Scientific Production in Computer Science: A comparative study between Brazil and other countries

Jacques Wainer[1*] Eduardo C. Xavier[1] Fábio Bezerra[1,2]
[1]Instituto de Computação - UNICAMP
Caixa Postal 6176
13084-971 – Campinas, SP, Brazil
[2]Centro de Ciências Exatas e Tecnologia – UNAMA
Av. Alcindo Cacela, 287
66060-902 – Belém, PA, Brazil

Abstract

In this paper we present a study about scientific production in Computer Science in Brazil and several other countries, as measured by the number of articles in journals and conference proceedings indexed by ISI and by Scopus. We compare the Brazilian production from 2001 to 2005 with some Latin American, Latin European, BRIC (Brazil, Russia, India, China), and other relevant countries (South Korea, Australia and USA). We also classify and compare these countries according to the ratio of publications in journals and conferences (the ones indexed by the two services).

The results show that Brazil has by far the largest production among Latin American countries, has a production about one third of Spain's, one fourth of Italy's, and about the same as India and Russia. The growth in Brazilian publications during the period places the country in the mid-range group and the distribution of Brazilian production according to impact factor is similar to most countries.

Key Words: Computer Science Scientific Production, Bibliometrics, Cross country comparison.

1 Introduction

There has been some research on the Brazilian scientific production.[1–3] There are also studies of Brazilian scientific production/productivity in specific areas such Psychiatry,[4,5] Life Sciences[6] and

*Corresponding author: wainer@ic.unicamp.br

Ecology.[7] The goal of this paper is to study Brazilian production in the particular scientific area of Computer Science, in the period from 2001 to 2005.

Bibliometric studies in Computer Science (CS) are uncommon because the standard methodology for such studies, using data from the *Institute for Scientific Information* (ISI) is not appropriate, as we will discuss below. In fact most bibliometric studies in CS are on the analysis of social networks of researchers, specially within subareas of CS, such as hypertext,[8] reverse engineering,[9] CSCW,[10] databases[11,12] among others.

1.1 Scientific Databases in Computer Science

In most scientific areas, the journals indexed by the ISI are considered the most prestigious ones and are the *de facto* standard of quality. ISI has a rigorous procedure to accept a new journal in its index, which includes formal aspects such as age of the journal, regularity in its publishing frequency, and impact in the science in general measured by the impact factor of the journal (but the measure of impact factor is not without criticism[13,14]).

But within Computer Science, the ISI is not an agreed upon standard. There are three distinct reasons. In CS, not all scientific production is published in journals, but also in conferences and workshops. The CS community has a strong respect for work published in some conferences, and has a long tradition of creating workshops to discuss cutting edge ideas and technologies. ISI does not index conference proceedings, with the exception of those published in Springer's Lecture Notes in Computer Science series (LNCS) (including the LNAI and LNBI subseries), and the ASIST conference. Furthermore, the conferences that publish their proceedings in the LNCS series are not an unbiased sample of CS conferences. Depending on the CS subarea, the most prestigious conferences are either associated with some professional association such as ACM (Association for Computer Machinery) or IEEE (Institute of Electrical and Electronics Engineers), or they are associated with an international/regional organizations whose sole purpose is the organization of the conference. LNCS, in general, publishes proceedings of regional conferences (for example the Brazilian AI conference), or restricted topic workshops.

Second, there are very well known and respected CS journals that are not yet in the ISI, for example ACM Journal on Experimental Algorithms, ACM Transactions on Algorithms, Journal of Discrete Algorithms, just to list a few in the area of algorithms. And finally, not all of the journals in the ISI Computer Science category are considered by most of the community as within the CS large area.

Scopus is a scientific database that probably provides a better picture of the CS production, since it indexes not only the whole ACM, IEEE, and Elsevier sets of journals, but also some ACM and IEEE conferences (it is not clear which ones and why). But Scopus also includes in its CS lists, journals that the community does not consider as realy CS journals. And the LNCS series is indexed in Scopus as both a journal and as conference proceedings.

There have been some papers comparing ISI, Scopus, and Google Scholar.[15–18] The papers compare coverage, bias on the coverage, and specially, the difference in results on citation analysis.

1.2 Scope of Study

This paper uses the four data sources to compare the Brazilian Computer Science production with that of the following countries: *Argentina, Australia, Chile, China, India, Italy, Mexico, Portugal, Russia, Spain, South Korea,* and *USA*.

We compare this countries according to the following measures: the total number of CS publications; the distribution of publications according to research sub-areas; the distribution of publications according to the impact factor of journals; and the relation between conference and journal publications.

The selected countries have different relations to Brazil. Argentina, Chile and Mexico are the other Latin American countries with some production in Computer Science. Italy, Portugal, and Spain are non-central Latin European countries, and they seem to share with Brazil a similar Latin culture and similar difficulties to write in English. China, India, and Russia, together with Brazil form the BRIC countries, countries that are believed to be strong emergent economic powers. Despite a large difference in culture, South Korea is used in Brazilian science and technology discussions as a prototypical example of a country that followed different government policies for industry and education. In the middle of the 80s, Brazil and South Korea had similar gross domestic product per capita, proportion of the the economy due to agriculture, proportion of the youth in college, and other demographic and economic indicators.[19,20] Australia is an example of a non-central Anglo-Saxon country, for which language is not a hinder to scientific publication. Finally, the US is the most prolific country in Computer Science publications.

The first of the four data sources analysed in this paper is called "ISI journals," and includes the ISI CS references **without** the LNCS series and ASIST proceedings (352 CS journals according to JCR 2005). The second source "ISI conferences" is the ISI data for the LNCS series and ASIST conference proceedings. "ISI total" is the sum of the ISI journal and conference data.

The third data source ("Scopus journals") is the Scopus Computer Science journal references, which includes in February of 2008 454 journals. The forth data source ("Scopus conferences") are the Scopus CS conference proceedings references (130 conferences in February 2008).

The "total" CS production, for this paper, is the result of adding all four data sources. It is important to notice that it by no means represents the set of all Computer Science references, and that there are many duplications - there are journal references in both the ISI and the Scopus data sets, and at least one triplication - LNCS is in the ISI conference, the Scopus journal, and the Scopus conferences data sources.

All data collected refer to the period from 2001 to 2005, inclusive.

1.3 Organization

The remainder of this paper is organized as follows: Section 2 describes the methodology used to collect the data; Sections 3 and 4 describes the results on number of journal publications, and conference publications. Section 5 compares Brazil with different sets of countries, such as Latin American countries, Latin European countries, the BRIC (Brazil, Russia, India, and China), and other countries. Finally, Section 6 discusses the limits of this work, summarizes some important conclusions and points some possible explanations to the results.

2 Method

The ISI data was obtained from the ISI site in November 2006, using the following procedure:

- in the JCR site, we obtained all journals classified in the Computer Science subject area, that is, all journals in the subject areas:
  - COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE;
  - COMPUTER SCIENCE, CYBERNETICS;
  - COMPUTER SCIENCE, HARDWARE & ARCHITECTURE;
  - COMPUTER SCIENCE, INFORMATION SYSTEMS;
  - COMPUTER SCIENCE, INTERDISCIPLINARY APPLICATIONS;
  - COMPUTER SCIENCE, SOFTWARE ENGINEERING;
  - COMPUTER SCIENCE, THEORY & METHODS

  and removed the LNCS series from the list.

- in the Web of Science site, we selected the advanced search and entered the queries of the form:

  `PY=2001 AND CU=BRAZIL AND`
  `(SO=(`*journal1*`) OR SO=(`*journal2*`) OR`...)

  which selects the publications from Brazil for the year 2001, and where *journal1*, *journal2*, and so on, are journals from the list above. The query cannot contain more than 50 clauses, thus only a subset of the journals can be entered in each query.

- We repeated the query with LNCS and ASIST conference proceedings as the *journal1* and *journal2* entries above to obtain the ISI conference data.

The Scopus data was obtained from the Scopus site using the advanced search option. The query:

```
SUBJAREA(COMP) AND AFFILCOUNTRY(brazil) AND PUBYEAR IS 2001
```

returns the list of all references which has one author whose affiliation country is Brazil, in the subject area of computer science, and published in 2001. The Document Type summary in the result page listed the number of references classified as "Articles", "Conference Proceedings", and other types. The first count was used as the Scopus journal data for Brazil for 2001. The second, the Scopus conferences data. The queries were performed in February 2008.

3 Journal Publications

Table 1: Publications for each year from ISI without LNCS and ASIST

Country	2001	2002	2003	2004	2005	Total	Yearly Growth
Brazil	212	215	240	283	292	1242	8.5%
Argentina	50	52	42	50	68	262	10.0%
Australia	579	510	562	597	681	2929	4.6%
Chile	35	37	44	37	48	201	9.6%
China	978	1106	1459	1756	1883	7182	18.1%
India	319	345	364	438	440	1906	8.6%
Italy	880	961	1035	1019	1072	4967	5.1%
Mexico	64	82	87	95	112	440	15.3%
Portugal	100	104	146	178	175	703	16.2%
Russia	285	284	256	277	269	1371	-1.2%
South Korea	574	631	724	825	797	3551	8.8%
Spain	573	678	754	807	918	3730	12.6%
USA	7310	7189	7798	8038	8462	38797	3.8%

The results for the ISI indexed journals can be seen in Table 1. The "Total" column adds the number of publications in each of the years. The "Yearly Growth" column displays the average yearly relative growth. The relative growth for each year is calculated as the difference between the production in year y and in year $y-1$ divided by the latter. The average yearly relative growth is the average of the relative growth for 2002, 2003, 2004, and 2005.

Table 2 reports the evolution of number of publications in the Scopus journal set, the total number of publications to the period, and the average yearly relative growth.

The distribution of the number of publications among the countries when considering the Scopus data is very similar to the distribution of the ISI data. For instance USA publishes almost half of the total publications in both sources and is followed by China.

Table 2: Publications for each year from Scopus journals

Country	2001	2002	2003	2004	2005	Total	Yearly Growth
Brazil	227	227	396	592	346	1788	20.6 %
Argentina	66	59	78	84	90	377	9.1%
Australia	600	571	876	1099	741	3887	10.4%
Chile	30	35	56	87	54	262	23.5%
China	2372	1828	2329	5443	7746	19718	45.1%
India	426	442	571	890	694	3023	16.7%
Italy	1076	1013	1570	1697	1255	6611	7.8
Mexico	94	88	183	219	183	767	26.2%
Portugal	106	107	262	280	193	948	30.4%
Russia	542	492	598	510	465	2607	-2.8%
South Korea	745	817	1630	2422	1252	6866	27.4%
Spain	656	658	1431	1597	928	5270	21.9%
USA	7453	6861	8671	9094	8089	40168	3.1%

Considering the ISI journals USA is responsible for 57% of the publications, China 10%, each of Australia, Italy, South Korea, and Spain is responsible for 4% to 7% of the publications while each of the other countries are responsible for less then 3% of the publications.

Considering the Scopus journals, USA is responsible for 43% of the publications and China 21%. Again each of Australia, Italy, South Korea, and Spain is responsible for 4% to 7% of the publications and each of the other countries are responsible for no more than 3% of the publications.

3.1 Further analysis of the ISI journal publications

ISI divides Computer Science into 7 subareas: artificial intelligence, cybernetics, hardware and architecture, information systems, interdisciplinary applications, software engineering, and theory and methods, with 79, 18, 44, 83, 83, 79, and 71 journals each subarea respectively. The same journal may be classified into two or more subareas. Table 3 reports the total production from 2001 to 2005 for each country, in these seven subareas of computer science.

The differences among the distribution of research effort in the different subareas of CS, although statistically significant, are not that salient, with a few exceptions. Artificial intelligence research varies from 14% of the total production in the US and Chile, to 28% for Spain. Cybernetics attract a low volume of publications (from 2% to 5%) with the exception of Russia. Hardware and architecture varies from 1-2% of the total publication for Russia and Chile, to 11-12% for China, India, South Korea and the US. Information systems, surprisingly, concentrates a fourth of the South Korean production; interdisciplinary applications is the most common area of publications for most countries, with the exception of Russia. Software engineering is very homogeneous, ranging from a low 11% for Mexico to a high of 18% for South Korea and the US. Finally, theory and methods, with the exception of Russia, also receives a homogeneous distribution of effort, ranging from 10% for South Korea, to 19% for Argentina and Chile.

Table 3: Distribution of the total ISI journal CS production for 2000 to 2005 in the seven subareas of computer science

Country	Art. Int.	Cyber.	HW & Arch	Inf. Sys.	Interd. App.	Soft. Eng.	Theory
Brazil	295 (19%)	36 (2%)	84 (5%)	157 (10%)	473 (30%)	267 (17%)	246 (16%)
Argentina	55 (17%)	16 (5%)	13 (4%)	10 (3%)	123 (39%)	41 (13%)	59 (19%)
Australia	646 (17%)	107 (3%)	258 (7%)	666 (17%)	911 (24%)	654 (17%)	565 (15%)
Chile	33 (14%)	11 (5%)	5 (2%)	26 (11%)	77 (33%)	36 (16%)	43 (19%)
China	1841 (20%)	316 (3%)	1005 (11%)	1346 (14%)	2056 (22%)	1533 (16%)	1210 (13%)
India	445 (18%)	58 (2%)	262 (11%)	303 (12%)	710 (29%)	319 (13%)	330 (14%)
Italy	1153 (18%)	158 (2%)	665 (10%)	842 (13%)	1390 (22%)	984 (15%)	1159 (18%)
Mexico	137 (24%)	35 (6%)	44 (8%)	49 (9%)	152 (27%)	62 (11%)	88 (16%)
Portugal	218 (26%)	17 (2%)	71 (9%)	59 (7%)	251 (30%)	113 (14%)	104 (12%)
Russia	552 (23%)	512 (21%)	28 (1%)	102 (4%)	331 (14%)	294 (12%)	610 (25%)
South Korea	781 (16%)	83 (2%)	595 (12%)	1196 (25%)	821 (17%)	874 (18%)	467 (10%)
Spain	1220 (27%)	154 (3%)	370 (8%)	423 (9%)	1019 (22%)	609 (13%)	753 (17%)
USA	6947 (14%)	1323 (3%)	6273 (12%)	9079 (18%)	11010 (22%)	8912 (18%)	7288 (14%)

ISI also computes the impact factor of each of its indexed journals. The 2005 JCR Impact factors in CS journals range from 7.4 to 0.027, with 0.798 as the median impact factor. Table 4 lists for each country, the proportion of the journal ISI CS production that was published in the journals with impact factor above 0.798 (the median).

The USA and Argentina are the countries with a higher proportion of publications on the top 50% impact factor journal., but the USA and Australia are the countries with higher mean-IF. Australia, Brazil, China, India, Italy, Mexico, and Spain are in a group that publishes around 50% of its papers in the top 50% higher impact factor journals. On the other end, Russia published only 30% in the top higher impact factor journals, whereas South Korea, Portugal and Chile, publish 40%, 42%, and 44% respectively.

Table 4: The proportion of the 2001 to 2005 CS ISI journal publications in the top 50% journals with higher impact factor.

Country	proportion	Country	proportion
Brazil	50.3 %	Mexico	48.9 %
Argentina	61.5 %	Portugal	41.7 %
Australia	51.9 %	Russia	29.2 %
Chile	43.8 %	South Korea	39.7 %
China	48.1 %	Spain	47.6 %
India	50.1 %	USA	62.7 %
Italy	51.8 %		

4 Conference Publications

Table 5 reports on the number of LNCS articles and ASIST conference articles in ISI. The growth column is the average yearly relative growth. It is interesting to notice that the growth of the ISI conferences proceedings are very large for all countries. It is more likely that Springer have been increasing the number of conferences that have their proceedings published in the LNCS series, than to assume that all the countries had a large effort in increasing their conference publications. If we assume that the US production is the most stable of them, then we have to assume that the LNCS series is increasing the number of conferences articles by around 44% each year. In fact, the number of publications from LNCS increased from 3165 in 2001 to 19329 in 2005.

The last column in Table 5 is the ratio between the total number of conferences publications and journal publications of each country. This figure is a first estimate of the effort of each country in publishing in journals and in conferences, given the characteristics of the ISI journals and conferences data sets discussed above.

Table 5: Publications for each year from LNCS and ASIST

Country	2001	2002	2003	2004	2005	Total	Growth	Conf/Jor
Brazil	45	119	227	410	314	1115	78.1%	0.89
Argentina	3	14	20	25	29	91	112.6%	0.34
Australia	106	285	539	697	774	2401	74.6%	0.81
Chile	5	24	37	75	76	217	134.6%	1.07
China	161	221	586	2240	3665	6873	137.1%	0.95
India	38	112	52	351	349	902	178.9%	0.47
Italy	224	466	874	854	1120	3538	56.1%	0.71
Mexico	24	40	112	244	196	616	86.2%	1.4
Portugal	49	75	247	191	222	784	69.0%	1.11
Russia	75	58	175	73	172	553	64.1%	0.4
South Korea	102	316	846	1578	2145	4987	125.0%	1.4
Spain	200	464	1038	1020	1154	3876	66.8%	1.03
USA	1000	1919	3023	3292	3948	13182	44.6%	0.33

Regarding the ISI conference to journal ratio, countries seems to be grouped in three sets. Argentina, India, Russia, and USA with the lowest ratio; Australia, Brazil, and Italy with a ratio below 1; and the other countries, with ratio above 1. The figures for Argentina and Russia seems to be explained by the recent decrease in science funding in these countries. Since journal publication in CS is usually free, and conference publication is not, researchers in these countries would prefer to submit their work to journals, whenever possible. This may also explain India's results. As for the USA, most conferences in that country are not published in the LNCS series. It is likely that American researchers would prefer publishing their results in other conferences.

Table 6 reports the data for the Scopus indexed conferences.

Since Scopus considers a broader range of conferences, one can get a better estimation of the effort of each country in publishing articles in journals or conferences. The last column of Table 6 is the ratio between the total of conferences and journal publications of each country, considering the Scopus data. From this data, one can conclude that most of the countries seems to give similar importance to publishing articles in journals and conferences. Again, Argentina, India, and Russia

Table 6: Publications for each year Scopus conference

Country	2001	2002	2003	2004	2005	Total	Growth	Conf/Jor
Brazil	209	232	305	314	495	1555	25.8%	0.86
Argentina	16	30	15	25	48	134	49.0%	0.35
Australia	392	310	584	538	1070	2894	39.6%	0.74
Chile	13	13	22	39	85	172	66.1%	0.65
China	438	722	2379	1691	4908	10138	113.9 %	0.51
India	88	141	255	189	490	1163	68.6%	0.38
Italy	557	656	1095	1077	1811	5196	37.8 %	0.78
Mexico	47	53	145	152	224	621	59.6%	0.80
Portugal	94	112	206	195	303	910	38.3%	0.95
Russia	76	89	206	151	264	786	49.2%	0.30
South Korea	404	413	924	615	2710	5066	108.3%	0.73
Spain	276	386	647	667	1473	3449	57.9%	0.65
USA	6095	6663	7079	3821	5580	29238	3.9%	0.72

are in a group of their own with a low ratio. It is interesting to notice that the USA is not in the low ratio group. China is also noteworthy; despite large growth, its ratio of conference to journals is around 0.5.

But the figures of Conf/Journal above points out a different problem. There is a rule of thumb (that as far as we know has not been rigorously evaluated) that CS researchers publish from 2 to 3 papers in conferences for each paper in a journal. If this rule is true, and if the Scopus Conference set is a more representative set of CS conferences, the figures should have been higher. We call this problem the *invisible work* problem - it seems that there is a large set of conferences which are not (yet) indexed by Scopus (or ISI for that matter) where CS researchers publish at least a third to half of their work, and this work is not indexed by either service. Thus measures of CS scientific production using either ISI or Scopus underestimate it by at least a third. That may be a problem when such production measures are used to define funding between different scientific areas.

5 Comparison of Brazil and other countries

In this section we give a detailed comparison of Brazilian CS production against other countries. Regarding the Latin American countries considered in this work, Brazil is by far the country that publishes more articles in computer science, with 57% of the published articles considering these four countries. Mexico appears in second followed by Argentina and Chile. As discussed, Argentina places lower emphasis in conference publications and has the higher proportion of published papers in journals with higher impact factor. There are no important differences in subjects, with the exception of a lower proportion of publications in Information Systems in Argentina.

Considering Brazil and other Latin European, we can see that Italy and Spain are the most prolific countries, followed by Brazil and Portugal. But the distribution of effort regarding conference to journal, the distribution of publications in the different classes of impact factor, and in the different subjects seems very similar among the four countries. The exception is Portugal's lower proportion of papers in the higher impact journals.

Regarding Brazil with the BRIC countries, we can see that China has a large production, and in 2005 it published four times the number of articles from India, the second top country. Russia and Brazil have more or less the same number of publications. The ratio conferences to journal are very different. As discussed Russia and India place lower effort in conference proceedings. The ratio for Brazil and China are comparable for the LNCS conferences, but not for the Scopus indexed conferences - China seems to place less effort in publishing in these conferences. Regarding distribution among the different impact factor classes for journals, with the exception that more than 70% of Russia's production is in the lower impact factor journals, the other three countries seem to have a similar distribution of publications. On subjects, Russia also is very different from the others, with a lot of publications in the Cybernetics and Theory categories. China and India seems to have a surprisingly similar distribution on subjects.

Finally we discuss the production of Brazil with the remaining countries considered in this article (USA, South Korea and Australia). The distribution of effort regarding the Scopus conferences and journals is similar among the four countries. The ISI ratio shows that the USA places lower effort in publishing in the LNCS conferences, whereas South Korea places a higher effort for that. The USA have a higher proportion of its papers in higher rated journals, and South Korea has a lower proportion.

Thus, Brazil has a good production considering Latin America, but regarding the other countries, Brazil is not well positioned. Considering the BRIC countries, China has much more publications than any other country, although Brazil has a comparable scientific production with Russia and India. Considering the Ibero-American countries, Brazil is far away from Spain and Italy and considering some developed countries, Brazil is the last positioned country, although not so far from these other countries.

Table 7 summarizes all countries production in the four classes in comparison to Brazil. For exemple, USA has 31.2 times the ISI journal production of Brazil, while Argentina has 0.2. Table 8 ranks the countries according to the average total CS production and average yearly growth.

6 Conclusions

This work has some limitations that must be made explicit. We already discussed the limitations of using the ISI and Scopus data sources in CS. Second, this paper does not deal with **productivity**, at all. Standard measures of productivity would be to divide the production figures above by the

Country	ISI journal	ISI conf.	Scopus journal	Scopus conf.	Total
Argentina	0.2	0.1	0.2	0.1	0.2
Australia	2.4	2.2	2.2	1.9	2.1
Chile	0.2	0.2	0.1	0.1	0.1
China	5.8	6.2	11.0	6.5	7.7
India	1.5	0.8	1.7	0.7	1.2
Italy	4.0	3.2	3.7	3.3	3.6
Mexico	0.4	0.6	0.4	0.4	0.4
Portugal	0.6	0.7	0.5	0.6	0.6
Russia	1.1	0.5	1.5	0.5	0.9
South Korea	2.9	4.5	3.8	3.3	3.6
Spain	2.9	3.5	2.9	2.2	2.9
USA	31.2	11.8	22.5	18.8	21.3

Table 7: Comparison of number of publications of the different countries to Brazil

Total CS production	Growth
USA	China
China	South Korea
South Korea	Mexico
Italy	Chile
Spain	Portugal
Australia	Spain
India	India
Brazil	**Brazil**
Russia	Australia
Portugal	Italy
Mexico	Argentina
Argentina	Russia
Chile	USA

Table 8: Countries ordered by decreasing total production and growth

number of CS researchers, or the amount of research investment for each country. But this data is not publicly available, as far as we could find, for any of the countries. UNESCO[21] collects some statistics on both the total amount of R&D investment for each country (called GERD or Gross Expenditure on Research and Development) and the total number of researchers (or full time equivalent researchers) for each country. But such aggregate data is not useful to evaluate the CS productivity unless one adopt the strong assumption that each of the countries have the same proportion of its researchers working in Computer Science, and that each country assigns the same proportion of its GERD to CS research.

An important point of this paper is to provide some intuitions to measure and if necessary improve the Brazilian Computer Science production. The first comparison is with the USA, and although no meaningful comparison of production can be made, we find that the similarity regarding the efforts in publication on conferences and journals using the Scopus data is encouraging. To further approximate the USA proportions, Brazilian researchers should increase the number of papers published in the ISI indexed journal, in particular the top ranked impact factor journals.

China is a scientific puzzle. The volume and growth of Chinese CS production are surprising. We are accustomed to large figures of Chinese economic growth, but science has a different dynamics than the economy. One cannot create scientists in a few years even with very large investments. It is possible that China is now reaping the benefits of a long term policy of sending CS students to study abroad, specially in the USA. It is also possible that such growth is only possible under a more authoritarian control of the scientists themselves. Another explanation is that the index services are with time, including more Chinese publications in their set of indexed journals. In fact, not only Scopus has some Chinese journals within its list, but there are a few Elsevier journals published in English that seems to have a majority of Chinese editors and authors. If, on one hand, it is interesting to know how Chinese CS achieved such success in terms of publications, how CS research is organized in China, and how the CS researchers overcome the problem of publishing in English, on the other hand, it is unlikely that many of these policies and practices can be adapted to Brazil, given the size and culture differences.

Closer to the Brazilian scenario are the countries of South Korea, Spain, Italy, Portugal, and Australia. We believe that the Brazilian CS community should carefully look into how CS research is organized in these countries, and should search for data that would allow some evaluation of the productivity of CS research in these countries.

Let us start with South Korea. The country exhibits a large production and a large growth, and again, that may be because of some of the reasons used to explain China's combination of production and growth. Regarding production, South Korea has almost three times the Brazilian production in ISI journals, almost four times in Scopus journals. Such differences in production have two extreme explanations: either the number of CS researchers in South Korea is three to four times the Brazilian, and thus researchers in both countries have the same productivity, or the two countries have the same number of researchers and the South Korean researchers are three time as productive as Brazilians. As we mentioned, we have no data on the number of researchers, and regarding South Korea we have no intuition if the number of researchers are three to four times the Brazilian. But the fact that South Korea has also shown a large growth seems to indicate that an increase in productivity is more likely than an increase in the number of researchers.

Our intuitions are that for the other countries, such as Spain, Italy, and Australia, it is unlikely that the size of the CS research community it the best explanation. If indeed CS researchers in these countries have higher productivity than Brazilian CS researchers, it would be very interesting to compare the cultural and organizational conditions that foster this increased productivity. Is the amount of time dedicated to research (as opposed to teaching and administration) in these countries larger than in Brazil? Do researchers in these countries have a better acceptance rate in journals (because of better English writing, better access to editors, better knowledge of what are the hot research topics)? Do Brazilian CS researchers produce more "invisible work" than other countries's? Do researchers in these countries have a more competitive environment, or a more collaborative one? Are international co-authors a factor in the increased productivity? These and other questions are of particular interest if the Brazilian CS community hopes to achieve a production level comparable to these countries.

References

[1] E. M. ALBUQUERQUE, R. SIMOES, A. BAESSA, B. CAMPOLINA, and L. SILVA. A distribuição espacial da produção cientifica e tecnologica brasileira: uma descrição de estatisticas de produção local de patentes e artigos cientificos. *Revista Brasileira de Inovação*, 1(2):225–251, July/December 2002. in Portuguese.

[2] W. GLANZEL, J. LETA, and B. THIJS. Science in Brazil. Part 1: A macro-level comparative study. *Scientometrics*, 67(1):67–86, April 2006.

[3] J. LETA, W. GLANZEL, and B. THIJS. Science in Brazil. Part 2: Sectoral and institutional research profiles. *Scientometrics*, 67(1):87–105, March 2006.

[4] J. LETA, R. JACQUES, I. FIGUEIRA, and L. MEIS. Central international visibility of Brazilian psychiatric publications from 1981 to 1995. *Scientometrics*, 50(1):241–254, February 2001.

[5] I. FIGUEIRA, R. JACQUES, and J. LETA. A comparison between domestic and international publications in Brazilian psychiatry. *Scientometrics*, 56(3):317–327, March 2003.

[6] J. LETA, J. C. R. PEREIRA, and H. CHAIMOVICH. The life sciences - the relative contribution of the University of So Paulo to the highest impact factor journals and to those with the largest number of articles, 1980 to 1999. *Scientometrics*, 63(3):599–616, June 2005.

[7] A. S. MELO, L. M. BINI, and P. CARVALHO. Brazilian articles in international journals on Limnology. *Scientometrics*, 67(2):187–199, May 2006.

[8] C. CHEN and L. CARR. Trailblazing the literature of hypertext: Author co-citation analysis (1989-1998). In *Proceedings of the ACM Conference on Hypertext*, pages 51–60, 1999.

[9] A. E. HASSAN and R. C. HOLT. The small world of software reverse engineering. *Proceedings of the 11th Working Conference on Reverse Engineering*, pages 278–283, 2004.

[10] D.B. HORN, T.A. FINHOLT, J.P. BIRNHOLTZ, D. MOTWANI, and S. JAYARAMAN. Six degrees of Jonathan Grudin: a social network analysis of the evolution and impact of CSCW research. *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, pages 582–591, 2004.

[11] M.A. NASCIMENTO, J. SANDER, and J. POUND. Analysis of SIGMOD's co-authorship graph. *ACM SIGMOD Record*, 32(3):8–10, 2003.

[12] E. ELMACIOGLU and D. LEE. On six degrees of separation in DBLP-DB and more. *ACM SIGMOD Record*, 34(2):33–40, 2005.

[13] The PLoS Medicine Editors. The Impact Factor Game: It is time to find a better way to assess the scientific literature. *PLoS Medicine*, 3(6):e291, 2006. doi:10.1371/journal.pmed.0030291.

[14] M. ROSSNER, H. EPPS, and E. HILL. Show me the data. *The Journal of Cell Biology*, 179(6):1091–1092, 2007. doi:10.1083/jcb.200711140.

[15] A. A. GOODRUM, K. W. MCCAIN, S. LAWRENCE, and C. L. GILES. Scholarly publishing in the internet age: a citation analysis of computer science literature. *Information Processing & Management*, 37(5):661–675, 2001.

[16] L. I. MEHO and K. YANG. A new era in citation and bibliometric analyses: Web of science, Scopus, and Google Scholar. arXiv:cs/0612132v1, 2006. http://arxiv.org/abs/cs/0612132v1.

[17] J. BAR-ILAN, M. LEVENEB, and A. LIN. Some measures for comparing citation databases. *Journal of Informetrics*, 1(1):26–34, 2007.

[18] P. JACSO. Comparison and analysis of the citedness scores in Web of Science and Google Scholar. In *Proceeding of Digital Libraries: Implementing Strategies and Sharing Experiences*, number 3815 in Lecture Notes in Computer Science, pages 360–369, 2005.

[19] Y.B. PARK, D. ROSS, and R.R. SABOT. *Educational Expansion and the Inequality of Pay in Brazil and Korea*. Center for Development Economics, Williams College, 1995.

[20] E.B. VIOTTI. National learning systems - a new approach on technological change in late industrializing economies and evidences from the cases of Brazil and South Korea. *Technological Forecasting and Social Change*, 69(7):653–680, 2002.

[21] UNESCO Science Report 2005. `http://www.unesco.org/science/psd/publications/science_report2005.shtml`, 2006.