Fall 10-3-2021

Assessing Q&A Trends in Scholarly Communications: a Quantitative Study of ResearchGate

Subhajit Panda
Chandigarh University, Mohali, Punjab, suvapanda007@gmail.com

Simran Kaur
Panjab University, Chandigarh, India

Heenam Gakhar
Panjab University, Chandigarh, India

Pooja Kumari
Panjab University, Chandigarh, India

Rupak Chakravarty Prof
Panjab University, Chandigarh, India

Assessing Q&A Trends in Scholarly Communications: a Quantitative Study of ResearchGate

Subhajit Panda
Assistant Librarian, Central Library, Chandigarh University, Mohali, Punjab, India
Email: suvapanda007@gmail.com

Simran Kaur
Senior Research Fellow, DLIS, Panjab University, Chandigarh, India
Email: sim199429@gmail.com

Heenam Gakhar
Panjab University, Chandigarh, India
Email: heenamgakhar0111@gmail.com

Pooja Kumari
Panjab University, Chandigarh, India
Email: poojathakur385@gmail.com

Rupak Chakravarty
Professor, DLIS, Panjab University, Chandigarh, India
Email: rupak@pu.ac.in

Abstract

The purpose of the study is to explore and analyse the disciplinary question and answer (Q&A) trends prevailing in the scientific social network (SCN)/scientific collaborative network (SCN) also known as an academic social network (ASN), ResearchGate (RG) and to examine the subject wise status of Q&A, content sharing trends of Q & A on RG and statistical analysis of various tests. The present study adopted a quantitative research design (Correlational as well as explanatory) to pursue the research questions and objectives for conducting the study. The data was collected from the research gate (RG) platform and 35 subjects were selected from the subject area and seven different types of variables were selected from each subject i.e., Posts, Questions, Publications, Articles, Preprints, Conference paper, Literature review. Various tests viz, Descriptive statistics, t-test, normality test and correlation, JASP software were performed for analysis of the data. The findings of the study reveal Findings of the study reveals the subject-wise status in different variables such as Statistics has the maximum number (4575) and Genealogy has the minimum number of posts (48) and so on. On the basis of the total content of the subject areas, the Literature review is (1%) followed by Preprints (5%), Questions (6%), Conference paper (7%), Posts (16%), Articles (65%). Publications possess maximum value while measuring central tendency (mean & median) & dispersion (std. deviation & std. error) and those are the minimum in the case of Literature Review. Publications, Preprints & Conference paper shows normal distribution *(p>0.05)* with low *p*-value, and the data-set for other four variables don't possess a normal distribution *(p<0.05)*. The statistics of the Student's t-test shows a maximum value in the case of 'articles' (viz. 50.19) and the 'p' value of all the variables taken to be <0.001 except literature review (i.e., 0.005). Among the 21 correlation types, 7 possess negative correlation value suggests the presence of no correlation between those variables. Whereas, the other 14 possess a positive correlation value ranges from 0.013 to 0.906 depending upon respective p-values.

1. Introduction:

Asking questions is a basic trait of humans. Queries originating in minds often make one impatient or annoyed. Being inquisitive and skeptical is a good and positive for approach for clearing our doubts and progress towards achieving our goals. This hold good for maintaining the true spirit of research as timely and to-the-point answers to those queries removes the hurdles and obstacles coming in the way of research pursuit. No library in the world can acquire all the resources, no academician or researcher have the knowledge of every single aspect of his/her field. In circumstances when an academician or researcher gets stuck in his/her existing research topic, social Q&A sites become an influential venue to seek information related to their scholarly & research areas and discuss openly with the experts of similar interest. Social Q&A sites facilitate people to share and distribute knowledge, expertise and experience every day (Xiang et al., 2012). The success of these sites can be attributed to the fact that people find it more convenient to have customized answer to their queries instead of going through long lists of answers provided by search engines (Ayoub et al., 2020). Knowledge sharing & effective communication, the backbone of today's academia & research community, have changed their way with the dynamic approach of Social Q&A sites. In addition, Social Q&A sites significantly decrease the manual effort to the traditional mail-list for formally communicate with other academicians and/or researchers.

The present study deals with the Q&A trends among the academia & research community in the Research Gate (RG). Further, the study tries to evaluate the approach of the current age academician & researcher with available dataset in the RG Q&A platform and compare it to the other different approaches. RG holds a reputed and strong position in the domain of 21st century "Academic Social Network" (ASN) or "Scientific Collaboration Network" (SCN) through its robust scholarly communication ecosystem.

2. Knowledge Management & Academic Social Networking Sites (ASNS):

In the current research environment, ASNS are designed to promote scholarly communication in which research collaboration and knowledge sharing play a fundamental role. The KM system integrates both technology & the social environment. The technology helps users to interact with knowledge resources on the internet to transform "individual tacit knowledge" into "collectively shared tacit knowledge", while the information environment assists in successful R&D, open communication, and information access. When using non-generic social media sites like Facebook, Twitter, and others for academic purposes, academics frequently find difficulties in reaching their target audiences. However, ASNS leverages on the characteristics of social networks but also offers additional controls to ensure user comportment (Ovadia, 2014). Therefore, users are frequently and actively engaged with the research-based features than the social-based features (Jeng et al. 2015). Unlike generic online networks, ASNS provides more reliable methods for managing personal profiles, thus users can develop a valid academic identity and also build a professional reputation (Barbour & Marshall 2012). ASNS facilitates in solving their knowledge queries through providing necessary reliable information from experts in the different respective fields and thus harnessing collective wisdom by the users.

KM aims to simplify knowledge retention. Q&A sites are one of the techniques to create dynamic and valuable knowledge bases. Q&A sites currently enabled people to help each other to find solutions to thousands of problems. In these sites, users exchange their expertise through activities like posting questions, answers, or comments and voting on the quality if they find worthy content in these posts. The

academic Q&A sites support the different phases of the KM process from Knowledge capturing to dissemination. These sites contribute to a comprehensive knowledge management system by creating, capturing, and sharing both explicit and tacit knowledge among users most efficiently and reliably.

3.   RG & Its Q&A Platform:

Research Gate (RG) is the professional network for scientists and researchers to connect, share and access scientific output, knowledge, & expertise and lets them make their own research visible. The idea of developing such a platform was conceived in 2008 by physicians Dr. Ijad Madisch and Dr. Soren Hofmayer, and computer scientist Horst Fickenscher (ResearchGate, 2019). Presently, RG has more than 20 million members contributing to over 135 million scholarly publication pages. RG users can also discover scientific knowledge, connect with their peers and even find a relevant job using RG's Job Board from anywhere through the Research Gate iOS app (ResearchGate, 2018). RG Q&A is an effective platform where registered members/users engage with other researchers/scientists on a professional level to share knowledge, connect with specialists, and identify oneself as an expert in a given field of study. RG Q&A has the following options (ResearchGate, n.d.):

a) Asking questions: Users can ask research-related questions by accessing the Q&A overview page and then enter "Ask a technical question" or "Start a discussion". The question may attach with a description and support by publications, images, graphs, or links (optional). This helps to increase the visibility of the question and enhance the number of responses. Question title should be clear, concise, interrogative, informative, and have a character limit.

b) Adding answers: While answering, the responder may attach supporting resources such as publications, files, images, graphs, or links to other content. When an answer is added, the question gets automatically followed. A good answer is a comprehensive and thoughtful response to the original question. Good answers often originate from researchers who can draw on their own research experience.

c) Following and recommending questions: In Research Gate researcher can follow a particular question and recommend both question & answer. The following demonstrates interest while recommending reflects appreciation. So, more followers help to increase the visibility of a question and get addressed by an expert, while the recommendation assists a user to identify the most relevant discussions in a particular field.

d) Finding questions and answers: A user/members can also browse Q&A using the other three filters: Recent questions in a given field, Questions user follow, and Questions user asked.

e) Deleting and editing questions: Questions asked can be edited or deleted at any later stage. Questions may be deleted if they belong to some specified categories like advertisements; quick answered general knowledge questions; broad, vague, or unscientific (e.g., personal stories) questions; job applications; duplicate questions; copyright claimed questions; private messages and general usage queries of RG. To keep Q&A relevant, RG may also edit researcher questions for language and other inconsistencies. For minor edits, such as typos or grammar changes, users are notified on-site and for larger edits, they receive an email.

f) Closed questions: Sometimes, RG prefers to close Q&A discussions that may be considered as a breach of Terms and Service or Community Guidelines of RG, conversation going off-topic, complaints about content shared in the discussion, etc. Users can still read the questions and answers, but they cannot add new responses to the individual Q&A discussions. Thus, the members can still benefit from any valuable ideas or insights exchanged in the closed discussions, rather than deleting that Q&A discussion entirely.

Thus, Q&A is a collaborative approach to problem-solving, and a successfully answered question is a valuable resource for other researchers to deal with a similar problem. Because of this, RG keeps questions visible so that other researchers can see them and learn from them and perhaps add their own knowledge.

4. Study Background:

Li et al. (2015) explored characteristics and prediction of academic answer quality on RG Q&A where web coded features (like RG scores, impact points, institutional RG score, institutional impact points) and human coded features (like social elements, consensus building, factual information) was explored. The results demonstrated that web-captured features have more influence and are associated with high-quality answers than human-coded features. *Li, He and Zhang (2016)* conducted a pilot study among 15 scholars to investigate how they assess the quality of content on the academic social networking platforms. The study included the Library Information Services domain with a dataset of 15 questions and 157 answers on RG Q&A. The results of the study identified nine criteria of quality evaluation. Out of nine aspects, the content of academic text and the users' beliefs & preferences were the two most commonly used aspects that affect quality judgment. *Jeng, DesAutels, He and Li (2017)* studied how academic social networking sites could facilitate the information exchange between scholars. Content analysis was performed on 1128 posts among three disciplines namely LIS, Arts, and Astrophysics on RG Q&A. The results found that in Arts, questioners request external and known resources to satisfy their answers and was not similar in the LIS discipline. Information-typed questions took longer response time by scholars than discussion ones. In Astrophysics, scholars preferred factual information as well as writing longer responses. Moreover, behaviors such as providing personal opinions and socio-emotional reactions were also common among the three disciplines. *Deng, Tong and Fu (2018)* studied the interaction pattern using Bales' Interaction Process Analysis (IPA) between scholars on RG Q&A, assessed the quality of academic answers and socio-emotional reactions during interactions. The authors chose the Library and Information Science domain and the dataset consisted of 371 questions with 7530 answers. The findings suggested that scholars tend to describe their own experiences when asking for information and suggestion questions. The authors accentuated that academic social networking sites can strengthen the interaction between scholars through the promotion of recommendation mechanisms, classification of questions, and strengthening socio-emotional experience. *Li et al. (2018)* proposed a study to investigate how scholars evaluate the quality of academic answers on academic social Q&A sites. To assess the characteristics of high-quality academic answers, 1021 responses were collected and compared using statistical analysis across three domains i.e., Library and information services, History of arts, and Astrophysics on RG Q&A. The results revealed that in Astrophysics, peer-judged academic answer quality is affected most by answer length as longer answers are more likely to be peer-judged as quality than short ones with fewer subjective opinions. In the Arts domain, however, neither characteristic is significantly linked to answering quality. Further, the academic credibility of the answerer affects peers' assessment of the answer quality, especially in LIS and Astrophysics. *Deng et al. (2019)* explored the influential & motivational factors of the scholars to respond in academic social networking platforms. The study was performed on RG Q&A by coding the 445 questions and 8457 related answers. The findings resulted that positive action-oriented statements to questions garnered more responses and entice reads from other scholars while positive procedural statements and negative action-oriented statements attracted responses to answers such as recommendations.

5. Research Gap:

Successful research is highly productive and brings a positive change in the societal well-being. At the same time, it is highly competitive, be it doctoral research or research publishing. In absence of timely answers to their research-based queries, the researches may tend to find a workaround or they may ignore the problem at the first place defeating the basic trait of research activity. It is the responsibility of the LIS professionals to either develop such systems or to promote such services which are already in place.

Scanning the literature reveals that Q&A systems have not being the importance they deserve in the LIS literature. One could hardly find relevant literature discussing the disciplinary adoption of Q&A culture by the global community of researchers. Q&A is a natural phenomenon facilitating learning, enhanced understanding and higher-order thinking skills (HOTS). LIS research as reported in the published literature heavily lack studies based on this unexplored aspect. Further, none of the researches try to evaluate the Q&A trends of such platforms on the basis of statistical approach. Therefore, it is is highly recommended to to undertake such a study to unravel the knowledge deficit in these domain.

6. Study Rationale & Contribution of Present Study:

Q&A play a vital role in the progression of research. Researchers throughout their journey along the strenuous path face many challenges and are surrounded by doubts, confusion and uncertainties. During such difficult situations they not only need help from their peers and mentors but also need to reach out to global community for finding answers to their queries. Unfortunately, their exits is no such institutional mechanism with global base to build a sustainable knowledge base (KB) system or model to expedite the pace of research. This lack of a well managed and ever-growing Q&A model adversely affects the researchers worldwide in general and researchers of the developing and underdeveloped countries who perhaps need more help. However, the LIS professionals and libraries must sensitise their user community regarding the importance of such a system motivating them to utilise the RG Q&A service which comes at no cost. This study, by highlighting the disciplinary Q&A usage trends available on RG platform will encourage the research community to take advantage of this service by actively participating in Q&A based scholarly discourses and to undertake their research endeavour with confidence. The findings of the study with several statistical test results adequately figure out & fill the current gap in the related literature.

7. Originality/Value:

Very few studies have been conducted on the Q&A trends in scholarly communication. This study focuses on the Q&A trends among the academia & research community in the Research Gate (RG). Further, the study also tries to evaluate the approach of the current age academician & researcher with available dataset in the RG Q&A platform and compare it to the other different approaches.

8. Study Objectives:

The main goal of the study is to explore and analyse the disciplinary Q&A trends prevailing in the scientific social network (SSN)/scientific collaborative network (SCN) also known as academic social network (ASN), ResearchGate (RG). To accomplish this goal, research objectives have been framed as enumerated under:

a) To examine the subject-wise status of Q&A on ResearchGate (RG).
b) To explore the content sharing trends on RG Q&A.
c) To study the descriptive statistics of the RG Q&A data.
d) Assessing or evaluating the Shapiro- Wilk Normality test on RG Q&A data.
e) Study the parametric one sample t- test on RG Q&A data.
f) Compute and interpret the Pearson's correlation in RG Q&A data.

9. Research Design:

Research design is the framework of research methods and techniques followed in a given study. The present study adopted a quantitative research design (Correlational as-well-as Explanatory) to pursue the research questions and objectives. The sample data of the current research gathered from the RG platform availed via URL, https://www.researchgate.net/. A total of 35 subjects were picked from the subject areas on RG

Q&A (without login), with seven different types of variables identified for each subject: posts, questions, publications, articles, preprints, conference paper, and literature review (Table 1). For the data cleansing process, the collected data were tabulated in an excel spreadsheet. It enhanced data visualization and easily identified any anomaly & errors attached inside the sample data.

To understand the arrangement and correlation between different variables of the dataset, various statistical tests were performed using JASP (Journal of Statistical Software) statistical software (v. 0.14.1.0). It is a free and open-source graphical program for statistical analysis. It has user- friendly interface and has standard analysis procedures for integration with the open science framework (OSF) and support for APA format (copy graphs and tables directly into word). (Admin, 2021). Among the statistical tests, Descriptive Statistics measure the central tendency (i.e., Mean, Median, Mode & Sum) and dispersion (i.e. Std. Error of Mean, Std. Deviation, Minimum & Maximum) of the sample data; parametric t-test and Shapiro- Wilk test check the normality among study variables; and the Pearson Correlation test was done to check the correlation between the variables.

10. Results and Discussion:

Table 1: Subject-wise categorization of collected data according to different variable types

SN	Subject	Posts	Questions	Publications	articles	preprints	Conference paper	Literature review	Total
1.	Agriculture	2185	925	882344	6340	254	933	19	893000
2.	Anthropology	697	300	158958	7339	233	425	32	167984
3.	Architecture	1123	513	760456	5201	1115	1599	34	770041
4.	Arts	1022	270	505426	5911	909	930	63	514531
5.	Astronomy	354	163	257892	7261	516	1047	19	267252
6.	Bibliography	436	128	126977	7096	185	614	147	135583
7.	Biology	2110	745	300085	7009	828	544	631	311952
8.	Botany	1577	1067	194180	7358	54	844	87	205167
9.	Chemistry	3131	1433	708771	6758	289	1058	142	721582
10.	Computer Science	2380	945	533984	6822	466	1382	42	546021
11.	Earth Sciences	563	324	176646	7897	74	915	20	186439
12.	Education	3250	1028	696279	6056	276	817	4	707710
13.	Engineering	2524	1356	723162	7559	346	770	26	735743
14.	Epistemology	673	280	150110	7714	402	483	47	159709
15.	Ethics	2338	844	399011	7076	510	574	94	410447
16.	Genealogy	48	26	30726	8166	325	370	81	39742
17.	Geography	536	231	220764	7026	405	624	33	229619
18.	Geology	1505	820	387674	3716	122	986	0	394823
19.	History	769	200	723693	6450	430	428	10	731980

20.	Language	1201	403	937030	6070	851	937	7	946499
21.	Law	503	185	558358	6767	725	618	5	567161
22.	Linguistics	1076	525	505114	7072	721	669	21	515198
23.	Literature	1283	206	310782	7017	574	703	33	320598
24.	Manuscripts	2610	733	90619	7186	1303	306	617	103374
25.	Mathematics	4239	1544	1164588	6917	1236	760	3	1179287
26.	Philosophy	2588	788	377684	6944	576	464	18	389062
27.	Physics	2943	1006	1143535	6287	730	900	2	1155403
28.	Political Science	468	218	161418	6943	154	477	14	169692
29.	Psychology	3652	1382	804869	6721	516	719	61	817920
30.	Religion	1549	434	536660	7369	384	552	16	546964
31.	Science	2319	695	10000	6746	852	565	93	21270
32.	Sociology	1601	751	548191	6950	354	422	19	558288
33.	Statistics	4575	2799	612039	6607	892	848	7	627767
34.	Technology	287	145	1037138	6604	404	1177	25	1045780
35.	Zoology	642	426	195078	7880	61	654	45	204786
	Total	**58757**	**23838**	**16930241**	**238835**	**18072**	**26114**	**2517**	**17298374**
	Average	**1678.77**	**681.09**	**483721.17**	**6823.86**	**516.34**	**746.11**	**71.91**	**-**

**Shaded areas indicates maximum coverage of specific content-types under subject categories

9.1. Distribution according to Subject of the Content:

The distribution of total RG contents (study samples) among the 35 subject categories is shown in Figure 1. Out of the 35 subject areas, Mathematics comprises the maximum number of topics (1179287 or 6.81%). It is followed by Physics and Technology, with 6.68% (1155403) and 6.05% (1045780) coverage, respectively.
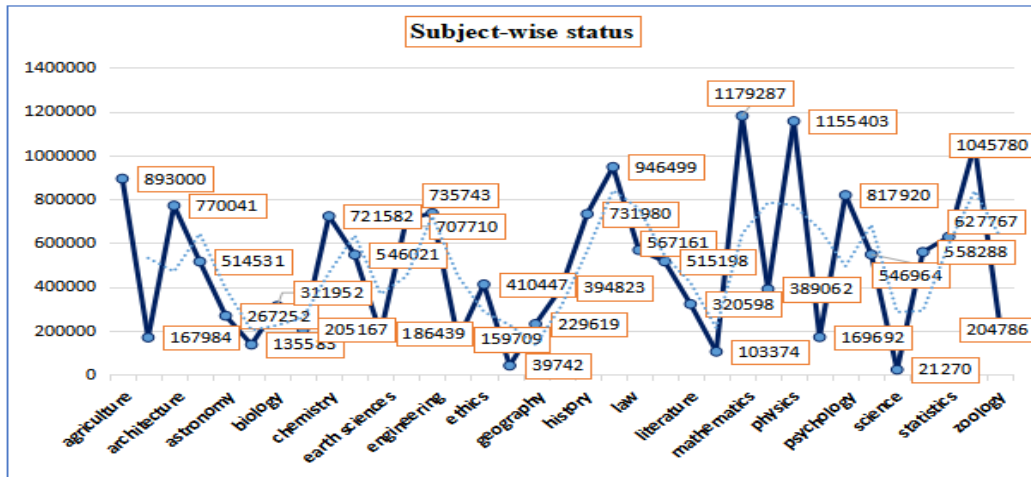
Figure 1: Subject-wise distribution of total contents in RG

9.2. Distribution according to Content types:

Figure 2 below shows the trends of content sharing on RG Q&A in which the majority of the contents comprises Publications with almost 98% coverage (16930241), followed by Articles (238835, ~2%). The rest of the content types covers less than 1% in total.
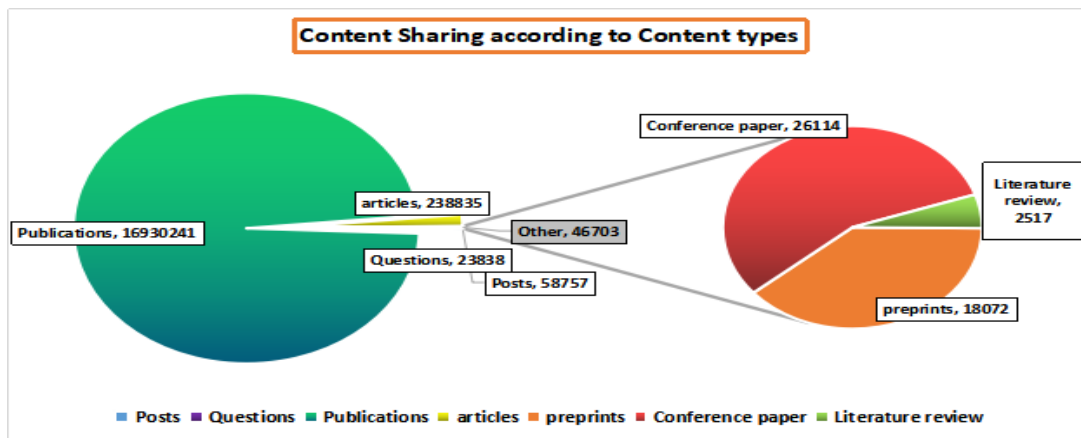
Figure 2: Content-wise distribution of total RG Contents

9.3. Content -sharing Trends in RG:

The distribution patterns of available contents (in percentage) of 35 selected subject categories in RG are shown in Figure 3 according to 6 distinct content types. According to the distribution trends of total contents, the "Literature Review" is the top unequally distributed field across different subject categories. It possesses a top pic of 25.07% (631) for Biology and a minimum of 0.00% (0) for the Geology subject

field (pic diff. 25.07%). On the contrary, "articles" is the most equally distributed field with a pic-difference of 1.86% only. In that case, the top pic consists of 3.42% (8166) coverage for Genealogy, whereas, lowest pic consists of 1.56% (3716) for the Geology subject field. The rest of the content types mostly possess in-between trends.

Content Sharing Trends

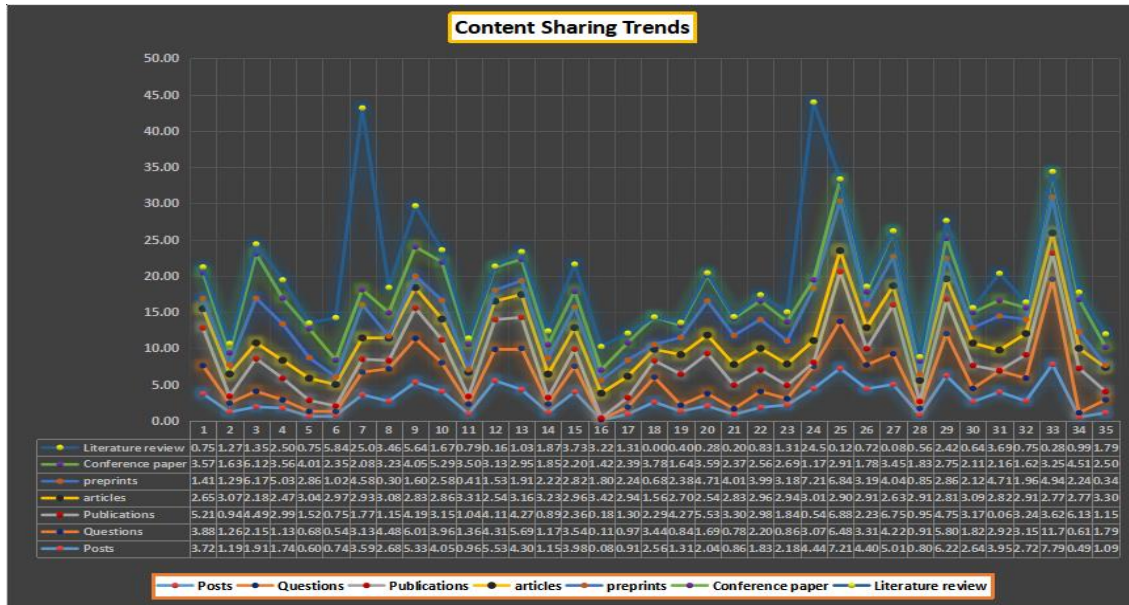|	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
Literature review	0.75	1.27	1.35	2.50	0.75	5.84	25.03	3.46	5.64	1.67	0.79	0.16	1.03	1.87	3.73	3.22	1.31	0.00	0.40	0.28	0.20	0.83	1.31	24.50	0.12	0.72	0.08	0.56	2.42	0.64	3.69	0.75	0.28	0.99	1.79
Conference paper	3.57	1.63	6.12	3.56	4.01	2.35	2.08	3.23	4.05	5.29	3.50	3.13	2.95	1.85	2.20	1.42	2.39	3.78	1.64	3.59	2.37	2.56	2.69	1.17	2.91	1.78	3.45	1.83	2.75	2.11	2.16	1.62	3.25	4.51	2.50
preprints	1.41	1.29	6.17	5.03	2.86	1.02	4.58	0.30	1.60	2.58	0.41	1.53	1.91	2.22	2.82	1.80	2.24	0.68	2.38	4.71	4.01	3.99	3.18	7	6.84	3.19	4.04	0.85	2.86	2.12	4.71	1.96	4.94	2.24	0.34
articles	2.65	3.07	2.18	2.47	3.04	2.97	2.93	3.08	2.83	2.86	3.31	2.54	3.16	3.23	2.96	3.42	2.94	1.56	2.70	2.54	2.83	2.96	2.94	3.01	2.90	2.91	2.63	2.91	2.81	3.09	2.82	2.91	2.77	2.77	3.30
Publications	5.21	0.94	4.49	2.99	1.52	0.75	1.77	1.15	4.19	3.15	1.04	4.11	4.27	0.89	2.36	0.18	1.30	2.29	4.27	5.53	3.30	2.98	1.84	0.54	6.88	2.23	6.75	0.95	4.75	3.17	0.06	3.24	3.62	6.13	1.15
Questions	3.88	1.26	2.15	1.13	0.68	0.54	3.13	4.48	6.01	3.96	1.36	4.31	5.69	1.17	3.54	0.11	0.97	3.44	0.84	1.69	0.78	2.20	0.86	3.07	6.48	3.31	4.22	0.91	5.80	1.82	2.92	3.15	11.70	0.61	1.79
Posts	3.72	1.19	1.91	1.74	0.60	0.74	3.59	2.68	5.33	4.05	0.96	5.53	4.30	1.15	3.98	0.08	0.91	2.56	1.31	2.04	0.86	1.83	2.18	4.44	7.21	4.40	5.01	0.80	6.22	2.64	3.95	2.72	7.79	0.49	1.09

Legend: Posts — Questions — Publications — articles — preprints — Conference paper — Literature review

Figure 3: Content-sharing trends (in percentage) according to subject categories

9.4. Statistical Analysis:

9.4.1. Descriptive Statistics:

Descriptive statistics is the branch of statistics that develops where necessary, methods for collecting, processing, quantitative and qualitative analysis of data. The objective of the descriptive statistics is to summarize or represent, through statistics, the data available when they are numerous. There are various types of statistics that are used to describe data, e.g., measures of central tendency (i.e., Mean, Median, Mode & Sum), measures of dispersion (i.e., Std. Error of Mean, Std. Deviation, Minimum & Maximum) etc.

Table 2: Descriptive Statistics Table					
	N	**Mean**	**Median**	**SD**	**SE**
Questions	35	681.1	525	558	94.3
Posts	35	1678.8	1505	1187	200.7
Publications	35	483721.2	505114	320727	54212.7
Articles	35	6823.9	6944	804	136.0

Table 2: Descriptive Statistics Table					
	N	Mean	Median	SD	SE
Preprints	35	516.3	430	329	55.6
Conference paper	35	746.1	703	288	48.7
Literature review	35	71.9	26	143	24.1

Table 2 above discuss the descriptive values of the sample data, where each of the seven types of sample consist of 35 sample values (N). Among them, Publications possess maximum value while measuring central tendency (mean & median) & dispersion (std. deviation & std. error) and those are the minimum in the case of Literature Review.

9.4.2. Normality Test:

Shapiro-Wilk tests was conducted in order to determine whether the distributions of Applications and Registrations were significantly different from a normal distribution. The Shapiro-Wilk test is the most powerful test to tell if random sample comes from a normal distribution. The null hypothesis for this test is that the data are normally distributed. The Prob < W value listed in the output is the p-value.

If the chosen alpha level is 0.05 (i.e., critical value, α =0.05) and,

(i) the p-value is less than 0.05, then the null hypothesis that the data are normally distributed is rejected (do not assume a normal distribution).

(ii) the p-value is greater than 0.05, then the null hypothesis is not rejected (assume a normal distribution)

Table 3: Shapiro-Wilk Normality Test Results		
	W	p
Questions	0.848	< .001
Posts	0.929	0.027
Publications	0.950	0.110
Articles	0.865	< .001
Preprints	0.944	0.072

Table 3: Shapiro-Wilk Normality Test Results		
	W	***p***
Conference paper	0.942	0.065
Literature review	0.460	< .001

Note. A low *p*-value suggests a violation of the assumption of normality

Shapiro-Wilk tests were conducted (taken critical value, α =0.05) in order to determine whether the distributions of seven different types of samples were significantly different from a normal distribution. Among the seven variables only three, viz Publications, Preprints & Conference paper shows normal distribution *(p>0.05)* with low *p*-value, and the data-set for other four variables don't possess a normal distribution *(p<0.05)*. Overall, the above variables had distributions that were significantly differ from normality based on an alpha of 0.05 and a very low p-value. The results are presented in Table 3.

9.4.3. One Sample T-Test:

The parametric one sample t-test determines whether the sample mean is statistically different from a known or hypothesized population mean. To check whether the one sample t-test is applicable for the sample data or not to provide a valid result, some assumptions are required to maintain,
- The test variable should be measured on a continuous scale.
- The test variable data should be independent i.e., no relationship between any of the data points.
- The data should be approximately normally distributed
- There should be no significant outliners.

But in previous Shapiro-Wilk test it is determined that the sample variable of the study is not significantly normally distributed. Though, the assumption check of normality with Shapiro-Wilk is not significant suggesting that the sample variables are normally distributed, therefore this assumption is not violated.

The following table 4 displays the result of one sample t-test and identified that there is a significant difference between the mean samples of all the seven variables. The statistics of Student's t-test shows a maximum value in the case of 'articles' (viz. 50.19) and the 'p' value of all the variables taken to be <0.001 except literature review (i.e., 0.005). Moreover, the mean difference is maximum (viz. 483721.2) for 'Publications', whereas, 'articles' shows a maximum of Cohen's d effect size (viz. 8.483).

Table 4: One Sample T-Test

		Statistic	df	p	Mean difference	95% Confidence Interval Lower	95% Confidence Interval Upper		Effect Size	95% Confidence Interval Lower	95% Confidence Interval Upper
Questions	Student's t	7.22	34.0	<.001	681.1	489.4	873	Cohen's d	1.221	0.776	1.655
Posts	Student's t	8.37	34.0	<.001	1678.8	1270.9	2087	Cohen's d	1.414	0.938	1.880
Publications	Student's t	8.92	34.0	<.001	483721.2	373547.7	593895	Cohen's d	1.508	1.016	1.990
articles	Student's t	50.19	34.0	<.001	6823.9	6547.5	7100	Cohen's d	8.483	6.377	10.464
preprints	Student's t	9.29	34.0	<.001	516.3	403.3	629	Cohen's d	1.570	1.066	2.062
Conference paper	Student's t	15.33	34.0	<.001	746.1	647.2	845	Cohen's d	2.591	1.890	3.284
Literature review	Student's t	2.98	34.0	0.005	71.9	22.9	121	Cohen's d	0.504	0.148	0.852

Note. For the Student t-test, effect size is given by Cohen's d.
Note. For the Student t-test, location estimate is given by the sample mean d.
Note. For the Student t-test, the alternative hypothesis specifies that the mean is different from 0.
Note. Student's t-test.

9.4.4. Correlation Matrix

A correlation expresses the strength of linkage or co-occurrence between two variables. The most common measure of correlation in stats is the Pearson Product Moment Correlation (PPMC) which shows the linear relationship between two sets of data. The correlation coefficient between two continuous-level variables is called Pearson's r or Pearson product-moment correlation coefficient which is represented typically as the letter r and has a single value between -1 and +1. This value measures the strength of the linkage.

In the present study, total seven variables possess 21 different types of correlation results a variety of Pearson's r value ranges between -1 and +1. Among the 21 correlation types, 7 possess negative correlation value suggests presence of no correlation between those variables. Whereas, the other 14 possess a positive correlation value ranges from 0.013 to 0.906 depending upon respective p-values. The correlation link is highly strong in between 'Post & Question' (viz. 0.906), and weakest in between 'Literature Review & Question' (viz. 0.013). *(see Table 5 & Figure 4)*

<div align="center">

Table 5: Pearson's Correlations

		n	Pearson's r	p	Lower 95% CI	Upper 95% CI	VS-MPR†
Posts	- Publications	35	0.454 **	0.006	0.142	0.684	11.766
Posts	- preprints	35	0.392 *	0.020	0.067	0.641	4.698
Posts	- Conference paper	35	0.100	0.568	-0.241	0.419	1.000
Posts	- Literature review	35	0.124	0.478	-0.218	0.439	1.000
Posts	- articles	35	-0.184	0.290	-0.487	0.159	1.025
Posts	- Questions	35	0.906 ***	7.073e -14	0.821	0.952	1.718e +11
Publications	- preprints	35	0.294	0.087	-0.044	0.571	1.733
Publications	- Conference paper	35	0.475 **	0.004	0.168	0.698	16.830
Publications	- Literature review	35	-0.314	0.066	-0.586	0.021	2.051
Publications	- articles	35	-0.409 *	0.015	-0.653	-0.088	5.929
Publications	- Questions	35	0.392 *	0.020	0.067	0.641	4.715
preprints	- Conference paper	35	0.080	0.649	-0.260	0.402	1.000
preprints	- Literature review	35	0.352 *	0.038	0.022	0.614	2.968
preprints	- articles	35	-0.201	0.246	-0.501	0.141	1.066
preprints	- Questions	35	0.230	0.183	-0.112	0.523	1.183
Conference paper	- Literature review	35	-0.281	0.102	-0.562	0.058	1.578
Conference paper	- articles	35	-0.485 **	0.003	-0.705	-0.182	20.474
Conference paper	- Questions	35	0.188	0.279	-0.155	0.491	1.033
Literature review	- articles	35	0.149	0.393	-0.194	0.459	1.000
Literature review	- Questions	35	0.013	0.940	-0.321	0.345	1.000
articles	- Questions	35	-0.156	0.370	-0.465	0.187	1.000

</div>

* p < .05, ** p < .01, *** p < .001

† Vovk-Sellke Maximum p -Ratio: Based on the p -value, the maximum possible odds in favor of H_1 over H_0 equals 1/(-e p log(p)) for $p \leq .37$ (Sellke, Bayarri, & Berger, 2001).

#shaded rows of relations denote positive correlation values *(r = +ve)*
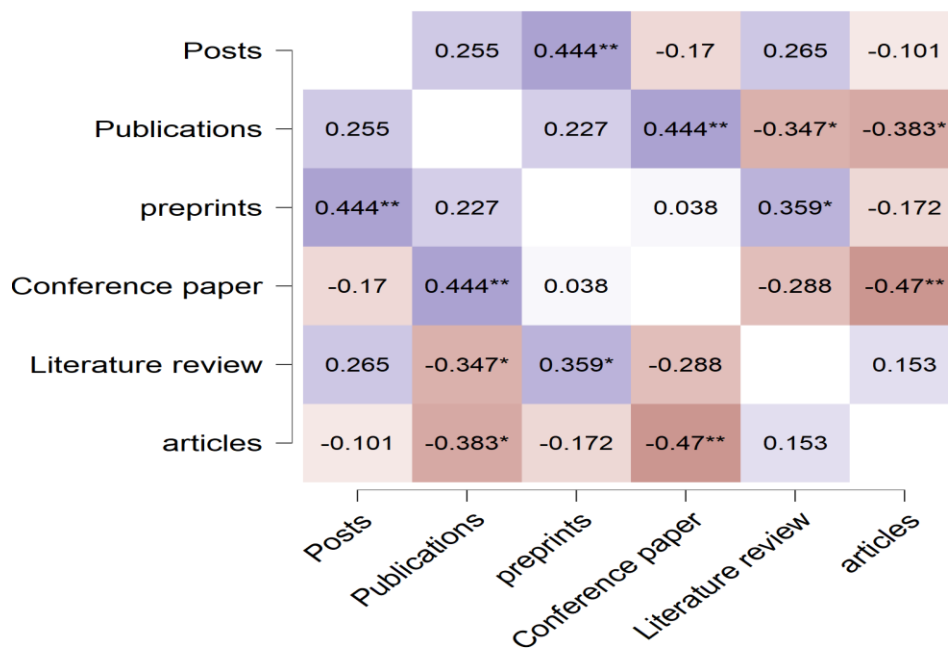
Figure 4: Partial Pearson's r Heatmap

11. Major Findings:

i.   RG Q&A has covered seven types of sample variables (i.e., Posts, Questions, Publications, Articles, Preprints, Conference paper, Literature review) consists of 35 sample value "N" (Subjects) among which Statistics has the maximum number of posts (4575) and Genealogy has the minimum number of posts (48). Statistics has the maximum number of questions (2799) and Genealogy has the minimum number of questions (26). Mathematics has the maximum number of Publications (11,64588) and Science has the minimum number of publications (10000). Genealogy has the maximum number (8166) and Geology has the minimum number of articles (3716). Manuscripts has the maximum number (1303) and Botany has the minimum number of preprints (54). Architecture has the maximum number (1599) and Manuscripts has the minimum number of conference papers (306). Biology has the maximum number (631) and Physics has the minimum number of literature review (2). Geology has zero (0) number of Literature review. One of the most popular subject Economics does not include in this study due to non-availability of sufficient amount of data of the particular in ResearchGate.

ii. On the basis of the total content of the subject areas, total of Literature review variable is 2517(1%) followed by Preprints 18071(5%), Questions 23838(6%), Conference paper 26114(7%), Posts 58757(16%), Articles 238835(65%).

iii. Various types of descriptive statistics were applied to the data i.e., measurement of central tendency and measures of dispersion and it is observed from the descriptive values of the sample data that Publications possess maximum value while measuring central tendency (mean & median) & dispersion (std. deviation & std. error) and those are the minimum in the case of Literature Review.

iv. Shapiro- wilk Normality test were conducted and it is observed that among the seven variables only three, viz Publications, Preprints & Conference paper shows normal distribution *(p>0.05)* with low *p*-value, and the data-set for other four variables don't possess a normal distribution *(p<0.05)*.

Overall, the seven variables had distributions that were significantly differ from normality based on an alpha of 0.05 and a very low p-value.

v.   One sample t- test were also conducted and from the results it is observed that there is a significant difference between the mean samples of all the seven variables. The statistics of Student's t-test shows a maximum value in the case of 'articles' (viz. 50.19) and the 'p' value of all the variables taken to be <0.001 except literature review (i.e., 0.005). Moreover, the mean difference is maximum (viz. 483721.2) for 'Publications', whereas, 'articles' shows a maximum of Cohen's d effect size (viz. 8.483).

vi. The study also reveals that seven variables possess 21 different types of correlation results a variety of Pearson's r value ranges between -1 and +1. Among the 21 correlation types, 7 possess negative correlation value suggests presence of no correlation between those variables. Whereas, the other 14 possess a positive correlation value ranges from 0.013 to 0.906 depending upon respective p-values. The correlation link is most strengthful in between 'Post & Question' (viz. 0.906), and weakest in between 'Literature Review & Question' (viz. 0.013).

12. Conclusion:

Question-Answer or Q&A or Q-A, is a time-tested phenomenon in scholarly communication. Answers given by experts help information seekers the necessary inputs to continue their scientific endeavour with confidence. RG offers a great platform for rich Q&A thus strengthening the academic collaboration and knowledge sharing. This paper focuses on the status, qualities and role of this academic social network in maintaining and nurturing a robust Q&A ecosystem in a sustainable manner. To assess the status and trends of Q&A, the relevant data was extracted for as many as 35 disciplines and for subsequent analysis using statistical techniques.

RG may close a Q&A discussion (Closed Questions) if it violets its policy. RG endeavours to maintain good standards of civility while establishing the right balance between community respect and free expression can be challenging. RG recommends to adhere to its Q&A guidelines in order to derive maximum benefit of Q&A ensuring relevancy, professional ethics and standards, and usefulness.

The Q&A enables participants and community as a whole to derive benefit from any valuable insights or ideas that were exchanged. RG connects the world of science and accelerate scientific progress by facilitating meaningful scientific discussions that happen in Q&A forum. Q&A help researchers exchange high-quality questions and answers in a research-focused environment. This collaborative approach to problem-solving not only supports academic networking in true sense but also help the researchers to build their reputation, at any stage of their career.

References

1) Admin. (2021, June 1). *Top Free and Open-Source Statistical Analysis Software*. Top IT Firms - Result of In-Depth Research & Analysis. https://www.itfirms.co/top-open-source-statistical-analysis-software/#9

2) Ayoub, A., Amin, R., & Wani, Z. A. (2020). The Leading Answerers of Social Q&A Sites: Exploring their Characteristics and Contributions. *International Journal of Information Studies and Libraries*, *5*(2). http://www.publishingindia.com/ijisl/107/the-leading-answerers-of-social-q-a-sites-exploring-their-characteristics-and-contributions/912/6285/

3) Barbour, K., & Marshall, D. (2012). The academic online: Constructing persona through the World Wide Web. *First Monday*, *17*(9). https://doi.org/10.5210/fm.v0i0.3969

4) Deng, S., Tong, J., & Fu, S. (2018, May 23). Interaction on An Academic Social Networking Sites. *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries*. https://doi.org/10.1145/3197026.3197062

5) Deng, S., Tong, J., Lin, Y., Li, H., & Liu, Y. (2019). Motivating scholars' responses in academic social networking sites: An empirical study on ResearchGate Q&A behavior. *Information Processing & Management*, *56*(6), 102082. https://doi.org/10.1016/j.ipm.2019.102082

6) JASP Team (2021). JASP (Version 0.14.1.0) [Computer Software]

7) Jeng, W., DesAutels, S., He, D., & Li, L. (2016). Information exchange on an academic social networking site: A multidiscipline comparison on researchgate Q&A. *Journal of the Association for Information Science and Technology*, *68*(3), 638–652. https://doi.org/10.1002/asi.23692

8) Jeng, W., He, D., & Jiang, J. (2014). User participation in an academic social networking service: A survey of open group users on Mendeley. *Journal of the Association for Information Science and Technology*, *66*(5), 890–904. https://doi.org/10.1002/asi.23225

9) Li, L., He, D., Jeng, W., Goodwin, S., & Zhang, C. (2015). Answer Quality Characteristics and Prediction on an Academic Q&A Site. *Proceedings of the 24th International Conference on World Wide Web*, 1453–1458. https://doi.org/10.1145/2740908.2742129

10) Li, L., He, D., & Zhang, C. (2016). Evaluating Academic Answer Quality: A Pilot Study on ResearchGate Q&A. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9751*, 61–71. https://doi.org/10.1007/978-3-319-39396-4_6

11) Li, L., He, D., Zhang, C., Geng, L., & Zhang, K. (2018). Characterizing peer-judged answer quality on academic Q&A sites. *Aslib Journal of Information Management*, *70*(3), 269–287. https://doi.org/10.1108/ajim-11-2017-0246

12) Ovadia, S. (2014). ResearchGate and Academia.edu: Academic Social Networks. *Behavioral & Social Sciences Librarian*, *33*(3), 165–169. https://doi.org/10.1080/01639269.2014.934093

13) ResearchGate. (n.d.). *Q&A Guidelines*. ResearchGate Help Center. Retrieved June 18, 2021, from https://explore.researchgate.net/pages/viewpage.action?pageId=14155852

14) ResearchGate. (2018). *ResearchGate App (Designed for iPhone)*. Mac App Store. https://apps.apple.com/us/app/researchgate/id1208163937?ls=1

15) ResearchGate. (2019). *About ResearchGate*. ; https://www.researchgate.net/about

16) Sellke, T., Bayarri, M. J., & Berger, J. O. (2001). Calibration ofρValues for Testing Precise Null Hypotheses. *The American Statistician*, *55*(1), 62–71. https://doi.org/10.1198/000313001300339950

17) Xiang, Y., Wang, R., Wang, Z., & Chen, Z. (2012, October). Extracting Keyphrase for Social Q&A Sites. *Proceedings of the 2012 International Conference*