

# A comparison of Scopus and Web of Science for a typical university

ELIZABETH S. VIEIRA, JOSÉ A. N. F. GOMES

*Requimte/Departamento de Química, Faculdade de Ciências, Universidade do Porto,  
Rua do Campo Alegre, 687, 4169-007 Porto, Portugal*

For many years, the ISI Web of Knowledge from Thomson Reuters was the sole publication and citation database covering all areas of science thus becoming an invaluable tool in bibliometric analysis. In 2004, Elsevier introduced Scopus and this is rapidly becoming a good alternative. Several attempts have been made at comparing these two instruments from the point of view of journal coverage for research or for bibliometric assessment of research output.

This paper attempts to answer the question that all researchers ask, i.e., what is to be gained by searching both databases? Or, if you are forced to opt for one of them, which should you prefer? To answer this question, a detailed paper by paper study is presented of the coverage achieved by ISI Web of Science and by Scopus of the output of a typical university. After considering the set of Portuguese universities, the detailed analysis is made for two of them for 2006, the two being chosen for their comprehensiveness typical of most European universities. The general conclusion is that about 2/3 of the documents referenced in any of the two databases may be found in both databases while a fringe of 1/3 are only referenced in one or the other. The citation impact of the documents in the core present in both databases is higher, but the impact of the fringe that are present only in one of the databases should not be disregarded as some high impact documents may be found among them.

## Introduction

Until recently, the Web of Science (WoS) was the only source for the assessment scientific output worldwide because of its multidisciplinary and international coverage. It includes more than 10000 journals and comprises three citation databases: Arts & Humanities Citation Index (with coverage going back to 1975), Social Sciences Citation Index (with coverage going back to 1956) and Science Citation Index Expanded (with coverage going back to 1900) [THOMSON REUTERS, 2008A]. The WoS contains over 38 million records and, each year, over 1.5 million new records and 23 million new cited references from more than 250 disciplines of the sciences, the social sciences and the arts and humanities are introduced [THOMSON REUTERS, 2008B].

Elsevier Science launched in 2004 the database Scopus that can already be considered a good alternative to that of Thomson Reuters. As described by the owners,

---

Received December 9, 2008; Published online April 16, 2009

*Address for correspondence:*

JOSÉ A. N. F. GOMES

E-mail: jfgomes@fc.up.pt

0138–9130/US \$ 20.00

Copyright © 2009 Akadémiai Kiadó, Budapest

All rights reserved

Scopus contains 33 million records of which 16 million include references going back to 1996 and 17 million pre-1996 records going back as far as 1841. These records were collected from over 15,000 peer-reviewed journals from more than 4,000 international publishers, including over 1200 Open Access journals, 500 Conference Proceedings, over 600 Trade Publications and 200 book series from all areas of knowledge [SCOPUS FAQs, 2008]. However it is important to note that the coverage of a journal by Scopus can have breaks, that is, for some journals Scopus makes a partial coverage. If we consult the title list that contains the journals indexed in Scopus we can see that for various journals the coverage by Scopus is partial [SCOPUS FAQs, 2008]. Scopus classifies journals in four wide scientific areas: Physical Sciences (with more than 5500 titles), Health Sciences (more than 5300 titles), Social Sciences (more than 2850 titles) and Life Sciences (more than 3400 titles). In addition, documents are classified in one of 27 subject areas. Scopus claims a worldwide coverage with more than half of its contents originating from Europe, Latin America and the Asian-Pacific Region [SCOPUS FAQs, 2008] while WoS covers mainly North America and Western European [MEHO & AL., 2007]. The competition between the two providers is intense and has led to the frequent upgrade of the services offered by both databases in the last few years.

Beyond Scopus, another important database was launched in November 2004, the Google Scholar. Contrary to the other databases referred to above, Google Scholar is available free to the whole internet community. This is a tool that allows researchers to locate a wide array of scholarly literature on the Web, including journals, books, preprints, technical reports and abstracts taken from the repositories of universities, academic institutions, and research groups around the world. Each day, more information is available online in Google Scholar making this a tool increasingly attractive to students and researchers from all over the world.

Several studies can be found in the literature making detailed descriptions of the main features of Scopus [JACSÓ, 2005; FIGERMAN, 2005; LAGUARDIA, 2005] and comparing the databases WoS, Scopus and Google Scholar with the aim of assessing the number of citations obtained by a particular set of documents in each of them [JACSÓ, 2005; BAKKALBASI & AL., 2006; MEHO & AL., 2006]. Other studies, analyze the set of journals covered by each database [GAVEL & AL., 2008; BOSMAN & AL., 2006] as well as their interface accessibility and usability [FIGERMAN, 2005]. JACSÓ [2005] compared Scopus, WoS and Google Scholar from the point of view of the number of items included, searching for a specific journal (Current Science) with the aim of testing the breadth of coverage. He found that the Google Scholar coverage of the journal Current Science is good and that there is considerable overlap between Scopus and WoS. Beyond this study, JACSÓ [2005] also conducted several studies searching for documents citing the 1955 Science paper of Garfield and searching for the 30 most cited articles of Current Science. For these 30 articles he found that WoS and

Scopus have almost identical citedness scores for many of the articles while Google Scholar has the source records for less than half of the articles.

BAKKALBASI & AL. [2006] compare citation counts for articles in the areas of oncology and condensed matter physics published in 1993 and in 2003. His results showed that for oncology in 1993 WoS returned the highest average number of citations (45.3), while Scopus returned the highest average number for oncology in 2003 (8.9). For condensed matter physics, WoS returned the highest number of citations both in 1993 and in 2003 (22.5 and 3.9 respectively).

MEHO & AL. [2006] used WoS, Scopus and Google Scholar to locate citations to the publications of the 25 faculty members of the Library and Information Science Department of the University of Indiana, ranking the members according to the counting obtained. The rankings from Scopus and WoS coincide at the top and the bottom but diverge significantly in the middle positions; Google Scholar stands out for its good coverage of conference proceedings as well as international, non-English language journals. They suggested that the use of Scopus and Google Scholar, in addition to WoS, helps reveal a more accurate and comprehensive picture of the scholarly impact of authors. BAR-ILAN & AL. [2007] proposed a set of measures for evaluating the similarity between rankings. They compared rankings of publications of 22 highly cited Israeli researchers as measure by the citation counts in Google Scholar, Scopus and WoS. The results of these measures showed high similarities between the rankings of the WoS and Scopus and lower similarities between Google Scholar and the other databases. NORRIS & AL. [2007] used the journal articles submitted for the 2001 Research Assessment Exercise in the social sciences and the list of journals indexed in the International Bibliography of the Social Sciences to assess the coverage of four data sources (CSA Illumina, Google Scholar, Scopus and WoS). They found that Scopus provided the best coverage of social sciences literature from among these data sources. BAR-ILAN [2008] compared the h-indices of a list of highly cited Israeli researches based on citation counts from WoS, Scopus and Google Scholar. The results obtained through Google Scholar were considerably different from the results based on Scopus and WoS. MEHO & AL. [2008] examined the differences between Scopus and WoS in the citation counting, citation ranking and h-index of 22 top human-computer interaction researchers from EQUATOR – a large British Interdisciplinary Research Collaboration project. The results indicated that Scopus provides a significantly better coverage of human-computer interaction literature than WoS primarily due to coverage of relevant ACM (Association for Computing Machinery) and IEEE (Institute of Electrical and Electronics Engineers) peer-reviewed conference proceedings. GAVEL & AL. [2008] made the comparison between the active titles (titles currently being indexed) in WoS, Scopus and some specialized databases (Medline and Embase). The number of titles covered by Scopus was found to exceed that covered by WoS by 4789, even if Scopus titles of source types other than “Journal” are excluded. In all, 6256

journal titles out of the 13690 in Scopus are not covered in WoS. On the other hand, 1467 out of 8901 titles in WoS are not covered in Scopus. A comparison of the Scopus's coverage with respect to Ulrich<sup>1</sup> was published by MOYA-ANEGON & AL. [2006] taking a series of variables into account such as journal subject, geographical origin, publisher and the language of the publication. It is concluded that Scopus is balanced in terms of subject areas, languages and editors when compared with Ulrich's Core and that Scopus has quite a homogenous global representation in nearly all areas except Arts and Humanities.

To our knowledge, no study has been published focusing on the relative value of the WoS and Scopus when considered for subscription by a university library, a question of great relevance as Scopus appears to become a serious contender for paid subscription. The aim of this study is to identify the major differences between Scopus and WoS and how far they should be considered as complementary to one another from the point of view of a multidisciplinary institution. To achieve this, the scientific production referenced in Scopus and WoS from each of the 16 Portuguese universities belonging to CRUP<sup>2</sup> (the Council of Rectors of Portuguese Universities) in the period between 2000 and 2007 is considered. The search has been limited to this period for three reasons: (1) Portugal has had the highest growth of scientific production in the last 15 years of any European country so that very long time spans imply a very heterogeneous set of publications; (2) JACSÓ [2005] showed that the number of records indexed each year in WoS and Scopus is similar after 2000; (3) The recognition that both databases have introduced major changes in their coverage suggests that decisions about the current status must be based on recent data.

In the next section below, the methodology used to count the documents in Scopus and in WoS for the Portuguese universities is described and evaluated by comparison with other published studies. The process adopted to make the comparison between the two databases is also described. After showing the results obtained from Scopus and from WoS for the Portuguese universities, a very detailed study on a paper-by-paper basis is then reported for the Universidade de Coimbra and the Universidade de Lisboa. The analysis of the Scopus referenced publications of these two universities by wide scientific areas is also presented. The final section synthesizes the major conclusions that can be drawn from the study.

## Materials and methodology

The first difficult step in any bibliometric study of an institution is to define a reliable methodology to collect documents from the databases that should be associated

---

<sup>1</sup> Ulrich's Periodicals Directory, [www.ulrichsweb.com](http://www.ulrichsweb.com)

<sup>2</sup> CRUP, Conselho de Reitores das Universidades Portuguesas, <http://www.crup.pt>

with that particular institution. In this paper we are concerned with the Portuguese universities affiliated to CRUP, the Council of Rectors of Portuguese Universities.<sup>3</sup> Its membership includes all 15 public universities and the Catholic University.<sup>4</sup> In this study, one document is associated with a particular university when one of the author's addresses in the database makes a clear reference to that university, to one of its subunits or to one of its research entities. This methodology was tested against known data for some universities to certify that the count is reliable. Small differences, normally in the order of a few percent may be associated with database errors and to the criteria used to define the institutional contour. In defining these contours the description given in the university website is used, together with the affiliation of the majority of the senior researchers in the research units funded by the Fundação para a Ciência e a Tecnologia.<sup>5</sup> Publications originating from hospitals linked to universities were counted only when the university was mentioned in the address. In the case of the hospitals at Coimbra, all papers were counted as the university is mentioned in the official name of the hospitals without this meaning ownership or a special relation.

Other studies of the referenced scientific production of the Portuguese universities in WoS were made by SCImago,<sup>6</sup> using the method of correction of addresses, [UNIVERSIA, 2008] and by the Centre for Science and Technology Studies (CWTS)<sup>7</sup> [CARRONDO, 2008; VAN RAAN, 2008]. The results are consistent with those reported here within an error explained by the difference in criteria of association of each document to the institutions. For all documents, the error found for each of the 6 years common to both studies was 1% and the results presented here are in most cases larger than those found by SCImago or CWST. The few cases where the difference is larger than 1% are associated with universities where addresses are particularly difficult to discern or the subunits have a more dubious dependence when seen from afar. This suggests that some subunits were not associated to the universities by the SCImago and CWST while they are included in this study, even if they are formally autonomous, but driven by researchers affiliated with that particular university and with its formal permission. These peculiarities of the Portuguese scientific system are important if a fair comparative assessment of the universities is sought.

In a second step, two Portuguese universities were chosen (Universidade de Coimbra and Universidade de Lisboa) and a comparison is made between all referenced documents in the WoS and in Scopus for these two universities in 2006. These universities were chosen because of its significant size and multidisciplinary nature. As such, they may be considered as models for any other comprehensive university. All records were analyzed individually to ascertain the number of documents that occurred

---

<sup>3</sup> Conselho de Reitores das Universidades Portuguesas, CRUP, <http://www.crup.pt>

<sup>4</sup> Universidade Católica Portuguesa, <http://www.ucp.pt>

<sup>5</sup> Fundação para a Ciência e a Tecnologia, <http://www.fct.pt>

<sup>6</sup> <http://www.scimago.es>

<sup>7</sup> <http://www.cwts.nl>

in both databases and those that occurred only in one or the other. In order to establish the relevance of the documents that were identified in only one of the databases, the number of citations per document (up to mid-2008) was calculated and compared with the number of citations per document obtained for the total ensemble of documents identified in each database and for the documents common to both databases.

To ascertain whether Scopus might over represent a particular scientific area, (considering those areas used by Scopus for classifying the journals, Physical Sciences, Health Sciences, Social Sciences and Life Sciences) we considered the documents identified only in Scopus for the two selected universities and distributed them by the four areas. This distribution was done by establishing the correspondence between the field "Source Title" shown in the list of journals available from Scopus [SCOPUS INFO, 2008] and the journal in which each document originated by *Universidade de Coimbra* and *Universidade de Lisboa* was published. When the Source Title is not classified in one of the four areas, the subject area associated with the document was used to find the scientific area.

For the two universities analysed in more detailed all data were taken from Scopus and WoS in July of 2008. For the other universities the number of documents referenced in Scopus were also taken from Scopus in July of 2008 and the counts reported for WoS were taken from NOUWS & AL. [2008].

## Results

The number of documents referenced in Scopus and WoS for Portugal until 2007 is 76290 and 72413, respectively. The number of documents referenced in both databases until 2007 for the universities affiliated to CRUP is presented in Table 1.

For Portugal the number of documents referenced until 2007 in Scopus is higher than in WoS.

However we can see that for some universities the number of documents is higher in WoS than in Scopus. On the other hand, if we consider the period between 2000 and 2007, the difference between the two databases appears to be smaller. For some universities (Universidade Católica, Universidade do Porto) the number of documents referenced is even higher in Scopus, while it is lower until 2007. Now analysing the documents referenced in WoS and Scopus until 1999 and comparing with the period between 2000 and 2007, we find that the count is growing very fast with this last period collecting as much as twice the number of the past recorded history. This is the result of a particularly high rate of growth of Portuguese science and justifies that we limit our analysis to the last few years.

Table 1. Number of referenced documents in Scopus and in the WoS for Portugal and for Portuguese universities belonging to CRUP

Institution	Documents published					
	until 1999		2000–2007		until 2007	
	WoS	Scopus	WoS	Scopus	WoS	Scopus
ISCTE	21	24	124	187	145	211
Universidade Aberta	7	7	87	120	94	127
Universidade dos Açores	141	98	419	435	560	533
Universidade do Algarve	358	359	1622	1606	1980	1965
Universidade de Aveiro	1152	1334	4445	5003	5597	6337
Universidade da Beira Interior	107	153	536	648	643	801
Universidade Católica	351	184	556	531	907	715
Universidade de Coimbra	3634	3228	5615	5757	9249	8985
Universidade de Évora	206	159	749	785	955	944
Universidade de Lisboa	4084	3450	5728	5604	9812	9054
Universidade da Madeira	96	105	312	334	408	439
Universidade do Minho	817	873	2996	3338	3813	4211
Universidade Nova de Lisboa	1949	1672	4105	4018	6054	5690
Universidade do Porto	4518	3898	8410	8423	12928	12321
Universidade Técnica de Lisboa	4945	4763	8260	9052	13205	13815
UTAD	157	164	875	916	1032	1080
Portugal	26206	26486	46207	48151	72413	74637

The scientific production of Portuguese universities as measured by the methodology described above is presented in Table 2.

Table 2. Number of referenced documents in Scopus and in the WoS for the Portuguese universities in 2000 to 2007

Institution	2000		2001		2002		2003		2004		2005		2006		2007	
	WoS	Scopus	WoS	Scopus	WoS	Scopus	WoS	Scopus	WoS	Scopus	WoS	Scopus	WoS	Scopus	WoS	Scopus
ISCTE	5	3	5	5	7	7	18	17	22	37	18	34	27	41	22	43
Univ Aberta	6	7	12	14	7	7	16	24	15	18	13	19	11	15	7	16
Univ Açores	36	37	45	36	32	24	51	58	48	50	56	60	88	92	63	78
Univ Algarve	113	107	135	118	156	134	203	201	187	203	254	248	293	303	281	292
Univ Aveiro	279	282	337	353	427	405	513	575	644	709	703	838	845	942	697	899
Univ Beira Interior	36	43	55	59	55	59	56	64	73	92	82	101	87	114	92	116
Univ Católica	51	45	54	50	57	43	67	70	69	71	72	69	88	84	98	99
Univ Coimbra	458	475	515	473	612	532	678	699	756	796	741	816	943	969	912	997
Univ Évora	53	51	48	39	62	66	85	89	88	105	109	112	161	164	143	159
Univ Lisboa	552	446	534	453	609	518	632	632	759	765	748	800	1022	1010	872	980
Univ Madeira	18	20	28	26	26	24	36	41	33	32	47	49	59	68	65	74
Univ Minho	192	189	249	229	269	259	339	380	450	495	435	484	581	644	481	658
Univ Nova Lisboa	371	314	411	309	437	380	460	483	528	531	586	597	717	732	595	672
Univ Porto	684	638	713	637	837	687	901	914	1122	1146	1260	1300	1484	1529	1409	1572
Univ Téc. Lisboa	747	785	830	776	918	907	986	1047	1112	1268	1154	1325	1355	1479	1158	1465
UTAD	35	33	66	61	70	60	83	99	128	131	134	158	174	177	185	197
Portugal	3792	3864	4115	4082	4748	4526	5221	5637	5991	6372	6407	7038	7990	8378	7983	8254

The general conclusion is that Scopus has a larger coverage, representing, on average, 104% of WoS. In the latter years, the relative coverage of Scopus appears to increase further, reaching ca. 110% in the last three years. As WoS is supposed to have an entry criterion based on the measured impact of the journal, one may suppose that Scopus includes more documents with a smaller impact as measured by the average number of citations generated. Before going into the analysis of the relative impact of

the documents in the two databases, we compare the number of documents recorded for each of the Portuguese universities in each year 2000 to 2007.

Comparing the different universities, the outliers are the ISCTE and Aberta with much higher results in Scopus than in WoS. These two institutions are relatively small and specialized with a relatively high quota of the social sciences and this may be the cause for this result but the absolute figures are small and difficult to trust if extrapolated for larger institutions.

For a more detailed understanding of what we may expect from the two databases for large institutions, we select a single year, 2006, and take the analysis into more detail for the Universidade de Coimbra and the Universidade de Lisboa. These are relatively large universities covering all traditional academic areas, including medicine, so that they may be considered as representative of the typical European university. The year 2006 was selected as the most recent for which a reasonable citation span could be analyzed. In fact the databases have been changing quickly in the last few years so that considering an earlier year might be less representative of the current situation. The detailed counts in Scopus and in WoS are shown in Figure 1 and Table 3.

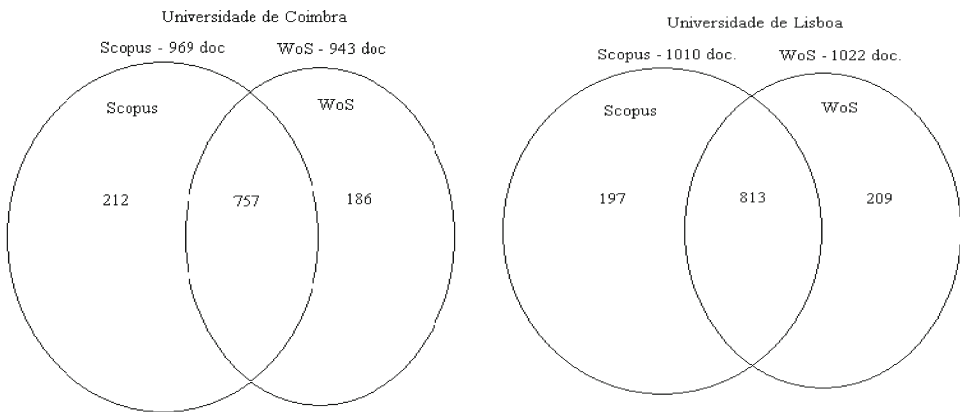


Figure 1. Number of documents identified in Scopus and in WoS for 2006, as well as in only one and the other and in both

The total number of documents referenced in any of the two databases is 1155 for the Universidade de Coimbra and 1219 for the Universidade de Lisboa for the year 2006. Of the 1155 documents of Coimbra, 969 are found in Scopus and 943 in WoS; while 757 are common to both databases, 212 are found only in Scopus and 186 only in WoS. Of the 1219 documents of Lisboa, 1022 documents are found in WoS and 1010 in Scopus; while 813 are common to both databases, 197 documents only are referenced in Scopus and 209 only in WoS.



As recognized above, the number of documents in each database may differ by not more than 3% but this hides a much larger difference in their contents. About 2/3 of the total documents referenced for the Universidade de Coimbra and Universidade de Lisboa are listed in both databases, while 1/6 is listed only in one and another (different) 1/6 is listed in the other. The differences between the two databases may be due to different abstracting policies but are frequently due to apparent errors in the transfer process from the journal to the database. To clarify this, a document by document analysis was done with the results summarized in Table 3.

Table 3. Explaining the different counts of referenced documents in Scopus and WoS for the Universidade de Coimbra and Universidade de Lisboa in 2006: Possible explanations for the missing documents.

Comments	Universidade de Coimbra		Universidade de Lisboa	
	WoS	Scopus	WoS	Scopus
(A1) Absence of journal, book series or conference proceedings	137	6	122	3
(A2) Scopus referenced documents only in the ISI Proceedings	57	0	20	0
(A3) Only overall reference to the conference proceeding	0	87	0	87
(B1) Errors of address or date	14	21	50	41
(B2) Documents unidentified)	3	25	5	33
(B3) Volume, issue or supplement of journal not found	0	47	0	43
(B4) Duplicate documents	0	0	0	-2

The comparison between the counts for the two universities may be explained by a variety of situations:

A. Different abstracting policies

B. Apparent errors in the process of constructing the database

A1) Documents were not identified because the journal in which they were published was not indexed in Scopus or in WoS. In Scopus, we found 137 for Coimbra (and 122 documents for Lisboa), for which the journal was not indexed in WoS for the year studied. A similar situation occurs for 6 documents of Coimbra (and 3 of Lisboa) that are found in WoS for sources not indexed in Scopus.

A2) Documents not found in WoS but recorded in ISI Proceedings (ISI Proceedings is part of the set of instruments offered by ISI Web of Knowledge that allow researchers access to the published literature from what Thomson Reuters considers to be the most significant conferences, symposia, seminars, colloquia, workshops and conventions worldwide). In this situation we identified 57 documents for the Universidade de Coimbra and 20 documents for the Universidade de Lisboa.

A3) Scopus policy of making a reference to the conference as a single item without individual entries for each contribution, therefore missing the names and addresses of authors. As WoS considers each item delivered in the conference as an autonomous document, we found 87 WoS documents of the Universidade de Coimbra (and 87 of Lisboa) without individual correspondence in Scopus.

B1) Errors of address or date. We found 21 documents for the Universidade de Coimbra and 41 for the Universidade de Lisboa where Scopus has errors of address or date that explain why they were not extracted by our methodology. The WoS has,

respectively, 14 and 50 errors of the same type. A variety of errors are found related to the absence or incorrect reference to the Universidade de Coimbra, to the Universidade de Lisboa or to the country in the field of the author's addresses. In the case of the Universidade de Lisboa, the most common error in WoS is the incorrect reference to the Universidade Nova de Lisboa and not to the Universidade de Lisboa as stated in the original document.

B2) Missing documents. Several documents were not found although the journal (volume, issue and supplement) where they were published was indexed in the database. This was the apparent reason for 3 Scopus documents for the Universidade de Coimbra (and 5 for the Universidade de Lisboa,) to be missing in WoS. Similarly, 25 WoS documents of the Universidade de Coimbra (and 33 of the Universidade de Lisboa) were missing from Scopus in this situation.

B3) Missing volume or issue or supplement. Some documents were not identified in Scopus because the volume, issue or supplement of a journal retained in Scopus is missing from the database. This was the apparent reason for 47 WoS documents of the Universidade de Coimbra (and 43 of the Universidade de Lisboa) to be missing from Scopus.

B4) Duplications. The last situation observed was the duplication of two documents in Scopus for the Universidade de Lisboa. These two documents contain the reference to the same journal, but with (erroneously) different pagination.

Another point that deserves clarification is the relative impact of the different classes of documents, those referenced in one of the databases analyzed or in both of them. The results are shown in Figure 2.

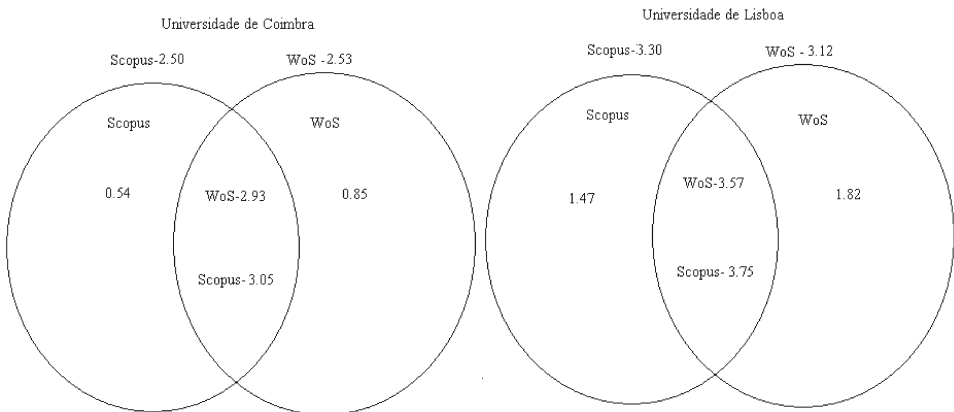


Figure 2. Number of citations per document obtained for the documents referenced in each of the databases, only in one of them, and for documents common to both

From the point of view of citations (and it should be stressed that we counted only the citations of the 2006 documents that were recorded up to July/2008), the documents referenced in both databases are clearly the most cited. For the Scopus documents left out of the WoS, the impact is much lower for Coimbra than for Lisboa.

Overall, it may be stated that the core set of the documents present in both Scopus and WoS are the most cited but the sets present only in one of the databases should not be disregarded as they will include many publications of high impact.

To gain a better understanding of the values found for the average citation counts of the documents referenced only in Scopus or only in WoS, we look in greater detail into these sets.

Figure 3 shows the number of documents only referenced in Scopus or in WoS for the Universidade de Coimbra and Universidade de Lisboa, in 2006, with zero citations so far and the impact of the documents with one or more citations.

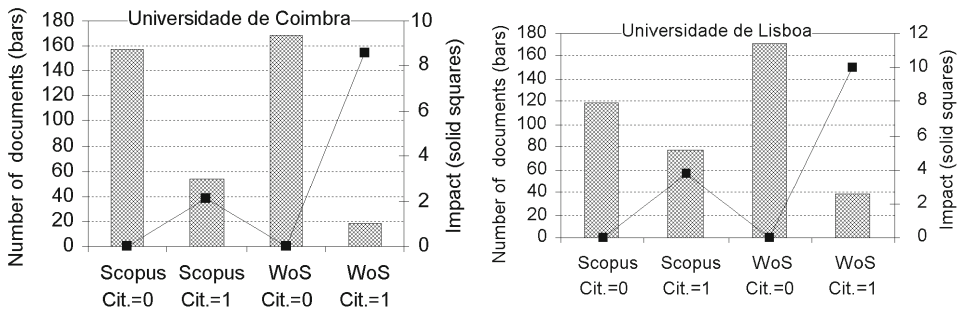


Figure 3. Number of documents referenced only in one of the databases with zero citations and one or more citations. The impact of these documents is shown by a square bullet

One feature stands out, the large number of documents that received no citations. These represent 75% and 61% of the documents referenced only in Scopus for the Universidade de Coimbra and Universidade de Lisboa, respectively. This percentage of never cited documents is still higher (90% and 82%, respectively) for those referenced only in WoS. If we calculate the average number of citations for the documents that received at least one citation, we find that the value obtained (2.13 and 3.73 for the Universidade de Coimbra and Universidade de Lisboa, respectively) is close to that obtained for the whole set of Scopus documents. Typically, these are documents extracted from journals not taken in the WoS (and some possible errors, as discussed above) but it is now recognized that 25% to 40% of these have a relatively good impact. The same procedure may be followed for referenced documents only in WoS, but here the results show that the number of citations per document is higher than the average found for the total ensemble of documents (8.6 and 10.05 for the Universidade de Coimbra and Universidade de Lisboa, respectively). This can be understood if we notice

that most of the relatively few cited documents in this set are left out due to errors of address. The uncited documents are mostly proceeding type documents that the WoS considers, but Scopus includes an overall reference of the conference proceedings without the descriptor of each individual contribution and, therefore, of the names or addresses of their authors.

An attempt is now made at explaining the variation between the two universities by their different scientific spectrum. In Table 4, we consider the journal classification of Scopus to calculate the share of Scopus documents in different areas that are left out of WoS.

Table 4. Distribution of the Scopus documents by scientific area for the Universidade de Coimbra and Universidade de Lisboa in 2006

Area	Universidade de Coimbra			Universidade de Lisboa		
	Documents identified only in Scopus	Scopus documents in common with WoS	Percentage documents only Scopus	Documents identified only in Scopus	Scopus documents in common with WoS	Percentage documents only Scopus
Physical Sciences	137	465	23%	102	448	19%
Health Sciences	50	57	49%	49	88	36%
Social Sciences	10	38	21%	28	25	53%
Life Sciences	16	197	8%	18	249	7%
Multidisciplinary	0	0	0%	0	3	0%

The Physical Sciences take the larger share and appear to be evenly represented as it should be remembered that, on average, about 20% of the documents are left out of the other database. For documents classified as Health Sciences and, to a lesser extent, for the Social Sciences, Scopus may include a larger than expected percentage of documents not found in the WoS. For the Life Sciences, on the other hand, Scopus appears to under represent the scientific production in both universities. These variations may be related to the different emphasis put in proceedings publications in different scientific areas. The study about the distribution of documents by scientific areas was only made for Scopus documents because in the WoS the same study can't be reproduced.

## Conclusions

This paper presents a comparison between the reference databases of ISI Web of Science and Scopus from the point of view of a typical comprehensive university. This was done by looking in detail at the documents referenced in these databases from two Portuguese universities, Coimbra and Lisbon. In each case, about 2/3 of the total documents referenced for the Universidade de Coimbra and the Universidade de Lisboa are referenced in both databases and a set of 1/3 appear in one of the databases only.

The detailed analysis of the documents referenced in only one of the databases showed the following situations:

Different abstracting polices related with: (1) journals, book series or conference proceedings not indexed in Scopus or in WoS; (2) WoS documents not identified in Scopus because Scopus has overall reference to the conference proceeding without descriptor of each individual contribution and, therefore, of the names or addresses of their authors and (3) some Scopus documents were not identified in WoS but were recorded in ISI Proceedings.

Apparent errors in the process of constructing the database related with: (1) errors introduced in the address or date; (2) documents not found although the journal (volume, issue or supplement) where the were published was indexed in the database; (3) issue, volume or supplement of journal not found although the journal to be indexed in the database for the year studied and (4) duplicated documents.

Considering the counting of citations for the referenced documents in both databases for the Universidade de Coimbra and the Universidade de Lisboa in the year 2006 we concluded:

The documents from Universidade de Coimbra and Universidade de Lisboa that are referenced in both databases are the most cited (about 2/3 of the total documents for each university);

The documents referenced only in Scopus come mostly from journals not taken in WoS but some 25% to 40% of them have a good citation record.

Most of the documents referenced only in WoS are proceeding-like with no citations that do not appear in Scopus as an overall reference to the conference proceedings is made without a descriptor of each individual contribution; the few highly cited documents in the set of those referenced only in WoS correspond to those that do not appear in Scopus due to addressing errors.

The distribution of the referenced documents in Scopus for the Universidade de Coimbra and the Universidade de Lisboa by scientific areas considered in Scopus has shown that the area Physical Sciences has the largest number of documents. Regarding the Scopus documents that were not found in WoS, it appears that the highest percentage of documents are classified in the area Health Sciences.

## References

- BAKKALBASI, N., BAUER, K., GLOVER, J., WANG, L. (2006), Three options for citation tracking: Google Scholar, Scopus and Web of Science, *Biomedical Digital Libraries*, 3 : 7. (Available at: <http://www.bio-diglib.com/content/pdf/1742-5581-3-7.pdf>).
- BOSMAN, J., VAN MOURIK, I., RASCH, M., SIEVERTS, E., VERHOEFF, H. (2006), *Scopus Reviewed and Compared*. (Avaliable at: <http://igiturarchive.library.uu.nl/DARLIN/2006-1220-200432/Scopus%20doorgelicht%20&%20vergeleken%20-%20translated.pdf>) (Accessed: July, 2008).
- BAR-ILAN J., LEVENE M, LIN A. (2007), Some measures for comparing citation databases, *Journal of Informetrics* 1 (1) : 26–34.

- BAR-ILAN, J. (2008), Which h-index? – A comparison of WoS, Scopus and Google Scholar, *Scientometrics* 74 (2) : 257–271.
- CARRONDO, M. A. (2008), *Bibliometric Study of UNL – Partial Characterization Based On The Publications (2000-2006) in Journals Indexed to Web of Science*. (Available at: <http://www.unl.pt/noticiastodas/MACarrondo.pdf>) (Accessed July, 2008).
- FINGERMAN, S. (2005), Scopus: profusion and confusion, *Online Medford*: Mar/Apr, 29 : 2, 36–38.
- GAVEL Y., ISELID, L. (2008), Web of Science and Scopus: a journal title overlap study, *Online Information Review*, 32 (1) : 8–21.
- JACSÓ, P. (2005), As we may search – Comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases, *Current Science*, 89 (9) : 1537–1547.
- LAGUARDIA, C. (2005), *E-Views and Reviews: Scopus vs Web of Science*. Library Journal.com. (Available at: <http://www.libraryjournal.com/article/CA491154.html%22>) (Accessed: July, 2008).
- MEHO L. I., YANG, K. (2007), Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar, *Journal of the American Society for Information Science and Technology*, 58 (13) : 2105–2115.
- MEHO L. I., ROGERS Y. (2008), Citation counting, citation ranking, and h-index of human-computer interaction researchers: A comparison of Scopus and Web of Science, *Journal of the American Society for Information Science and Technology*, 59 (11) : 1711–1726.
- MOYA-ANEGÓN, F., CHINCHILLA-RODRÍGUEZ, Z., VARGAS-QUESADA, B., CORERA-ÁLVAREZ, E., MUÑOZ-FERNANDÉZ, F. J., GONZÁLEZ-MOLINA, A., HERRERO-SOLANA, V. (2007), Coverage analysis of Scopus: A journal metric approach, *Scientometrics*, 73 (1) : 53–78.
- NORRIS, M., OPPENHEIM, C. (2007), Comparing alternatives to the Web of Science for coverage of the social sciences' literature, *Journal of Informetrics*, 1 (2) : 161–169.
- NOUWS, H., ALBERGARIA, J. T., VIEIRA, E. S., DELERUE-MATOS, C., GOMES, J. A. N. F., *Documentos Indexados no ISI, Web of Knowledge, 2000-2007*, Nota Técnica Research Metrics nº 1, 14. Jan.2008, in <http://www.requimte.pt/metrics> and <http://www.fc.up.pt/pessoas/jfgomes>
- SCOPUS FAQs (2008), <http://www.info.scopus.com>, (Accessed July, 2008).
- SCOPUS INFO (2008), [http://www.info.scopus.com/detail/what/documents/title\\_list.xls](http://www.info.scopus.com/detail/what/documents/title_list.xls), (Accessed July, 2008).
- THOMSON REUTERS, (2008A), *Web of Science*, (Available at: [http://isiwebofknowledge.com/media/pdf/WoSFS\\_08\\_7050-1.pdf](http://isiwebofknowledge.com/media/pdf/WoSFS_08_7050-1.pdf)) (Accessed July, 2008).
- THOMSON REUTERS (2008B), *ISI Web of Knowledge*. (Available at: [http://scientific.thomson.com/media/scpdf/wok\\_multidisc\\_fs.pdf](http://scientific.thomson.com/media/scpdf/wok_multidisc_fs.pdf)) (Accessed July, 2008).
- UNIVERSIA, *Ranking Iberoamericano de Instituciones de Investigación*, <http://investigacion.universia.net/isi/isi.html>, (Accessed July, 2008).
- VAN RAAN, A. F. J. (2008), *Measuring Academia Research Performance with Advanced Bibliometric Methods, Center for Science and Technology Studies (CWTS)*, Leiden University, <http://www.unl.pt/noticiastodas/ApresentacaovanRaan.pdf> (Accessed July, 2008).
- VAUGHAN, L., SHAWB, D. (2008), A new look at evidence of scholarly citation in citation indexes and from web sources, *Scientometrics*, 74 (2) : 317–330.