

World Literature on Semantic Web Research: A Scientometric Study Based On Science Citation Index

*R. Prabu, **S. Srinivasaragavan, ***Mr. Jeyapragash

**Research Scholar*

DLISc.,

***Librarian & Head*

DLISc.,

****Assistant Librarian*

DLISc., Bharathidasan University,
Trichy

Correspondence to:

R. Prabu

praburkp@gmail.com

ABSTRACT

Analysis of the scholarly research output of researchers in the field of semantic web, the research papers published by the scholars in the field of semantic web covered in the annual version of Science Citation index database were taken as the prime source for the present study. This paper attempts to highlight quantitatively the growth and development of world literature on semantic web in terms of publication output as per Science Citation Index (1999-2010). During 1999-2010 a total of 428 papers were published by the scientists in the field 'Semantic Web'. Authorship pattern and author productivity are examined to identify the pattern of research contributions in the field of semantic web research productivity. Area-wise research performance is analysed to identify core area of research and the trend of subjects. Further, an attempt is made to measure the performance of researchers and their research concentration in the field of Semantic Web Research. This article is mainly identifying research output on Semantic Web Research Productivity

Introduction

Scientometrics is the branch of science that describes the output traits in terms of organizational research structure, resource inputs and outputs, develops benchmarks to evaluate the quality of information output. The term Scientometrics originated as a Russian term for the application of quantitative methods to the history of science, but its scope and objectives have widened considerably. Scientometrics studies characterize the disciplines using the growth pattern and other attributes. These studies have potential particularly in assessing the emerging disciplines.

Semantic Web

Semantic Web is the emerging web-technology that may pose a great revolution the way in which information is designed and retrieved in the web. This is the concept mainly emphasizes on meaningful and relevant description of products and objects that are appearing in web. The detailed description of the websites, the objects, the people, the products, services, the organization with existing meta-data would support the people to have more meaningful correlation of the search in retrieving the precise information.

Objectives of the Study

The main objective of the study is to present the growth of

world literature on Semantic web and make the quantitative assessment of status of the research by way of analyzing the following features of research output:

- ◆ To identify the annual growth of publications.
- ◆ To identify the geographical distribution of research output
- ◆ To analyze the authorship and collaboration pattern in the publications
- ◆ To analyze the Institution wise research distribution.
- ◆ To assess the document types used by the researcher.
- ◆ To identify the language wise distribution of publications.
- ◆ To test the Bradford's law of scattering in semantic web Research Productivity output.

Materials and Methods

The publication of semantic web is mostly in the form of primary Journals, Notes, Letters, Reviews, Editorial-materials, Meeting-abstracts, Bibliographic-items and Discussions. The research papers published by Researchers in the field of semantic web covered in the annual version of Science Citation index database were taken as the prime source for the present study. The papers published from 1999 to 2010 by the Researchers are accounted totally 428. They were retrieved from SCI

(Web of Science) database, which is considered to be a prime source of data for the present study. The bibliographical details of publications were entered in the catalogue cards. Finally the cards were arranged in different ways with a view to identify the research performance on semantic web Research Productivity

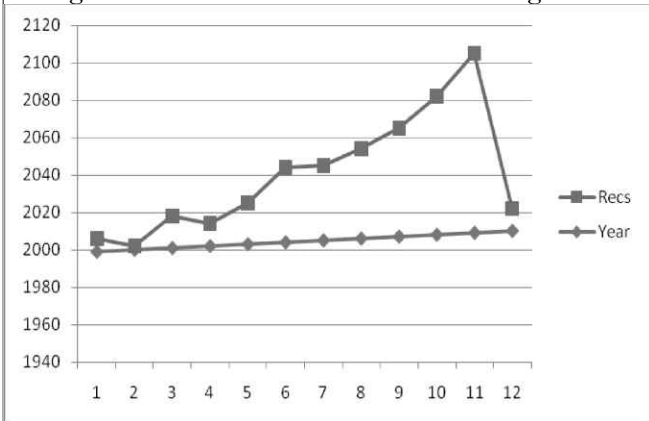
Results and Discussion

1. Annual Growth of publication on Semantic Web Literature

Table 1 could clearly see that during the period 1999-2010 a total of 428 publications were published at International level. The highest publication is 96 in 2009 with 21 Global Citation Scores followed by 74 papers in 2008 with 86 Global Citation Score and 58 papers in 2007 with 140 Global Citation Scores. The lowest publication is 2 in 2000 with 66 Global Citation Scores. It shows that even minimum numbers of records were scored higher global citations. The study also reveals all these 428 publications have 14026 cited references it shows that there is a healthy trend in citing reference is found in semantic web research globally.

Sl.No	Year	Records	%	TLCS	TGCS
1	1999	7	1.6	0	27
2	2000	2	0.5	5	66
3	2001	17	4.0	8	123
4	2002	12	2.8	13	209
5	2003	22	5.1	13	339
6	2004	40	9.3	29	409
7	2005	40	9.3	10	340
8	2006	48	11.2	17	263
9	2007	58	13.6	6	140
10	2008	74	17.3	3	86
11	2009	96	22.4	3	21
12	2010	12	2.8	0	0

Fig. 1 Year wise distribution of literature growth



2. Authorship Pattern

Table 2 indicates ranking of authors by number of publications. Authors Gasevic D and Young CC are published highest number of articles for the study period

with 5 records each. Fergerson RW, Crubezy M, Noy N F having highest global citation score of 133 with just 2 publications each, while Musen MA with 3 publications. Thus the most-cited authors are distinguished from the most-published ones. It is found from the analysis that Lotka's law may not be applicable with regard to author productivity in proliferation of research in semantic web as the research papers equally distributed by a large number of authors.

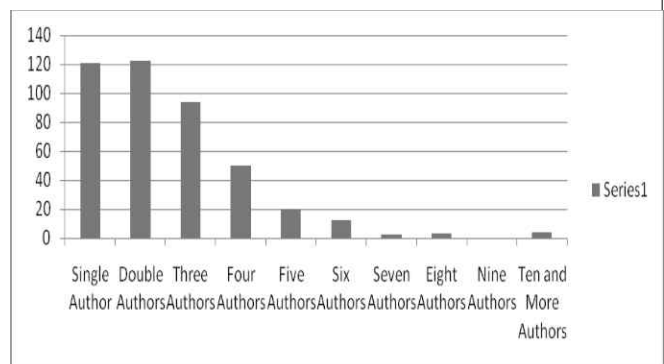
Sl. No.	Author	Records	TLCS	TGCS
1	Gasevic D	5	7	32
2	Yang CC	5	4	29
3	Brooks TA	4	0	1
4	Chi YL	4	0	0
5	Herrera-Viedma E	4	0	0
6	Morales-Del-Castillo JM	4	0	0
7	Motta E	4	3	47
8	Peis E	4	0	0
9	Singh R	4	1	11
10	Storey VC	4	1	26
11	Wei CP	4	1	17

3. Single Vs Multiple Authors

It is found from the study that collaborative research is ensured between the authors in semantic web research as 72% of publications made by joint authors. More than 60% of the research papers are published by 2 to 4 authors. It is also found that the papers published by more than 10 authors.

Sl. No.	Author	Records	%
1	Single Author	121	28.27
2	Double Authors	122	28.50
3	Three Authors	94	21.97
4	Four Authors	50	11.69
5	Five Authors	20	4.67
6	Six Authors	12	2.81
7	Seven Authors	2	0.46
8	Eight Authors	3	0.70
9	Nine Authors	0	0
10	Ten and More Authors	4	0.93
	Total	428	100

Figure: 2 showing collaboration pattern of authors

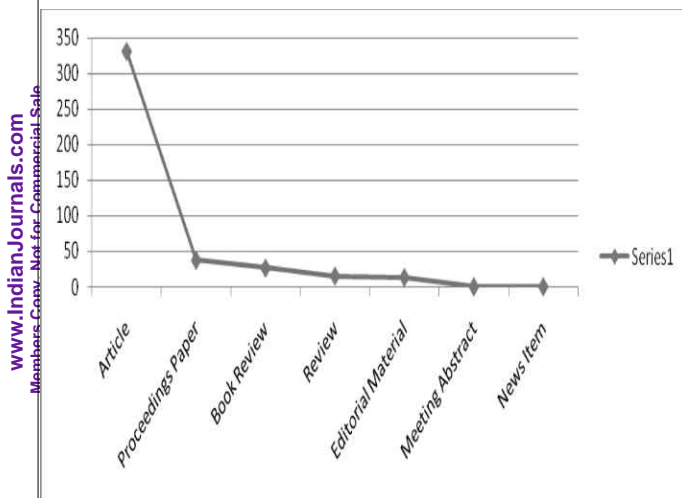


4. Source Wise Distribution of Publications

The study reveals that the major source of publications covered by web of science on semantic web research in journal articles (77.3%), while conference proceedings comprises (8.9%) and book reviews with (6.3%) of the remaining literature.

Sl. No.	Document Type	Records	%	TLCS	TGCS
1	Articles	331	77.3	79	18
2	Proceedings Papers	38	8.9	15	239
3	Book Reviews	27	6.3	0	0
4	Reviews	16	3.7	11	136
5	Editorials	14	3.3	2	28
6	Meeting Abstracts	1	0.2	0	2
7	News Items	1	0.2	0	0

Figure: 3 Document type Wise distributions

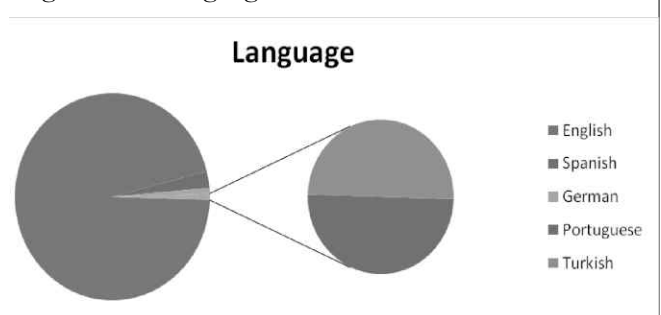


5. Language wise Distribution of Publications

The above table shows the language-wise distribution of research output. It clearly reveals that 409 (95.6%) of the research publications are in English language, while 11 (2.6) articles in Spanish language, 4 (0.9) in German language, 2 (0.5) articles each from Portuguese and Turkish languages. The remaining publications are with other languages like Hungarian, Polish, Dutch, Russian, Spanish, Chinese and Italian etc.

Sl. No.	Language	Records	%	TLCS	TGCS
1	English	409	95.6	107	2022
2	Spanish	11	2.6	0	1
3	German	4	0.9	0	0
4	Portuguese	2	0.5	0	0
5	Turkish	2	0.5	0	0

Figure-4 : Language wise Distribution of Publications



6. Country wise Distribution of Publications

The above table Indicates that among the country wise distribution of semantic web research literature covered by the study United States of America tops with 129(30.1%) articles which is followed by United Kingdom with 55(12.9%) publications, while Spain with 34(7.9%),China with 30(7.0%), Germany with 29(6.8) and Canada with 25(5.8%) research publications respectively. India is one among the top 20 countries in terms of the semantic web research literature and positioned at 19th rank with only 5 publications. USA and UK are having the highest total global citation score of 938 and 332 respectively. Sweden secured 3rd rank in terms of GCS with 154 but with only 6 publications. It seems that there is no significant relation with number of articles and citation scores among the countries produced semantic web research literature.

Sl. No.	Country	Records	%	TLCS	TGCS
1	USA	129	30.1	31	938
2	UK	55	12.9	24	332
3	Spain	34	7.9	8	57
4	China	30	7.0	6	77
5	Germany	29	6.8	6	100
6	Canada	25	5.8	4	43
7	Taiwan	23	5.4	1	83
8	Netherlands	18	4.2	24	139
9	Greece	17	4.0	6	80
10	Italy	17	4.0	3	118
11	South Korea	13	3.0	3	50
12	Australia	11	2.6	0	25
13	Unknown	11	2.6	0	5
14	Singapore	10	2.3	9	66
15	New Zealand	7	1.6	1	14
16	France	6	1.4	2	12
17	Sweden	6	1.4	5	154
18	Austria	5	1.2	1	30
19	India	5	1.2	0	4
20	Ireland	5	1.2	1	9
21	South Africa	5	1.2	0	0

Figure 5 : Country wise Distribution of Publications



5. Surulinathi, M et al. (2008), "Scientometric Dimensions of Knowledge Management Research in India: A Study based on Scopus database", Sri Lankan Journal of Librarianship and Information Management, Vol. 2 No. 2, PP. 13-24.

Conclusion

The scientific study on semantic web based on Science Citation Index database shows that USA is the major producer of scientific output with 129 authorships to its credit in this field also it clearly reveals that 409 (95.6%) of the research publications are in English language. More than 77 percent of the publications were published in the journals with high impact factors is suggestive of the publication behavior of scientists who preferred to publish their papers in highly reputed journals. Collaborative research is ensured between the authors in semantic web research as 72% of publications made by joint authors. The highest publication is 96 in 2009 with 21 Global Citation Scores followed by 74 papers in 2008 with 86 Global Citation Score and 58 papers in 2007 with 140 Global Citation Scores. As the Semantic Web is gaining momentum in taking the world wide web and internet in to next generation web, it is appropriate to analyze the research literature in semantic web by various subject areas, institutions, countries and the contributions of authors.

REFERENCES

1. Angadi, M. et al., (2006), "Publication productivity of Tata Institute of Social Sciences: a scientometric study", *Journal of Information Management*, Vol. 43, No. 4, pp. 363-374.
2. Cahlik, T. and Jioina, M. (1996), "Scientometric analysis of artificial neural networks scientific field", *Neural Network World*, Vol. 6, No. 5, pp. 847-860.
3. Gupta, B.M. and Dhawan, S.M. (2005), "Computer science research in India: a scientometric analysis of research output during the period 1994-2001", *DESIDOC Bulletin of Information Technology*, Vol. 25, No. 1, pp. 3-11.
4. Kumaravel, J.P.S. (2005), "Economic growth, population and research output: a scientometric analysis of genetic engineering, 1988-2000", *IASLIC Bulletin*, Vol. 50, No. 1, pp. 52-59.