SCIENTOMETRIC ANALYSIS OF BIOINFORMATICS LITERATURE DURING 2008-2017

P.Veeramuthu *   C. Baskaran **

*   Research Scholar, Alagappa University, Karaikudi, Tamil Nadu, India

**Librarian & Research Supervisor, Dept. of Library & Information Science, Alagappa University, Karaikudi, Tamil Nadu, India

QR Code

Abstract - *The study analysis the Bioinformatics Literature during 2008-2017. For this study a total of 93460 publications were analyzed. This article evaluated 10 years of bioinformatics publications with the aid of scientometric tools to find out the Year wise distribution, Relative Growth Rate, Doubling Time, Annual Growth Rate, Exponential Growth Rate, Document Type, Country wise Distribution and Language wise Distribution. The findings revealed that a maximum of 14675 publications were published in 2017. There is a decreasing trend in the Year wise Relative Growth Rate. The year 2011 shows RGR as 0.63. Since then it has decreased up to the year 2017. At the same time the Doubling time has increased. There is an decreasing growth rate found in Annual Growth Rate. The annual growth rate of publications over 10 years ranging between 88.54 and 18.63. It found that overall average exponential growth rate was 1.08 during the study period. In the Document type, Journal articles occupied the first position which contributed 57631 records. In the country wise Distribution, China has topped the list with 28266 records followed by United States with 24718 Publications. Majority of the publications (90704) were published in English language only.*

Keywords: Scientometrics, Bioinformatics, Year wise Distribution, Relative Growth Rate, Doubling Time, Document Type

Introduction

Bioinformatics is the application of computer technology to the management and analysis of biological data. Tan interdisciplinary subject which develops methods for storing, organizing, retrieving and evaluating biological data. Databases and information systems are used to store and organize biological data. Analyzing biological data may involve algorithms in artificial intelligence, image processing, data mining, and soft computing. The algorithms in turn depend on theoretical foundations such as discrete mathematics, control theory, system theory, information theory, and statistics. Bioinformatics uses many subjects such as computer science, statistics, mathematics and engineering to process biological data. The most important activity in bioinformatics is to develop software tools to generate useful biological knowledge. Bioinformatics is a different science from biological computation, the latter being a subfield of computer engineering using bioengineering and biology to build biological computers, whereas bioinformatics simply uses computers to better understand biology. Bioinformatics is similar to computational biology

and has similar aims to it but differs on scale: whereas bioinformatics works with basic biological data (e.g. DNA bases), i.e. it works on the small scale paying attention to details, computational biology is a subfield of computer science which builds large scale general theoretical models of biological systems seeking to expand our understanding of them from an abstract point of view, just as mathematical biology does with mathematical models. Bioinformatics as a science can provide input to all previously mentioned scientific fields, as the recording and processing of detailed biological data is the first step towards doing something with them.

Review of Literature

Baskaran, C. (2016)[1] examined the relative growth rate and doubling time of Bioinformatics Publication during 1999-2013. The mean relative growth was measures and doubling time observed from the analysis. Total number 20577 of records on bioinformatics publication during the study. The Maximum of Publication 2234 in 2012 was published compare to rest of the years. The highest publication published in Bioinformatics journal and Harvard University scientists contributed highest number of publication in the study. RGR and DT is exhibits that fluctuating trend happening whole period of study.

Baskaran, C. (2015)[2] analysed the enzymes publications total number of records found 4962

from 1999-2013.Relative Growth rate (RGR) and doubling time of publication were found RGR has been increasing from 2001, 2002 (0.001) to 2013 (0.023). This study and it confront the publications output trend among USA scientists, Wang Y has secured top level as measured 0.226%. USA scientists have contributed totally 15832 (30.815%) items and include 87.947% percent are appeared as journal articles. Harvard University scientists are much attention in produced large number of research papers and they hold top level among research collaboration in enzyme research.

Baskaran,C & Karuilancheran,C. (2015)[3] conducted a bibliometric study in Diabetes and Allied Diseases literature during the year 1995-2013. The study concentrated on Activity Index and Lotka's Law Application with Diabetes and Allied Diseases in India. In this research, the calculated values of Maximum Likelihood Estimator, n and k are 0.24, 2.66 and 0.78 respectively.

Baskaran, C & Sivakami, N. (2014)[4] conducted a bibliometric analysis on Swine influenza research output. In this study, a total of 2360 articles indexed in Pubmed database were taken for analysis. The study concentrated the publication frequency, country wise and institution wise Distributions. Findings revealed that majority of the research papers were contributed by multiple authors.

Ganganna (2017)[5] carried out an analysis of the journal of Academic Librarianship during 2012 – 2016. It was found that maximum number of the Articles published in the year 2014 and minimum number of the Articles published in the year 2012. In the authorship pattern, maximum contribution was from two authors.

Kuri Ramesh & Aadin Tayappa (2016)[6] evaluated the international journal of information Dissemination and technology during 2011-2015. The study analyzed 254 articles with the tool of bibliometrics. The average numbers of author per paper was 01.86. The degree of the collaboration was 0.59%.

Sudharani, Y and Nagaraju, K. (2013)[7] conducted a scientometric study on the research publications from Webology as the source journal during 2004 to 2012. Majority of the publications were contributed by single author 55(54.55%) followed by two authors 32 (31.68%). The degree of collaboration was 0.45. Among the country wise distribution, the majority of the contributions were from both India and Iran 11 (10.89%).

Viswanathan & Tamizhchelvan (2016)[8] analysed the authorship pattern of spacecrafts Research output during the year 2000-2014. In this study, a total of 2, 58,861 articles were analyzed. The average number of the authors per publication was 4.14. Research output. International Journal of Library and Information studies. 6 (2), 1 – 07.

Senthil Kumar PA and Baskaran, C. (2018). Scientometric Analysis of Synthetic Fiber Literature. Library Philosophy and Practice (e-journal).
2204.https://digitalcommons.unl.edu/libphilprac/2204.

METHODOLOGY

Bioinformatics literature Data from the year 2008-2017 were Down loaded from SCOPUS online database. The collected data were analysed with Open Office Spreadsheet and presented the results in the form of Tables and Graphs.

ANALYSIS AND INTERPRETATIONS

Year wise Distribution of Publications

Table 1 indicates that total of 93460 publications published on Bioinformatics publications during 2008-2017. It has been started from 8158 in 2008 publications. Maximum numbers of articles were published (14675) in 2017.Minimum articles were published in 2010 (6574). The average publication per year is 9346. The publication growth on this bioinformatics which could be found from the table an increasing trend from 2008-2017.

Table 1: Year wise Distribution of Publication in Bioinformatics

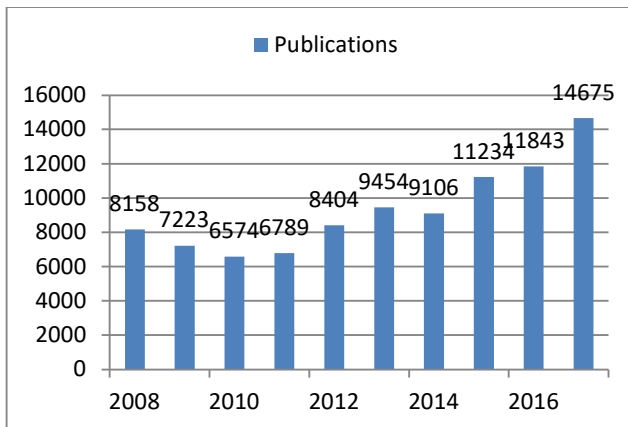| S. No	Year	Publications	Percentage	Cumulative	Cumulative Percentage
1	2008	8158	8.73	8158	8.73
2	2009	7223	7.73	15381	16.46
3	2010	6574	7.03	21955	23.49
4	2011	6789	7.26	28744	30.76
5	2012	8404	8.99	37148	39.75
6	2013	9454	10.12	46602	49.86
7	2014	9106	9.74	55708	59.61
8	2015	11234	12.02	66942	71.63
9	2016	11843	12.67	78785	84.30
10	2017	14675	15.70	93460	100
Total		**93460**	**100**		

Figure 1: Year wise Distribution of Publication in Bioinformatics

Relative Growth Rate and Doubling Time

It can be noted from table 2 that there is a decreasing trend in the Relative Growth Rate yearwise. The year 2009 shows RGR as 0.63. Since then it has decreased upto the year 2017. At the same time the Doubling time (Dt) has increased .The Doubling time was 1.09 in 2009 and it increased to 4.06 in 2017.

Table 2: Relative Growth Rate and Doubling Time

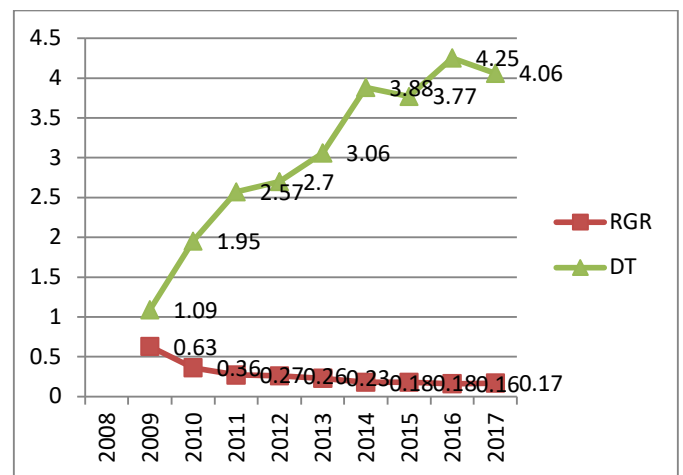| S. No	Year	Publications	Cumulative	W1	W2	RGR	DT
1	2008	8158	8158		9.01		
2	2009	7223	15381	9.01	9.64	0.63	1.09
3	2010	6574	21955	9.64	10.00	0.36	1.95
4	2011	6789	28744	10.00	10.27	0.27	2.57
5	2012	8404	37148	10.27	10.52	0.26	2.70
6	2013	9454	46602	10.52	10.75	0.23	3.06
7	2014	9106	55708	10.75	10.93	0.18	3.88
8	2015	11234	66942	10.93	11.11	0.18	3.77
9	2016	11843	78785	11.11	11.27	0.16	4.25
10	2017	14675	93460	11.27	11.45	0.17	4.06
Total		**93460**					

Figure 2: Relative Growth Rate and Doubling Time

Annual Growth Rate

Annual growth rate is a useful tool to identify trends. The formula used to calculate the annual growth rate uses the previous year as a

base. Table 3 shows the annual growth rate of publications over 10 years ranging between 88.54 and 18.63. There is an decreasing growth rate found in Annual Growth Rate. The maximum annual growth rate was (88.54) in 2011. The minimum Annual Growth Rate was in the year 2018 (17.69).

Table 3: Annual Growth Rate

S. No	Year	Publications	Cumulative	AGR
1	2008	8158	8158	
2	2009	7223	15381	88.54
3	2010	6574	21955	42.74
4	2011	6789	28744	30.92
5	2012	8404	37148	29.24
6	2013	9454	46602	25.45
7	2014	9106	55708	19.54
8	2015	11234	66942	20.17
9	2016	11843	78785	17.69
10	2017	14675	93460	18.63
Total		**93460**		

Exponential Growth Rate
Table 4: Exponential Growth Rate

S. No	Year	Publications	Exponential Growth Rate
1	2008	8158	
2	2009	7223	0.89
3	2010	6574	0.91
4	2011	6789	1.03
5	2012	8404	1.24
6	2013	9454	1.12
7	2014	9106	0.96
8	2015	11234	1.23
9	2016	11843	1.05

10	2017	14675	1.24
Total		**93460**	

Table 4 shows an exponential growth in the number of publications and this was observed during the period 2008-2017. The highest growth rate of 1.24 was found in the year 2012 (8404) and 2017 (14675). The lowest Growth rate of 0.89 was found in the year 2009 with 7223 publications. It found that overall average exponential growth rate was 1.08 during the study period.

Document Type wise Distribution
Table 5: Document Type in Bioinformatics

Document Type	Publications	Percentage
Article	57631	61.66
Conference Paper	22812	24.41
Review	5402	5.78
Editorial	2743	2.93
Book Chapter	2303	2.46
Note	679	0.73
Short Survey	371	0.40
Letter	359	0.38
Book	353	0.38
Erratum	315	0.34
Conference Review	280	0.30
Data Paper	51	0.05
Retracted	21	0.02
Undefined	140	0.15
Total	**93460**	**100**

Document type Distribution of Research output in the field of Bioinformatics Table 5 indicates the document wise distributions of publications on Bioinformatics. The data from the table reveals that most prevalent form of publication is journal article which were 57631

records. Article as a source of publication (57631 publications) take the first place, followed by Conference paper (22812 publications) and Review with (5402 Publications).

Country wise Distribution of Publications

The table 6 reveals that China has topped the list with 28266 records followed by United States with 24718 Publications, United Kingdom with 6135, Germany with 5528 records and India with 4599   publications during the period of study.

Table 6: Country wise distribution in Bioinformatics

S. No	Country	Publications	Rank
1	China	28266	1
2	United States	24718	2
3	United Kingdom	6135	3
4	Germany	5528	4
5	India	4599	5
6	Italy	3905	6
7	Canada	3630	7
8	France	3576	8
9	Spain	3100	9
10	Australia	2891	10
	Others	7112	

Language wise Distribution

Table 7 shows the language wise distribution of publications. Majority i.e., 90704 of the publications in bioinformatics research literature were written in English language. Much of the non English publications were written in Chinese 2174 records, trailed by Russian 133 articles followed by Spanish, Japanese, French,

German,Turkish, Persian and Portugese language so on.

Table 7 : Language wise Distribution of Publications in Bioinformatics

Language	Publications	Percentage
English	90704	97.05
Chinese	2174	2.33
Russian	133	0.14
Spanish	113	0.12
Japanese	51	0.05
French	48	0.05
German	47	0.05
Turkish	45	0.05
Persian	37	0.04
Portuguese	31	0.03
Korean	12	0.01
Polish	12	0.01
Czech	11	0.01
Italian	9	0.01
Hungarian	7	0.01
Arabic	4	0.00
Greek	4	0.00
Slovenian	4	0.00
Hebrew	3	0.00
Croatian	2	0.00
Ukrainian	2	0.00
Azerbaijani	1	0.00
Danish	1	0.00
Serbian	1	0.00
Undefined	4	0.00
Total	**93460**	**100**

Conclusion

The study analyzed the Bioinformatics Literature during 2008-2017. For this research, a total of 93460 publications indexed in Scopus were analyses with the tools of Scientometrics .The study covered 10 years only. The findings revealed that a maximum of 14675 publications were published in 2017. There is a decreasing trend in the Year wise Relative Growth Rate. The year 2009 shows RGR as 0.63. Since then it has decreased up to the year 2017. At the same time the Doubling time has increased. There is an decreasing growth rate found in Annual Growth Rate. The annual growth rate of publications over 10 years ranging between 88.54 and 18.63. It found that overall average exponential growth rate was 1.08 during the study period. In the Document type, Journal articles occupied the first position which contributed 57631 records. In the country wise Distribution, China has topped the list with 28266 records followed by United States with 24718 Publications. Majority of the publications (90704) were published in English language only.

REFERENCE

1. Baskaran, C. (2016). "A Scientometric study on Bioinformatics literature during 1999-2013." *International Journal of Library Science and Information Management* 2(4): 62-71.

2. Baskaran, C. (2015). Research productivity of enzymes literature: A Scientometric study. *International Journal of Library Science and Information Management (IJLSIM)*, *1*(2), 17-25.

3. Baskaran,C & Karuilancheran,C. (2015). Activity Index and Lotka's Law Application with Diabetes and Allied Diseases in India During 1995–2013, SRELS Journal of Information Management, 52(6).

4. Baskaran, C & Sivakami, N. (2014). Swine influenza research output: a bibliometric analysis,- SRELS Journal of Information Management, 51(1): 13-20

5. Ganganna C. (2017). A Bibliometric Analysis of the Journal of Academic Librarianship for the period of 2012 – 2016. International Journal of Library and Information studies. 7(2), 322 – 327.

6. Kuri Ramesh & Aadin Tayappa (2016). International Journal of Information Dissemination and Technology (IJIDT) : A Bbliometric Study. International Journal of Library and Information studies. 6 (4), 72 – 82.

7. Sudharani, Y., and Nagaraju, K. (2013). Webology: a bibliometric analysis. E-library science research journal, 2 (1), 1-7.

8. Viswanathan V. & Tamizhchelvan M. (2016). Authorship pattern on spacecrafts Research output. International Journal of Library and Information studies. 6 (2), 1 – 07.