DESIDOC Journal of Library & Information Technology, Vol. 39, No. 1, January 2019, pp. 31-38, DOI : 10.14429/djlit.39.1.13558

Machine Translation Research: A Scientometric Assessment of Global Publications Output during 2007-16

B.M. Gupta[#] and S.M. Dhawan[$,*]

[#]*1173, Sector 15, Panchkula 134 113, India*
[$]*114, Dayanand Vihar, Delhi 110 092, India*
**E-mail: smdhawan@yahoo.com*

ABSTRACT

The present study provides a quantitative and qualitative description of global machine translation research published during 2007-16 and as indexed in Scopus database. The study profiles research in the field on a series of measures, such as publications growth rate, global share, citation impact, share of international collaborative papers and distribution of publications by sub-areas. The study also profiles top contributing countries, organisations and authors in machine translation research on a series of bibliometric indicators. The study further reports characteristics of highly cited papers in the field.

Keywords: Machine translation; Global publications; Scientometrics; Bibliometrics

1. INTRODUCTION

Translation process is an imperative need to address linguistic barriers and attend information inequalities across peoples known to be speaking about 6,500 world languages. For translation to be complete and meaningful, a profound understanding of syntactic, semantic, and morphological features that characterise a language pair forms an important prerequisite. Translations were carried out earlier by human translators manually, but given the inherent limitations (subject expertise and expensive costs), the focus has since shifted to the development of machine translation[1].

Machine translation (MT) is automatic translation through a software program without human assistance. Machine translation has evolved as a sub-field of computational linguistics and language engineering that draws its ideas and techniques from linguistics, computer science, artificial intelligence, translation theory and statistics[2].

There are three major approaches to machine translation: (i) rule-based approach - Here the software employs a set of grammar rules to automatically complete the translation process. These rules are developed by human language experts and programmers who know and understand how to map rules of two languages. In addition, this system also relies on manually built bilingual dictionaries for the purpose; (ii) statistical machine translation - Here the system uses predictive algorithms to teach computer how to translate a text from one language to another. Building algorithms (statistical translation models) is a quick process but the parameters to create an appropriate model stem from analysis of bilingual text corpora, a set of good translations examples. The statistical model is then used to automatically translate the given text for the most probable output. Statistical machine translation models can be organised into sub-groups such as word-based, phrase-based, and hierarchical-based translation models; and (ii) Hybrid machine translation- It combines the features of the above two approaches[4]. Neural machine translation (NMT) has evolved as the newest method of MT, modeled on neural networks found in human brain. The information goes through different "layers of algorithms" which process the input text for output translation. NMT software also learns from its mistakes, picks up on writing styles being used, and uses recurrent neural network (RNN) to maximise translation quality[5]. While it is still early to assess the potential impact of this approach, it is already very clear that this approach improves the output that looks more fluid sufficient to meet end-user expectations.

The machine translation market is advancing at a rapid rate and multinational companies are taking steps to localise content into more languages. The demand for location based content is also rising across various industrial verticals such as e-commerce, electronics, travel, e-commerce and hospitality. That technology aids in providing information to the end users in their native language is a widely accepted fact. Technology is indeed an absolute need of any organisation to gain prominence in the business and reach global machine translation market[3].

In India, several organisations are supporting the integration and deployment of MT tools and programs to reduce the communication barriers. With the launch of the government's "Digital India" initiative, India is poised to grow

Received : 29 August 2018, Revised : 13 November 2018
Accepted : 22 November 2018, Online published : 14 January 2019

significantly in the global IT industry. Digital India initiatives promise to offer ample opportunities to the national and global companies to expand and extend their reach to Indian markets in this area. Many premier academic and research organisations are also engaged in development of MT systems for Indian languages. The Government of India has facilitated the development of MT systems for translation from english to Indian languages and from Indian to Indian languages using different technological approaches under its Technology Development for Indian Languages (TDIL) Program initiated by Ministry of Electronics & Information Technology[6-7].

1.1 Literature Review

The literature on bibliometric studies on machine translation research and related areas in the field is limited to a few publications. Voss and Zhao[8] analysed machine translation research using publications indicators. Dong and Chen[9] analysed publications during 2000-2015, using Web of Science database. The authors reported the current status of research in machine translation in the field in terms of publication trends, geographic distributions, core literature, and the distinctive research areas of machine translation research: theoretical translation studies, translation and interpreting training and descriptive translation studies. Gile[10] traced the origins of bibliometric investigations in translation studies (TS), its history and recent developments, inter alia in China. Zanettin[11], *et al.* investigated how subfields within translation studies were defined and how research interests and foci have gradually shifted over time. The data for the study was sourced Translation Studies Abstracts (TSA) online database. Doorslaer and Gambier[12] examined academic publishing in translation studies, using translation studies bibliography and focused on the geographical spread of translation and interpreting research measured through academic affiliations and links between keyword frequency and journals, as well as languages of publication.

2. OBJECTIVES

The present study examines the quantitative and qualitative performance of machine translation research published during 2007-16, using Scopus database. The study seeks to look at the status of global machine translation on a series of measures: distribution of global publication output of the world, top 10 most productive countries, distribution document type and source publication type, annual and five-year publications growth rate, and national level share of international collaborative publications of select countries. The study also seeks to measure and analyse publications output by sub-areas of machine translation, publication output and citation impact of top 15 global organisations and authors, top 20 significant journals, and characteristics of high cited papers in the field.

3. METHODOLOGY

The publication data on machine global translation research for the present study was derived from the Scopus database (http://www.scopus.com) covering the period 2007-16. "Machine translation" was used as the search term and coupled it with "Keyword tag" or "Article Title tag" or "Source Title tag" to find out global publication output data in the field. This search string was refined to limit the search period to the period 2007-16. The search string was further refined to find country output by coupling 'country name' with "country tag" and accordingly research output of top 10 most productive countries (including India) in machine translation research was determined. The search string was further refined by using analytical tags as prescribed by Scopus database, such as "subject area tag", "country tag", "source title tag", "journal title name" and "affiliation tag", to retrieve data/information on distribution of publications output by subject, collaborating countries, author-wise, organisation-wise and journal-wise, etc. For citation data, citations to publications were also collected from date of publication till 25 November 2017. A few select bibliometric indicators were used to measure the global performance of machine translation research.

KEY (''machine translation'') OR TITLE (''machine translation'') OR SRCTITLE (''machine translation'') AND PUBYEAR > 2006 AND PUBYEAR < 2017))

4. ANALYSIS

Machine translation global research registered 12.35 per cent growth during 2007-16; its annual research output gradually went up from 260 in the year 2007 to 518 publications in 2016, and its 10-year global output cumulated to 5181 publications during the period. Its five-year cumulative output registered 3.10 per cent absolute growth, up from 2551 in 2007-11 to 2630 publications in 2012-16. The 10-year citation impact of machine translation research averaged to 6.03 citations per publication (CPP) during the period; and its five-year citation impact showed a downward trend, declining from 8.95 CPP in 2007-11 to 3.20 CPP in 2012-16 as shown in Table 1.

Table 1. World publication output and citations in machine translation, 2007-16

Publication period	World		
	TP	TC	CPP
2007	260	3960	15.23
2008	526	4776	9.08
2009	561	5297	9.44
2010	659	4998	7.58
2011	545	3451	6.33
2012	590	2710	4.59
2013	483	1785	3.70
2014	579	1843	3.18
2015	460	1558	3.39
2016	518	514	0.99
2007-11	2551	22842	8.95
2012-16	2630	8410	3.20
2007-16	5181	31252	6.03

TP=Total Papers; TC=Total Citations; CPP=Citations Per Paper

Of the total publications output across the world in machine translation, 74.97 per cent (3884) appeared as conference papers, 21.46 per cent (1112) as articles, 1.24 per cent (64) as book chapters, 0.95 per cent (49) as reviews, 0.68 per cent (35) as editorials, 0.27 per cent (14) as conference reviews, 0.14 per

Table 2. Global publication share of top 10 most productive countries in machine translation during 2007-16

Country	Number of papers			Share of papers			TC	CPP	ICP	%ICP	RCI
	2007-11	2012-16	2007-16	2007-11	2012-16	2007-16			2007-16		
USA	611	449	1060	23.95	17.07	20.46	13526	12.76	252	23.77	2.12
China	417	534	951	16.35	20.30	18.36	3316	3.49	154	16.19	0.58
Japan	222	223	445	8.70	8.48	8.59	1420	3.19	106	23.82	0.53
India	107	217	324	4.19	8.25	6.25	749	2.31	42	12.96	0.38
Spain	173	135	308	6.78	5.13	5.94	1223	3.97	99	32.14	0.66
U.K.	126	165	291	4.94	6.27	5.62	2521	8.66	145	49.83	1.44
Germany	129	155	284	5.06	5.89	5.48	2364	8.32	119	41.90	1.38
France	114	113	227	4.47	4.30	4.38	1744	7.68	103	45.37	1.27
Ireland	87	106	193	3.41	4.03	3.73	818	4.24	93	48.19	0.70
Canada	83	87	170	3.25	3.31	3.28	2312	13.60	63	37.06	2.26
Total	2069	2184	4253	81.11	83.04	82.09	29993	7.05	1176	27.65	1.17
World	2551	2630	5181				31252	6.03			
Share of 10 Countries in World Total							4.20				

TP=Total Papers; TC=Total Citations; CPP=Citations Per Paper; HI=h-index; ICP=International Collaborative Papers; RCI=Relative Citation Index

Table 3. Subject-wise breakup of global publications in machine translation during 2007-16

Subject*	Number of papers (TP)			Activity Index		TC	CPP	%TP
	2007-11	2012-16	2007-16	2007-11	2012-16	2007-16		
Computer Science	2041	2142	4183	99.10	100.88	23299	5.57	80.74
Arts & Humanities	756	725	1481	103.67	96.44	10834	7.32	28.59
Social Sciences	440	647	1087	82.21	117.26	11080	10.19	20.98
Mathematics	449	464	913	99.88	100.12	3603	3.95	17.62
Engineering	449	414	863	105.67	94.50	3979	4.61	16.66
Physics & Astronomy	61	51	112	110.62	89.70	795	7.10	2.16
Decision Science	45	35	80	114.24	86.19	246	3.08	1.54
Materials Science	31	28	59	106.71	93.49	435	7.37	1.14
Medicine	15	35	50	60.93	137.90	276	5.52	0.97
World Output	2551	2630	5181					

There is overlapping of literature covered under various subjects, TP=Total Papers; TC=Total Citations; CPP=Citations Per Paper

cent (7) each as books and short surveys, 0.12 per cent (6) as articles in press and 0.02 per cent (1) each as letter, note and miscellaneous.

4.1. Top 10 Most Productive Countries in Machine Translation

Even as machine translation research was conducted across 93 countries during the 10-year period 2007-16, but distribution of research output across contributing countries was highly skewed with select few countries contributing to research in bulk. As many as 39 countries published a low level output of 1-10 papers each, another 34 countries 11-50 papers each, 8 countries 51-100 papers each, 10 countries 101-500 papers each, and 2 counties contributed a high of 901-1060 papers each. The top 10 countries contributed 170 to 1060 publications each in the field during the period under study as shown in Table 2.

The top 10 most productive countries in machine translation research (Table 2) accounted for more than 82.09 per cent global publication share and 95.57 per cent global citation share during the period. Their five-year global publications share increased marginally from 81.11 per cent in 2007-11 to 83.04 per cent in 2012-16. Country-wise, their global publication share ranged widely from 3.28 per cent to

Table 4. Classification of MT literature by type of approaches used during 2007-16

Type of MT	Total papers (TP)			TC	CPP	Growth rate	%TP
	2007-11	2012-16	2007-16	2007-16			
Rule-based MT	250	332	582	4172	7.17	32.8	22.81
Direct MT	85	110	195				
Transfer-based MT	20	15	35				
Interlingua-based MT	5	3	8				
Corpus-based MT (or Data-Driven) MT	842	991	1833	15102	8.24	1.49	71.85
Statistical MT	792	963	1755				
Example-based or Case-based MT	45	38	83				
Hybrid MT	81	133	214	615	2.87	64.20	8.39
Neural MT	37	230	267	5296	19.84	521.62	10.47
			2551				22.81

Table 5. Scientometric profile of top 10 organisations have had registered top productivity and top citation impact in terms of citations per paper and relative citation index

Name of the Organisation	TP	TC	CPP	HI	ICP	%ICP	RCI
Dublin City University, Ireland	153	628	4.10	12	67	43.79	0.68
Carnegie Mellon University, USA	119	287	2.41	18	37	31.09	0.4
Japanese National Institute of Information & Communication Technology, Japan	115	543	4.72	12	30	26.09	0.78
Harbin Institute of Technology, China	99	293	2.96	9	12	12.12	0.49
IBM Thomas J Watson Research Centre, USA	88	555	6.31	13	14	15.91	1.05
Universitat Politecnica de Valencia, Spain	80	290	3.63	8	14	17.5	0.6
Microsoft Research, USA	78	1190	15.26	18	29	37.18	2.53
University if Edinburg, U.K.	69	1231	17.84	18	31	44.93	2.96
Institute for Intercommunication Research, Singapore	62	402	6.48	11	37	59.68	1.08
University of Maryland, USA	62	989	15.95	16	8	12.9	2.65
Rheinisch-Westfalische Technische Hochschule, Aachen, Germany	59	613	10.39	14	22	37.29	1.72
Google Inc, USA	58	1702	29.34	21	14	24.14	4.87
BBN Technologies, Cambridge, USA	49	761	15.53	13	4	8.16	2.58
John Hopkins University, USA	47	866	18.43	16	9	19.15	3.06
Microsoft Research Asia, China	46	431	9.37	11	8	17.39	1.55
University of Southern California, Information Sciences Institute, USA	43	1118	26.00	18	11	25.58	4.31
Columbia University in the City of New York, USA	42	343	8.17	11	11	26.19	1.35

20.46 per cent, with the USA leading with the highest global publications share (20.46 %), followed by China (18.36 %), Japan (8.59 % share), and 6 other countries (from 3.28 % to 6.254 %) during the period. The top 10 countries differed in the growth of their global publication share over time. India, China, U.K., Germany, Ireland, and Canada registered increase in their five-year global publications share by 0.06 per cent to 4.06 per cent. On the other hand, USA, Spain, Japan, and France registered decline in their five-year global publications share by 0.17 per cent to 6.88 per cent during the same period from 2007-11 to 2012-16. Five of top 10 countries scored relative citation index above the group average of 1.17: Canada (2.26), USA (2.12), U.K. (1.44), Germany (1.38) and France (1.27) during 2007-16. The top 10 most productive countries

contributed 12.96 per cent to 49.83 per cent of their national level output as international collaborative papers on machine translation research.

4.2 Subject-Wise Distribution of Research Output

According to the Scopus classification, machine translation research can be distributed across nine sub-fields for comparative performance by subject. Computer science has been found to be the most popular sub-area of research in the domain of machine translation, with highest publications share (80.74 %), followed by arts & humanities (28.59 %), social sciences (20.98 %), and 6 other sub-fields contributing between 0.97 per cent and 17.62 per cent global publications share.

The five-year research activity across 9 sub-fields of machine translation was further inter-compared on an indicator named 'activity index'. The world average activity index of a sub-field in machine translation is taken as 100, with sub-fields showing rise in their research activity as computer science, social sciences, mathematics and medicine and sub-fields showing decline in their research activity as arts & humanities, engineering, physics & astronomy, decision science and materials science from 2007-11 to 2012-16. Social sciences registered the highest citations impact per paper of 10.19 CPP and decision science (3.08) the least during 2007-16 as shown in Table 3.

4.3 Approaches to Machine Translation (MT)

There are different approaches to machine translation. Corpus-based MT accounts for the largest share of papers (71.85 %), followed by rule-based MT (22.81 %), neural MT (10.47 %) and hybrid MT (8.39 %) during 2007-16. In terms of publication growth from 2007-11 to 2012-16, the largest (521.61 %) was observed in case of neural MT, followed by hybrid MT (64.20), rule-based MT (32.8 %) and corpus-based MT (1.49 %). In terms of citation impact per paper, neural MT again registered the largest impact (19.84), followed by corpus-based MT (8.24), rule-based MT (7.17) and hybrid MT (2.87) during 2007-16 as shown in Table 4.

4.4 Profile of Top 50 Most Productive Global Organisations

Machine translation research was conducted by 683 organisations spread across 93 countries during 2007-16. The distribution of research across contributing organisations is highly skewed. Bulk of the participating organisations (539) contributed 1-10 papers each, another 107 organisations contributed 11-30 papers each, 20 organisations 31-50 papers each, 13 organisations 51-100 papers each and 4 organisations 101-154 papers each. The top 50 productive organisations in machine translation research contributed 47.38 per cent (2455) global publication share and 61.39 per cent (19187) global citation share during 2007-16. Their 10-year research output varied from 24 to 153 publications during 2007-16 as shown in Table 5.

- Seventeen of these organisations registered publications output above the group average of 49.1: Dublin City University, Ireland (153 papers), Carnegie Mellon University, USA (119 papers), Japanese National Institute of Information & Communication Technology, Japan (115 papers), Harbin Institute of Technology, China (99 papers), etc. during the period as shown in Table 5.
- Fifteen organisations registered citation impact and relative citation index above the group average of 7.82 citations per publication and 1.15: Google Inc., USA (29.34 and 4.87), University of Montreal, Canada (26.73 and 4.43), University of Southern California, USA (26.0 and 4.31), Stanford University, USA(23.20 and 3.85), John Hopkins University, USA (18.43 and 3.06), etc. during the period (Table 5)
- Twenty three organisations contributed international collaborative publications share above the group average of 29.16 per cent: Xerox Research Centre for Europe, France (70.0 %), Qatar Computing Research Institute (67.86 %), Institute for Intercommunication Research, Singapore (59.68 %), University of Montreal, Canada (50.0 %), Universitat Politecnia de Catalunya, Spain (48.08 %), etc. during the period.

4.5 Profile of Top 50 Most Productive Authors

A total of 2174 authors participated in global research on machine translation research during 2007-16. Of these, 1836 authors contributed 1-5 papers each, 218 authors 6-10 papers each, 120 authors 11-30 papers each, 6 authors 31-50 papers each and 4 authors 51-88 papers each. Top 50 authors contributed 14.29 per cent (2758) global publications share and 31.17 per cent (9742) global citation share during the period. Their research productivity in the field varied from 17 to 83 publications as shown in Table 6.

- Thirteen authors registered publications output above the group average of 28.58: A. Way (83 papers), E. Sumita (&3 papers), Q. Liu (57 papers), F. Casauberta (55 papers), T.Zhao(51 papers) , etc. during the period 2007-16 (Table 6)
- Eighteen authors **r**egistered impact and relative citation index above the group average of 6.18 citations per publication and 1.13: C. Collison-Burch (24.12 and 4.0), P. Koehn (23.41 and 3.88), K. Knight (23.0 and 3.81), R. Resnik(22.63 and 3.75), C. Dyer (19.10 and 3.17), etc. during the period (Table 6)
- Twenty three authors contributed international collaborative publications share above the group average of 26.66 per cent of all authors: M.R. Costa-Jussa (92.0 %), M. Turchi (60.0 %), J. Van Genabith (56.0 %), D. Xiong (55.56 %), Y. Lepage (52.63 %), etc. during the period.

4.6 Medium of Research Communication

Of the total world output in the field of machine translation research, 60.95 per cent (3158) appeared in conference proceedings, 22.33 per cent (1157) in journals, 15.61 per cent (809) in book series, 1.0 per cent (52) as books and 0.10 per cent (5) as trade publications during 2007-16.

A total of 705 journals published in all 1157 research papers. Of these, 293 journals published 1-10 papers each, 6 journals 11-20 papers each, 4 journals 21-30 papers each, 1 journal each published 31-40 papers and 120 papers during 2007-16.

Table 6. Scientometric profiles of top 10 authors ranked top in terms of their research productivity and in terms of their highest citation impact and relative citation index

Name of the	Affiliation of the Author	TP	TC	CPP	HI	ICP	%ICP	RCI
A.Way	Dublin City University, Ireland	83	395	4.76	10	37	44.58	0.79
E. Sumita	National Institute of Information & Communication Technology, Japan	73	284	3.89	9	20	27.4	0.65
Q. Liu	Institute of Computing Technology, CAS, China	57	343	6.02	10	29	50.88	1.0
F. Casauberta	University Politec. De Valencia, Spain	55	217	3.95	8	4	7.27	0.65
T. Zhao	Herbin Institute of Technology, China	51	114	2.24	5	8	15.69	0.37
H. Ney	RWTH, Aachen University, Germany	48	484	10.08	13	13	27.08	1.67
M. Zhou	Microsoft Research Asia, China	43	382	8.88	10	5	11.63	1.47
S. Li	Herbin Institute of Technology, China	42	102	2.43	6	4	9.52	0.4
L. Specia	University of Wolverhampton, U.K.	37	236	6.38	8	17	45.95	1.06
M. Utiyama	National Institute of Information & Communication Technology, Japan	33	192	5.82	8	11	33.33	0.96
M. Federico	FBK-IRST-Ricerca Scientifica e Tecnologica, Povo, Italt	30	302	10.07	7	11	36.67	1.67
P.Koehn	University of Edinburg, U.K.	27	632	23.41	12	13	48.15	3.88
C. Dyer	Carnegie Mellon University, Pittsburg, USA	21	401	19.1	9	10	47.62	3.17
A.Lavie	Carnegie Mellon University, Pittsburg, USA	21	301	14.33	8	5	23.81	2.38
K. Knight	USC, Information Sciences, CA, USA	20	460	23	12	2	10	3.81
R. Resnik	University of Maryland, USA	19	430	22.63	10	1	5.26	3.75
H.T. Nag	National University of Singapore	19	313	16.47	9	5	26.32	2.73
H.Wu	Toshiba (China) R & D Centre, Beijing, China	19	197	10.37	7	1	5.26	1.72
C. Collison-Burch	John Hopkins University, USA	17	410	24.12	9	3	17.65	4

The top 15 most productive journals reported 9 to 120 papers each on machine translation research; and together they accounted for 31.72 per cent (367 papers) share of total researchon machine translation that appeared in journal medium during the period. The top ranking journals are: *Machine Translation* (with 120 papers), followed by *Computer Speech & Language* (35 papers), *IEICE Transactions on Information & Systems* (27 papers*), IEEE Transactions on Audio Speech & Language Processing* (24 papers*), ACM Transactions on Asian Language Information Processing* (23 papers each), etc. during the period as shown in Table 7.

4.7 Highly Cited Papers

A total of 35 papers that received 100 to 100+ citations each since their publication during 2007-15 were identified as highly cited papers in machine translation research. Of these 35 highly cited papers, 29 were in the citation range 101-200 and 6 in citation range 201-300. Together these 35 papers cumulated a total of 5565 citations, averaging 159.0 citations per paper. Of the 35 highly cited papers, 24 were contributed (as non-collaborative papers) by sole organisations and remaining 11 by organisations in joint or multi-lateral collaboration (7 national level collaborative and 4 international level collaborative papers). The USA participated in the largest number of highly cited papers (23), followed by the Canada, Germany and U.K. (3 papers each), France and Italy (2 papers each), China, Greece, Hong Kong and Singapore (1 paper each). These 35 highly cited papers involved the participation of 112 authors from 48 research organisations across the globe.

The leading world organisations contributing to highly cited papers were: Google Inc. USA (4 papers), Carnegie Mellon University, USA, BBN Technologies, Cambridge, USA, University of Southern California, Information Sciences Institute, USA, Stanford University, USA, University of Montreal, Canada and University of Edinburg, U.K. (2 papers each), etc. Of the 35 highly cited papers, 7 were published as articles and 28 as conference papers. Seven journals published 1 highly cited paper each: *ACM Computing Surveys, Artificial Intelligence, IEEE Intelligent Systems, IEEE Software, IEEE Transaction on Computers, International Journal of Robotics*

Research and *Journal of Artificial Intelligence Research.*

5. SUMMARY & CONCLUSION

This paper provides a quantitative and qualitative description of machine research published during 2007-16. The data for the study was sourced from Scopus database covering the period 2007-16. Machine translation research registered a high 12.35 per cent growth, and cumulated 5181 publications during 2007-16. A total of 2174 authors from 683 organisations from 93 countries contributed to the research during the period. At qualitative level, machine translation research averaged a medium level citation impact of 6.03 citations per paper during the period. Only five countries could achieve relative citation index above the world average of 1.17: Canada (2.26), USA (2.12), U.K. (1.44), Germany (1.38) and France (1.27) during 2007-16.2007-16. The global output of highly cited papers in machine translation research has been small and insignificant limited to just 35 papers (accounting for 0.67 % share of world output).

Table 7. Top 10 Most productive journals in machine translation during 2007-16

Source Journal	Number of Papers		
	2007-11	2012-16	2007-16
Machine Translation	61	59	120
Computer Speech & Language	4	31	35
IEICE Transactions on Information & Systems	15	12	27
IEEE Transactions on Audio Speech & Language Processing	15	9	24
ACM Transactions on Asian Language Information Processing	13	10	23
Processamiento De Lenguaje Natural	5	16	21
Natural Language Engineering	6	12	18
Computational Linguists	0	17	17
Journal of Computational Information Systems	9	7	16
Journal of Artificial Intelligence Research	3	11	14
Language Resources & Evaluation	5	9	14
Journal of Information & Computational Science	6	5	11
ACM SIGPLAN Notices	4	5	9
Electronic Notes in Theoretical Science	8	1	9
IEEE ACM Transactions on Audio Speech & Language Processing	0	9	9
Total of 15 journals	154	213	367
Total global journal output	529	628	1157
Share of top 15 journals in global journal output	29.11	33.92	31.72

Computer science is the most sought after sub-area of machine translation research, with (80.74 % share) the highest publications share, followed by arts & humanities (28.59 %), social sciences (20.98 %), , and others. Social sciences, among various subjects registered the highest citations impact per paper of 10.19 CPP, followed by materials science (7.37), arts & humanities (7.32), physics & astronomy (7.10), computer science (5.57) etc. during 2007-16

Conclusion - Analysis of 10-year data on machine translation research across the world reveals that the subject under study is still in an infancy stage of its development, Top 10 most productive countries that dominate the research in the subject across the world contributed 82 per cent global publications share and 95 per cent global citations share. The USA leads the world in machine translation research, accounting for the highest global publications share (20.46 %), followed closely by China (18.36 %), Japan (8.59 % share), and 6 more countries (from 3.28 % to 6.254 %). The top 10 most productive organisations in the subject are from developed world countries including USA, Japan, Ireland and China. The top 10 most productive authors are also from developed countries including USA and Canada. Evidently, machine translation is still an area of research dominated exclusively by developed world countries like USA, Japan, and China. Developing world countries have yet to make their iimpact in this field. Of all the approaches to machine translation software that have evolved over time, it is found that neural machine translation approach has been most popular one across several different sectors including academic, scientific, industrial and defense sectors mainly due to its potential to offer high accuracy and speed in the translated works.

The main problems that India faces in the area of MT software are syntactic and semantic in nature since each Indian language has own distinct structure. It is not easy to capture such grammatical nuances across languages when it comes to software development for machine translation of Indian languages. Nevertheless, MT in India has over the years made a notable progress in the field. In order to catalyse machine translation research, India needs a long-term policy with a view to prioritise R&D areas in MT, identify role of private sector in system development and identify organisations that have major potential to undertake machine translation research. Initiatives on such lines for developing machine translation systems will help: (i) facilitate smooth communication between the Centre and the states so vital for promoting and accelerating research activity; (ii) providee local population grass root information of land, agriculture, health and education in regional languages; and (iii) convert existing manuscripts, books, reports, newspapers, etc., from regional languages to English.

REFERENCES

1. Oladosu, John et al. Approaches to machine translation: A

review. *FUOYE J. Eng. Technol.*, 2016, **1**(1), 120-126.
2. What is machine translation? n.d. https://www.sdltrados.com/solutions/machine-translation/ (Accessed on 22 June 2018)
3. Machine translation market size was over USD 400 million in 2016 with 19.0 % CAGR estimation from 2017 to 2024. https://www.gminsights.com/industry-analysis/machine-translation-market-size (Accesed on 22 June 2018
4. Okpor, M.D. Machine translation approaches: Issues and Challenges. *IJCSI Int. J. Comput. Sci.,* 2014, **5**(2), 159-165.
5. Brownlee, Jason (2017). A gentle introduction to neural machine translation. Natural language processing. https://machinelearningmastery.com/introduction-neural-machine-translation/ (Accessed on 22 June 2018)
6. Final draft standard on machine translation acceptance. Version 4.0. Ministry of Electronics & Information Technology, Govt. of India http://tdil-dc.in/undertaking/article/262568Machine%20Translation%20Acceptance%20Ver4.0.pdf (Accessed on 22 June 2018)
7. TDIL. http://tdil.meity.gov.in/ (Accessed on 22 June 2018)
8. Voss, Stefan & Zhao, Xin. Some steps towards a scientometric analysis of publications in machine translation. Conference: IASTED *In* Conference on Artificial Intelligence and Applications, part of the 23rd Multi-Conference on Applied Informatics, Innsbruck, Austria, February 14-16, 2005
9. Dong, Dahui & Chen, Meng-Lin. Publication trends and co-citation mapping of translation studies between 2000 and 2015. Scientometrics November 2015, **105**(2), 1111–1128
10. Daniel Gile. Analyzing translation studies with scientometric data: From CIRIN to citation analysis. *Perspect. Stud. Transl. Theory Pract.* 2015, **23**(2), 240-48
11. Zanettin, Federico; Saldanha, Gabriela & Harding, Sue-Ann. Sketching landscapes in translation studies: A bibliographic study. Perspectives 2015, **23**(2), 161-182.
12 Doorslaer, Luc van & Gambier, Yves. Measuring relationships in translation studies. On affiliations and keyword frequencies in the translation studies bibliography. Perspectives 2015, **23**(2), 305-319.

CONTRIBUTORS

Dr B.M. Gupta received his BLibSc, Associateship in Documentation, and PhD in Library and Information Science. He is retired as Scientist G and Emeritus Scientist from CSIR- NISTADS, Delhi. He has published more than 200 research papers mainly in the area of scientometrics in journals and conferences.
Contribution in the current study, he conceived the idea, downloaded the publications data and analysed the data and wrote the first draft.

Dr S.M. Dhawan received his MSc (Physics) from Sardar Patel University, MLIS from University of New York, USA, and PhD in Library Science. He is retired as Scientist 'F' from CSIR-National Physical Laboratory, Delhi. He has authored several research papers, research reports covering several areas of library science, library management systems and scientometrics.
Contribution in the current study, he contributed in writing of paper and giving it a final shape.