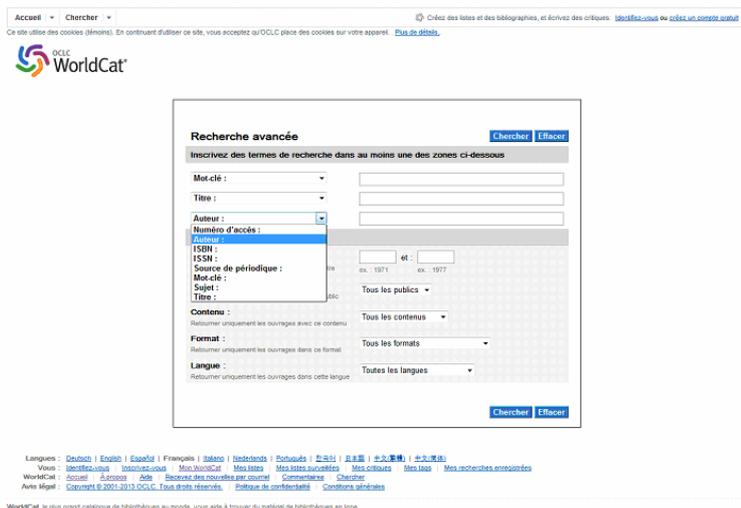


Fig. 2 : Page de recherche avancée de *WorldCat.org*.

l'utilisateur la possibilité de limiter la portée de sa recherche avant même de lancer celle-ci.

*WorldCat.org*¹⁵ propose par défaut un menu de recherche de type "Google", qui comporte un seul champ de recherche interrogeant l'ensemble de la base de données. Il est bien sûr facile de sauter rapidement vers la recherche avancée (Fig. 2), mais le choix par défaut d'un menu simple, comme ceux qu'on trouve dans les outils de découverte de nouvelle génération cités plus haut, pourrait être significatif d'une volonté de placer l'utilisateur dans un environnement qui lui est désormais familier. En témoignent les petites applications et widgets proposés sur la page d'accueil ainsi que l'accent mis sur les possibilités de partage des listes de résultats et des "impressions" des utilisateurs.

La recherche avancée offre la possibilité de sélectionner des champs de recherche dans un menu déroulant, mais les champs sélectionnables sont plus génériques que dans *Scopus* et adaptés à la diversité des documents répertoriés par *WorldCat*. *WorldCat.org* permet également de limiter préalablement la portée de la recherche

Les possibilités offertes par les deux systèmes en mode "Document search" pour l'un, "Recherche avancée" pour l'autre, semblent correspondre aux objectifs initiaux de chacun d'eux : la recherche documentaire spécialisée pour un public universitaire chez *Scopus* et la localisation de documents variés pour un public élargi chez *WorldCat.org*.

C'est en ayant ces orientations à l'esprit que nous avons effectué notre recherche initiale sur les "euplotes" dans les deux systèmes et obtenu des résultats que nous n'attendions pas, en termes de nombre d'articles scientifiques.

Contenu des bases de données

Pour comparer le contenu scientifique des deux bases de données en termes d'articles scientifiques, nous avons sélectionné 50 titres de journaux¹⁶ et comparé leur couverture dans *Scopus* et *WorldCat.org*.

Après divers essais, nous avons décidé d'effectuer la recherche sur l'ISSN qui est l'élément le moins équivoque dans les deux bases de données (risque de bruit de fond lorsque la recherche est effectuée sur le titre du périodique !). Nous avons exclu l'année 2013 de notre recherche afin de stabiliser dans le temps, autant que possible, le nombre de réponses obtenues. Nous avons utilisé l'onglet "Source" dans *Scopus* et la "Recherche avancée" dans *WorldCat.org*.

Nous avons relevé¹⁷ pour chaque base le nombre de réponses correspondant à des articles scientifiques et pris soin de ne pas tenir compte des titres des revues elles-mêmes que *WorldCat.org* comptabilise parfois à plusieurs reprises. Nous avons également noté les années dépouillées et finalement obtenu le tableau 1.

Tab. 1 : Comparaison du contenu de Scopus et WorldCat.org

Titre du journal	ISSN	Scopus		WorldCat.org	
		Réponses	Années couvertes	Réponses	Années couvertes
1 <i>Revue médicale de Liège</i>	0370-629X	7.670	1947-2012	9.834	1947-2012
2 <i>Annales de médecine vétérinaire</i>	0003-4118	490	1947-48; 1996-2011	517	1995-2011
3 <i>Astronomy and astrophysics</i>	0004-6361	30.118	1984; 1986-1988; 1990-1993; 1995-2012	52.345	1984; 1986-1988; 1990-2012
4 <i>Osteoporosis International</i>	0937-941X	3.923	1990-2012	6584	1990-2012
5 <i>Biotechnologie, Agronomie, Société et Environnement (=BASE)</i>	1370-6233	260	2009-2012	598	2002-2012
6 <i>Macromolécules</i>	0024-9297	37.205	1968-2012	31.359	1973-1978; 1989-2012
7 <i>Moyen Âge (Le)</i>	0027-2841	104	2001-2012	1.146	1974; 1978; 1997-1998; 2000; 2003; 2005-2012
8 <i>Journal of biological chemistry</i>	0021-9258	162.184	1945-2012	189.684	1945-2012
9 <i>Revue de jurisprudence de Liège, Mons et Bruxelles</i>	0774-8108	0	-	0	-
10 <i>Astrophysical Journal</i>	0004-637X	4.597	1984-2001; 2006-2007; 2009-2012	74.594	1895; 1900; 1966-1967; 1978; 1980; 1984-2012
11 <i>Acta chirurgica Belgica</i>	0001-5458	5.209	1946-2012	7.980	1946-2012
12 <i>Ciel (Le)</i>	0771-3010	0	-	0	-
13 <i>Acta clinica Belgica</i>	0001-5512	3.385	1946-2012	4.907	1946-2012
14 <i>Bulletin de la société géographique de Liège</i>	0770-7576	0 ¹⁸	-	0 ¹⁹	-
15 <i>Ciel et terre</i>	0009-6709	0 ²⁰	-	339	1887; 1907; 1912-1942; 1944; 1959; 1954-1978; 1998; 2002-2011
16 <i>Organohalogen compounds</i>	1026-4892	0	-	3.756	1997-2003
17 <i>Journal de pharmacie de Belgique</i>	0047-2166	2.670	1945-2012	2.776	1945-2012
18 <i>Biochemical Journal</i>	0264-6021	44.386	1945-2012	53.054	1909; 1912-1916; 1919-2012
19 <i>Monthly notices of the Royal Astronomical Society</i>	0035-8711	22.300	1987; 1989; 1996-2012	29.588	1987; 1989; 1992-2012
20 <i>Acta gastro-enterologica Belgica</i>	0001-5644	3.577	1946-2012	4.929	1946-2012
21 <i>Revue de la Faculté de droit de l'Université de Liège</i>	1780-5511	0	-	0 ²¹	-
22 <i>Veterinary Journal</i>	1090-0233	2.837	1997-2012	4.274	1994-2012 ²²
23 <i>Journal of dairy science</i>	0022-0302	14.346	1965-2012	16.942	1965-2012
24 <i>Bulletin de la Société Royale des Sciences de Liège</i>	0037-9565	325	1988; 1996-2012	362	1995-2005
25 <i>Culture, le magazine culturel de l'Université de Liège</i>	2032-7900	0	-	0 ²³	-
26 <i>Osteoarthritis and cartilage</i>	1063-4584	2.547	1993-2012	8.611	1993-2012
27 <i>Journal des tribunaux</i>	0021-812X	0	-	0 ²⁴	-
28 <i>Annals of the rheumatic diseases</i>	0003-4967	11.953	1945-1948; 1950-2012	13.793	1939-2012
29 <i>Theriogenology</i>	0093-691X	9.160	1974-2012	16.586	1974-2012
30 <i>PLoS one</i>	1932-6203 ²⁵	53.883	2006-2012	52.518	2006-2012
31 <i>Geophysical research abstracts</i>	1029-7006	0	-	0 ²⁶	-
32 <i>Bulletin des recherches agronomiques de Gembloux</i>	0435-2033	14	1971; 1973; 1978-1980; 1987	388	1995; 2002-2011
33 <i>Arthritis and rheumatism</i>	0004-3591	15.485	1958-2012	21.682	1958-2012
34 <i>Archives internationales de physiologie et de biochimie</i>	0003-9799	4.944	1949; 1954-1990	4.868	1949; 1954-1990; 1994 ²⁷
35 <i>Revue du rhumatisme</i>	1169-8330	3.924	1961; 1993-2012	9.287	1993-1994; 2000-2012
36 <i>Veterinary record : journal of the British veterinary association</i>	0042-4900	29.543	1945-1951; 1961; 1965-2012	36.917	1945-1951; 1961; 1965-2012
37 <i>Cephalalgia</i>	0333-1024	4.094	1981-2012	5.490	1981-2012
38 <i>Médecine et hygiène</i>	0025-6749	8.793	1945-1952; 1961-1962; 1965-2004	13.234	1945-1952; 1961-1962; 1965-1989; 1994-2004; 2007-2010
39 <i>Forêt wallonne</i>	1372-8903	0	-	0 ²⁸	-
40 <i>Acta anaesthesiologica Belgica</i>	0001-5164	2.089	1950-1951; 1960-2012	2.800	1950-1951; 1959-1961; 1963-2012
41 <i>Journal of clinical endocrinology and metabolism</i>	0021-972X	28.826	1945-2012	35.489	1945-2012
42 <i>Proceedings of the National Academy of Sciences of the United States</i>	0027-8424	103.441	1945-1951; 1961-2012	143.199	1915-2012
43 <i>Cahiers d'éthologie</i>	0778-7103	0	-	0 ²⁹	-
44 <i>Physical review letters</i>	0031-9007	104.411	1958-2012	109.805	1920; 1961; 1964; 1966; 1968-1971; 1975; 1977; 1980; 1982-1985; 1988-2012
45 <i>Astronomy and astrophysics. Supplement series</i>	0365-0138	1.712	1985; 1996-2000	21.252 ³⁰	1985; 1995-2010
46 <i>Nature</i>	0028-0836	301.371	1869-2012	137.218	1945-2012
47 <i>Lecture notes in computer science</i>	0302-9743	215.346	1981-1984; 1986; 1998-2012	352.160	1994-2012
48 <i>Literaturkritik.de</i>	1437-9317 ³¹	0	-	0 ³²	-
49 <i>Pflügers archiv</i>	0031-6768	11.948	1950-1959; 1968-2012	14.330	1950-1959; 1968-2012
50 <i>Geophysical research letters</i>	0094-8276	25.123	1977; 1979-2012	45.459	1982; 1984; 1987; 1989-2012

La première constatation est que le nombre de périodiques dépouillés dans *WorldCat.org* est aussi important et même supérieur au nombre de journaux dépouillés dans *Scopus*. Outre les périodiques communs aux deux bases, *WorldCat.org* dépouille également *Ciel et terre* (données fournies par *ArticleFirst*) et *Organohalogen compounds* (données fournies par *British Library Serials*).

D'une façon générale, les revues locales ne sont dépouillées par aucune des deux bases, même s'il peut exister des traces de leur existence : *Revue de jurisprudence de Liège, Mons et Bruxelles, Bulletin de la société géographique de Liège, Revue de la Faculté de droit de l'Université de Liège, Culture, le magazine culturel de l'Université de Liège, Journal des tribunaux, Forêt wallonne* et *Cahiers d'Éthologie*.

Au titre par titre, le nombre de réponses affichées est toujours supérieur dans *WorldCat.org*, à l'exception de trois cas (tableau 1 : lignes 30, 34 et 46) dont celui notable de *Nature* qui est dépouillé dans *Scopus* depuis le premier volume paru en 1869. Cependant, un examen plus attentif des réponses obtenues dans *WorldCat.org* a révélé que de nombreux articles apparaissaient plusieurs fois du fait de leur indexation parallèle par plusieurs participants à la base de données. Un nombre de résultats beaucoup plus élevé dans *WorldCat.org* n'est donc pas forcément significatif en comparaison de celui obtenu dans *Scopus*.

De fait, en première approche, la comparaison des états de collections (tableau 1) et du nombre de résultats obtenus année par année dans chacune des deux bases (détails non communiqués) semble indiquer que le contenu réel des deux bases serait comparable pour la moitié des périodiques (tableau 1 : lignes 1, 4, 8, 9, 11, 12, 13, 14, 17, 20, 21, 23, 25, 26, 27, 29, 30, 31, 33, 34, 36, 37, 39, 41, 48 et 49). Pour ces titres, les états de collection sont identiques tandis que nous avons relevé dans *WorldCat.org* des doublons d'articles : dans *Osteoporosis international* des doublons déposés à la fois par *ArticleFirst* et *Elsevier* ou par *ArticleFirst* et *MEDLINE/PubMed* ; dans *Osteoarthritis and cartilage*, des doublons déposés à la fois par *Elsevier* et *ScienceDirect* ; dans *Theriogenology*, des doublons déposés par *MEDLINE/PubMed* et *ScienceDirect* ou par *MEDLINE/PubMed* et *British Library Serials*...

Une analyse plus fine serait nécessaire pour déterminer si, pour un état de collection donné, les doublons apparaissant dans *WorldCat.org* peuvent expliquer en tout ou en partie le nombre plus élevé de réponses. Malheureusement, l'interface *WorldCat.org* ne permet pas de réaliser

une telle analyse dans un temps raisonnable. Nous n'avons pas trouvé de possibilité pour regrouper ou isoler les références identiques. De plus, le traitement des références et leur affichage est limité à 10 références à la fois, ce qui rend difficile la sauvegarde ou l'export de l'ensemble des notices bibliographiques dans une perspective d'analyse et de dédoublement a posteriori, via des programmes de gestion de références bibliographiques par exemple (*EndNote, Reference Manager, RefWorks*, etc.).

Par ailleurs, le filtre des résultats par facettes³³, qui est utilisé dans les deux systèmes et qui permet le choix de nouveaux critères pour affiner une recherche, est également très pauvre chez *WorldCat.org* par rapport à *Scopus*. Il semblerait que la version commerciale *FirstSearch* offre plus de possibilités de filtres, pour un usage professionnel, par exemple sur base du fournisseur de données³⁴.

Pour certains titres de périodiques, les données fournies par *British Library Serials, ArticleFirst, MEDLINE/PubMed, Life Sciences, Health & General Sciences* ou *Institute of Physics eJournals and Archive* (etc.) peuvent parfois compléter utilement les collections de *WorldCat.org* par rapport à *Scopus* (ex : *Astronomy and astrophysics*, 1994 ; *Biotechnologie, Agronomie, Société et Environnement*, 2002, 2003, 2004, 2005, 2006, 2007 et 2008 ; *Astrophysical Journal*, 2002-2005, 2008 ; *Biochemical Journal*, 1912-1916 et 1919-1944 ; *Proceedings of the National Academy of Sciences of the United States*, 1915-1944 et 1952 à 1960). À l'inverse, *Scopus* peut également présenter des états de collection plus complets qui ne sont pas nécessairement reflétés par le nombre de résultats affichés (ex : *Nature, Macromolecules, Revue du rhumatisme, Physical review letters*). Sous réserve de ce que pourrait révéler une étude plus détaillée, les états de collection de *WorldCat.org* semblent globalement plus complets que ceux de *Scopus*.

Vu les limites de l'approche par titres de périodiques, nous avons voulu poursuivre notre analyse par des recherches plus spécifiques, qui concernent des titres de périodiques associés à des années de publication, des titres d'articles, des noms d'auteurs et des mots-clés. Notre sélection de tests a été guidée par les centres d'intérêt de la Bibliothèque des Sciences de l'Université de Liège ou par les chercheurs de cette institution. Nos observations doivent donc être comprises dans leur contexte même si d'autres essais pourraient conduire à des observations similaires.

Recherches sur une ou quelques années précises d'un titre de périodique³⁵**Scopus**

ISSN 0021-8790 (journal of animal ecology) and Year = 2010 : 144 réponses dont il faut retirer 1 note d'éditeur et 1 erratum → **142 réponses**

WorldCat.org

ISSN 0021-8790 (journal of animal ecology) and Year = 2010 : 205 réponses auxquelles on retire 63 doublons, notes d'éditeur, erratum, pages liminaires et appendices → **142 réponses**

Aucun problème : la couverture est identique dans les deux bases de données.

Scopus

ISSN 0736-6299 (Solvent extraction and ion exchange) and Year 2009 → **40 réponses**

WorldCat.org

ISSN 0736-6299 (Solvent extraction and ion exchange) and Year = 2009 : 96 réponses auxquelles on retire 55 doublons = **41 réponses**

De nouveau, pas d'écart significatif entre les deux bases de données qui dépouillent l'ensemble de la revue. La réponse supplémentaire dans *WorldCat.org* est l'annonce d'une récompense attribuée à un professeur spécialisé dans le domaine couvert par le journal. Son absence chez *Scopus* n'a donc rien d'étonnant.

Autre exemple : une recherche par ISSN de la revue *Cataloging & Classification Quarterly* (ISSN 0163-9374) donne les résultats suivants pour les années 2005 à 2012 :

Scopus	WorldCat.org
2005 : 9 références	2005 : 126 références
2006 : 60	2006 : 140
2007 : 49	2007 : 105
2008 : 36	2008 : 103
2009 : 47	2009 : 144
2010 : 53	2010 : 147
2011 : 38	2011 : 105
2012 : 47	2012 : 90

Ici, l'écart entre les deux systèmes est assez étonnant mais se réduit quand on élimine les doublons (après dédoublonnage, le nombre de références obtenues dans *WorldCat.org* passe de 126 à 72 en 2005 ; de 140 à 85 en 2006 ; de 105 à 54 en 2007, etc.). Pour ce journal, *Scopus*

n'a notamment pas dépouillé le fascicule n°1 de 2012 et le fascicule n°7 de 2010 (phénomène déjà constaté avec la revue *Talanta*, notamment, dont plusieurs fascicules ne sont pas dépouillés par *Scopus*). Il faut cependant remarquer que si *Scopus* ne dépouille pas forcément l'entièreté des journaux qui font partie de sa base de données³⁶, il lui arrive aussi de compléter le dépouillement ultérieurement (un fascicule de l'année 2010 du *Journal of Animal Ecology* a été introduit en 2013 dans la base de données).

Recherche sur différents champs des références bibliographiques

La principale fonction d'une base de données bibliographique, dans un cadre universitaire, est d'obtenir les informations les plus pertinentes sur un sujet scientifique très précis. En l'absence de thesaurus, la recherche sur les mots-clés, les sujets, les mots des titres, est alors la plus appropriée. Parallèlement à ces possibilités, les éditeurs du *Web of Science* et de *Scopus* ont développé un volet bibliométrique qui permet une analyse des périodiques, des instituts scientifiques nationaux et des chercheurs par le biais d'indicateurs (facteur d'impact, citations, h-index, SJR, SNIP...). Ces fonctions impliquent une haute organisation des bases de données, notamment en ce qui concerne les auteurs qui doivent être identifiés de manière univoque.

Jusqu'à présent, ces aspects ne sont pas développés par *WorldCat.org* dont la vocation première n'est pas la même que *Web of Science* et *Scopus*. Sans volonté d'uniformiser les noms d'auteur ni de relier ceux-ci à un domaine, à un pays, à une institution, les recherches par auteur y sont relativement aléatoires.

Établir la bibliographie d'un auteur³⁷ est donc très difficile dans *WorldCat.org* (à moins d'avoir un nom peu courant). De plus, certains éléments liés à l'utilisation de guillemets ou de troncatures méritent d'être soulignés :

- La troncature (*) n'est pas reconnue : Kuty O* = Kuty O, mais pas Kuty Olgierd
- La position des mots joue un rôle si on utilise les guillemets. Au : "Kuty Olgierd" ≠ Au : "Olgierd Kuty"
- La meilleure solution est de composer soi-même la requête avec les opérateurs booléens : au:Kuty O OR au:Kuty Olgierd OR au:O Kuty
- La liste des différents collaborateurs (du moins de certains d'entre eux !) est donnée parmi les facettes, mais sans possibilité d'en sélectionner plus d'un à la fois.

Autres exemples de recherche sur un auteur³⁸

Scopus

Detrembleur C in AU permet de distinguer Christophe (chimiste) et Christine (orthopédiste). Total : **157 références** pour Christophe (81 pour Christine).

WorldCat.org

Detrembleur C in AU = 398 références (**380 articles**), mais impossible de distinguer (sauf manuellement) le chimiste (Christophe) de l'orthopédiste (Christine), ou d'autres éventuels homonymes.

Scopus

Hansenne M in AU = **100 références**. Michel Hansenne est le nom que portent un professeur de psychologie de l'ULg et un ancien ministre et directeur du Bureau International du Travail. Dans *Scopus*, seul le psychologue est présent.

WorldCat.org

Hansenne M in AU = 149 (**109 articles**). L'homonymie rend la recherche dans WorldCat.org particulièrement délicate et nécessite une analyse fastidieuse de toutes les notices. Celle-ci débouche sur 103 articles pour le psychologue et 6 pour le politicien.

Scopus

Laboury D in AU = **5 références**. Dimitri Laboury est égyptologue à l'ULg. Le petit nombre de références qui le concernent confirme un état de fait souvent constaté : quoi que multidisciplinaires, les bases de données telles que Scopus ou Web of Science ont une couverture de la littérature des sciences humaines moins large que celle des sciences pures ou appliquées.

WorldCat.org

Laboury in AU = 26 références dont il faut éliminer 10 doublons ou homonymies du nom d'auteur : **16 références** (parmi lesquelles les 5 références obtenues dans *Scopus*)

2 remarques : introduire l'initiale du prénom ou le prénom en entier limitera le nombre de réponses → pour un patronyme peu courant, le choix d'ignorer le prénom s'avérera utile ; dans le cas de nom de famille répandu, une telle recherche bibliographique peut devenir très ardue.

Par ailleurs, la complémentarité Articles/Monographies de *WorldCat.org* explique le plus grand nombre de réponses obtenues dans des disciplines (comme les sciences humaines) où les chercheurs publient indistinctement dans les revues scientifiques et dans la littérature non périodique.

Mots-clés, mots du titre,...³⁹

Scopus

Article title : shale gas AND all fields : hydraulic extraction = **55 réponses**

WorldCat.org

ti : shale gas AND kw : hydraulic extraction = **27 réponses**

Les réponses de *Scopus* sont plus nombreuses et plus récentes. Peu de recoupement avec celles de *WorldCat* dont plus de la moitié concerne des monographies. Sans conteste, les résultats obtenus chez *Scopus* sont plus intéressants.

Scopus

Article title : dendrochronology AND article title : tree-ring* = **18 réponses**

WorldCat.org

ti : dendrochronology AND ti : tree-ring* (limité au type de document "article") = 67 réponses - 18 doublons → **49 réponses**

Une explication de la grande différence entre le nombre de réponses trouvées chez *WorldCat.org* et chez *Scopus* vient du fait précédemment souligné que *Scopus* ne dépouille pas la totalité des revues qui font partie de sa base de connaissance : ici, 11 articles de la revue *Vernacular architecture*, qui répondent de façon pertinente à la question et qu'on trouve dans *WorldCat.org*, n'apparaissent pas dans *Scopus* (qui pourtant répertorie cette revue dont elle recense quelques articles isolés).

Sujet, Abstract

La recherche par sujet n'est possible que dans *WorldCat.org*, mais est peu fiable car seule une partie des références en est pourvue (sans doute à cause des sources différentes qui alimentent la base de données). Ex. : su :shale gas = 6.170 tandis que kw :shale gas = 14.746 (comprenant les 6.170 puisque : kw:shale gas AND su:shale gas = 6.170). Il n'empêche que les sujets peuvent être d'excellents critères de recherche pour une bibliographie sélective, avec des possibilités telles que : su:Grande-Bretagne Politique et gouvernement 20e siècle (1377 réponses).

Si la recherche par sujet n'est pas possible chez *Scopus*, par contre la majorité des références est accompagnée d'abstracts qui améliorent fortement la qualité des recherches "All fields".

Conclusions

Nos tests semblent confirmer que *WorldCat.org* est une base très bien fournie, y compris en terme de références d'articles scientifiques. Même si leur nombre est surestimé (doublons), la couverture des périodiques n'en est pas moins impressionnante, régulièrement (mais pas toujours) plus importante même que dans une base de données bibliographique comme *Scopus*.

Cependant, l'interface *WorldCat.org* a pour vocation de permettre la localisation de documents et pas la recherche documentaire. L'utilisateur qui voudrait effectuer une recherche documentaire via *WorldCat.org*, en particulier d'articles scientifiques, serait rapidement confronté à de multiples limites : champs de recherche spécialisés peu nombreux, syntaxe d'interrogation peu développée et/ou peu visible, difficulté d'éliminer les doublons dans les résultats, affichage et gestion des résultats limités à dix références à la fois, pauvreté des fonctions annexes (facettes, sauvegarde),...

Il est vrai qu'outre sa fonction de localisation de documents, offerte gratuitement pour l'utilisateur final, *WorldCat.org* est aussi une vitrine attractive pour d'autres développements, plus performants et qui font l'objet d'abonnements, comme *FirstSearch* et *WorldCat Local*.

Avec le développement d'outils de type "Web-Scale Discovery Tool", *WorldCat* ne peut être présenté comme un simple catalogue. Une grande force de *WorldCat* est probablement d'associer des contenus aussi divers que le principal catalogue mondial de bibliothèques académiques, des tables de matières d'éditeurs commerciaux, des données open access... La connaissance des pratiques et besoins des bibliothèques n'est certainement pas le moindre atout de l'OCLC dans le cadre du développement d'outils intégrés.

L'analyse comparée de *Scopus* et de *WorldCat.org* a permis de mettre en évidence une certaine complémentarité de ces outils plutôt qu'une concurrence entre eux. Mais cette analyse montre aussi combien les systèmes sont évolutifs et les frontières variables.

Il semble loin, le temps où les chercheurs passaient des heures dans les bibliothèques à dé-

pouiller d'imposants volumes alignés sur des rayonnages interminables pour faire leur bibliographie hebdomadaire.

La mise en ligne, par étapes successives, de ces bibliographies a été un progrès énorme pour le monde des scientifiques et des bibliothécaires, mais a nécessité d'abord une connaissance pointue du contenu des bases de données et des modes d'interrogation bibliographique : équations de recherche, opérateurs booléens... C'était peut-être l'apogée de ce genre de recherche, du moins l'époque où le chercheur pouvait encore avoir l'impression de maîtriser le système, pour obtenir les réponses les plus pertinentes à ses questions.

Petit à petit, l'évolution conduit à une prise en main de plus en plus complète des modes d'interrogation par les systèmes informatiques eux-mêmes, résultat d'une part de la concurrence des moteurs de recherche, en particulier de *Google* et de ses applications *Google Scholar* et *Google Books*, d'autre part d'un changement de comportement général de notre société qui réclame toujours plus de rapidité, voire d'immédiateté dans l'accès à l'information.

Les Web-Scale Discovery Tools prendront bientôt la première place sur les sites des institutions scientifiques. Les informations s'y retrouveront, et sans doute d'autres, mais organisées différemment et qui, sous une apparente facilité d'accès, exigeront toujours du lecteur sens critique, curiosité et persévérance, sous une forme un peu différente qu'aujourd'hui...

Philippe MOTTET
Ninfa GRECO

Université de Liège (ULg)
Bibliothèque des Sciences et
Techniques

Bât. B6 Bibliothèque des sciences et
techniques
Allée de la Chimie, 3
4000 Liège 1

pmottet@ulg.ac.be
N.Greco@ulg.ac.be

Mai 2013

Quelques lectures utiles

Ballew, Barbara S. Elsevier's Scopus database. *Journal of electronic resources in medical libraries* [en ligne, accès restreint], 2009 (consulté le 26 septembre 2012), vol. 6, p. 245-252.

<<http://search.ebscohost.com/login.aspx?direct=true&db=lih&AN=44081493&site=ehost-live>>

Bertot, John Carlo; Berube, Katy; Devereaux, Peter; Dhakal, Kerry; Powers, Stephen; Ray, Jennie. Assessing the usability of WorldCat Local: findings and considerations. *The Library quarterly* [en ligne, accès restreint], avril 2012 (consulté le 8 septembre 2012), vol. 82, n° 2, p. 207-221.
<<http://www.jstor.org/discover/10.1086/664588?uid=37943&uid=3737592&uid=2&uid=3&uid=37942&uid=67&uid=62&sid=21101306206021>>.

Boock, Michael ; Chadwell, Faye ; Reese, Terry. *WorldCat Local Task Force: report to LAMP* [en ligne]. 27 mars 2009 (consulté le 10 septembre 2012). <<http://hdl.handle.net/1957/11167>>.

Bosman, Jeroen ; Van Mourik, Ineke ; Rasch, Menno; Sieverts, Eric; Verhoeff, Huib. *Scopus reviewed and compared: the coverage and functionality of the citation database Scopus, including comparisons with Web of Science and Google Scholar* [en ligne]. Universiteitsbibliotheek Utrecht / Utrecht University Library, juin 2006 (consulté le 26 septembre 2012).
<<http://igitur-archive.library.uu.nl/DARLIN/2006-1220-200432/Scopus%20doorgelicht%20&%20vergeleken%20-%20translated.pdf>>.

Breeding, Marshall. WorldCat Local. *Library technology reports* [en ligne, accès restreint], 2007 (consulté le 26 septembre 2012), vol. 43, n° 2, p. 33-37.
<<http://search.ebscohost.com/login.aspx?direct=true&db=afh&AN=26074121&site=ehost-live>>.

Dess, Howard M. Database reviews and reports: Scopus. *Issues in science and technology librarianship* [en ligne], hiver 2006 (consulté le 26 septembre 2012). <<http://www.istl.org/06-winter/databases4.html>>.

Madarash-Hill, Cherie. The Effectiveness of librarian searching of Google, WorldCat, and a library online catalog. *College & undergraduate libraries* [en ligne, accès restreint], 2009 (consulté le 26 septembre 2012), vol. 16, p. 300-310.
<<http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=5&hid=18&sid=f30251f6-07f9-4154-8b66-ae8af62197ee%40sessionmgr10>>.

Les outils de découverte : rapport du groupe de travail du réseau des bibliothèques de l'Université de Monston. Version abrégée destinée au public, mars 2011. [en ligne]
<http://www.couperin.org/images/stories/documents/outils_decouverte_rapport_final_abrege_mars2011.pdf>
(consulté le 10 septembre 2012).

Vaughan, Jason. OCLC WorldCat Local. *Library technology reports* [en ligne, accès restreint], 2011 (consulté le 26 septembre 2012), vol. 47, n° 1, p. 12-21.
<<http://search.ebscohost.com/login.aspx?direct=true&db=afh&AN=57161082&site=ehost-live>>.

Ward, Jennifer. Why WorldCat Local? "Web scale" discovery and delivery of library resources. *Library technology reports* [en ligne, accès restreint], 2008 (consulté le 26 septembre 2012), vol. 44, n° 6, p. 7-16.
<<http://search.ebscohost.com/login.aspx?direct=true&db=afh&AN=34109253&site=ehost-live>>.

Notes

- 1 Produits proposés respectivement par les sociétés Ex-Libris, Ebsco et Serials Solutions.
- 2 Animal unicellulaire.
- 3 *Scopus* a été choisi en janvier 2012, pour 4 ans, comme base de données bibliographique de référence pour les universités de la Communauté française de Belgique et pour le Fonds de la Recherche scientifique (F.R.S. – FNRS).
- 4 WorldCat. *Wikipédia* [en ligne], 9 avril 2013 (consulté le 30 avril 2013). <<http://fr.wikipedia.org/wiki/WorldCat>>
WorldCat. *What is WorldCat?* [en ligne]. <<http://www.worldcat.org/whatis/default.jsp>> (consulté le 30 avril 2013).
- 5 1031 réponses dans *WorldCat* contre 448 chez *Scopus* au 25 février 2013, avant tout traitement des résultats.
- 6 *Scopus* [en ligne]. <<http://www.scopus.com/>> (consulté le 30 avril 2013).
- 7 Scopus. *What does Scopus cover?* [en ligne]. <<http://www.info.sciverse.com/scopus/scopus-in-detail/facts>>
(Consulté le 19 septembre 2012).
- 8 Ohio College Library Center, ASBL universitaire créée en 1965, devenu Online Computer Library Center, Inc. en 1981.
- 9 OCLC. *Available content* [en ligne]. <<http://www.oclc.org/worldcat-local/content.en.html>> (consulté le 7 mai 2013).

- ¹⁰ OCLC. *WorldCat* [en ligne]. <<http://www.worldcat.org/>> (consulté le 7 mai 2013).
- ¹¹ <http://www.oclc.org/support/services/worldcat-org/faq.en.html> consulté le 7 mai 2013.
- ¹² L'abonnement à *FirstSearch* permet non seulement une plus grande visibilité de son catalogue sur *WorldCat.org*, mais également sur *Google* et *Yahoo !*, partenaires de *WorldCat*.
- ¹³ OCLC. 1,814 familiar collections in one place [en ligne]. <<http://www.oclc.org/worldcat-local/content/dblist.en.html>> (consulté le 7 mai 2013).
- ¹⁴ *Scopus* [en ligne]. <<http://www.scopus.com/>> (consulté le 30 avril 2013).
- ¹⁵ OCLC. *WorldCat* [en ligne]. <<http://www.worldcat.org/>> (consulté le 7 mai 2013).
- ¹⁶ Le choix d'un échantillon de revues à analyser dans ce genre d'étude comparative n'est jamais idéal. Notre comparaison porte sur les 50 journaux les plus représentés dans les références publiées par les chercheurs de l'Université de Liège, à la date du 7 septembre 2012. La source est le répertoire institutionnel de l'Université de Liège, *ORBi*, consultable à l'adresse <<http://orbi.ulg.ac.be/>> (consulté le 7 mai 2013).
- ¹⁷ Mise à jour des résultats de la recherche effectuée le 15 avril 2013.
- ¹⁸ 104 articles parus de 1983 à 1994 sont cependant retrouvés via l'onglet "Document search".
- ¹⁹ La revue est connue de *WorldCat* (7 exemplaires différents pour la notice du périodique), mais n'est pas dépouillée.
- ²⁰ 1 article paru en 1978 peut être retrouvé *via* l'onglet "Document search".
- ²¹ La revue est connue de *WorldCat* (3 exemplaires différents pour la notice du périodique), mais n'est pas dépouillée.
- ²² Pour les années 1994-1996, il s'agit en fait de *British Veterinary Journal* (ISSN 0007-1935).
- ²³ La revue est connue de *WorldCat* (2 exemplaires différents pour la notice du périodique), mais n'est pas dépouillée.
- ²⁴ Même remarque (18 exemplaires différents).
- ²⁵ ISSN électronique, cette revue n'existe que sous cette forme.
- ²⁶ La revue est connue de *WorldCat* (23 exemplaires différents pour la notice du périodique), mais n'est pas dépouillée.
- ²⁷ 103 références sont erronément obtenues pour 1994 : il s'agit d'un titre suivi de "Archives internationales de physiologie, de biochimie et de biophysique" avec un autre ISSN (0778-3124).
- ²⁸ La revue est connue de *WorldCat* (2 exemplaires différents pour la notice du périodique), mais n'est pas dépouillée.
- ²⁹ Même remarque (1 exemplaire).
- ³⁰ Les résultats obtenus ne concernent pas uniquement le supplément de la revue *Astronomy and astrophysics*, mais également certains articles de la revue elle-même, malgré la distinction nette entre les deux ISSN.
- ³¹ ISSN électronique, cette revue n'existe que sous cette forme.
- ³² La revue est connue de *Worldcat* (4 exemplaires différents pour la notice du périodique), mais n'est pas dépouillée.
- ³³ Présentation, en marge de la liste des résultats, de divers éléments constitutifs des références obtenues (auteurs, années, types de documents, langues, etc.) dont la sélection permet de ne conserver que les réponses les plus pertinentes.
- ³⁴ OCLC. *WorldCat.org frequently asked questions* [en ligne]. <<http://www.oclc.org/support/services/worldcat-org/faq.en.html>> (consulté le 07 mai 2013).
- ³⁵ Recherches effectuées le 13 mai 2013.
- ³⁶ En novembre 2011, à la remarque que nous lui faisons concernant ce genre de lacunes assez inhabituelles, *Scopus* nous avait répondu ne pas avoir de contrôle sur le contenu des sources bibliographiques extérieures (il s'agissait de *Medline*) qui alimentent sa base de données.

³⁷ Les quelques exemples concernent des chercheurs de l'Université de Liège, pris au hasard dans différentes disciplines scientifiques. Pour avoir une bibliographie complète de ces auteurs, il est nécessaire de consulter *ORBi* (<<http://orbi.ulg.ac.be/>>)

³⁸ Recherches effectuées le 14 mai 2013.

³⁹ Recherches effectuées le 14 mai 2013.