

The anatomy of an electronic discussion list for librarians, KUTUP-L: Bibliometric and content analyses of postings

Yaşar Tonta¹; Doğan Karabulut²

¹Department of Information Management, Faculty of Letters,
Hacettepe University, 06800 Beytepe, Ankara, Turkey.
tonta@hacettepe.edu.tr

²Turkish Grand National Assembly Library,
06543 Bakanlıklar, Ankara, Turkey
dogank@tbmm.gov.tr

Abstract

Electronic discussion lists are widely used as a professional and scientific communication tool since late 1980s. Analysis of messages sent to discussion lists provides useful information on professional as well as scientific communication patterns. In this paper, we present the findings of a bibliometric analysis of some 20,000 messages sent to KUTUP-L, an electronic discussion list for Turkish librarians, between 1994 and 2008. We test if the distributions of messages and their authors conform to Pareto, Price and Lotka laws. We then analyze the contents of 977 messages based on a stratified sample. Findings indicate that the number of messages sent to KUTUP-L has increased over the years along with the number of authors. Two thirds (1,232) of about 1,900 list members posted at least one message to the list while the rest preferred to be so called "lurkers". Some 35 authors posted almost half (49%) the messages while 20% of the authors posted 83% of all messages. The distribution of messages to authors conform to Price ("the square root of all authors would post half the messages") and Pareto laws (so called "80/20 rule"), respectively. Of the 1,232 authors, one third (as opposed to 60% predicted by Lotka's Law) sent only one message to the list. Results of content analysis show that 40% of messages sent to the list were off-topic. Issues about or related with information management services (32%), library and information science (23%) and professional and scientific communication (19%) were discussed more often in the list. The intent analysis of the postings shows that three quarters of the messages were initiatory while the rest were reflexive. That's to say that the majority of messages posted on KUTUP-L to initiate a discussion did not seem to generate enough interest for others to reflect upon them by sending follow up

messages, suggesting that professional and scientific communication taking place on KUTUP-L on certain subjects can be characterized as more of a one-way communication than a participatory one.

Keywords: KUTUP-L; electronic discussion lists; electronic publishing; professional communication; scientific communication; bibliometric analysis; content analysis

1. Introduction

The history of computer-based communication dates back to mid-1960s. Host-based mail systems were later replaced by the electronic mail (or e-mail) system of the ARPANET computer network in early 1970s. E-mail has become a “killer app” on BITNET, the predecessor of the current day Internet. LISTSERV, an electronic discussion list management software, was introduced in early 1990s and e-mail based discussion lists such as PACS-L, LIBREF, PUBLIB and WEB4LIB proliferated thereafter. Messages posted to such discussion lists contain invaluable resources for historians, social scientists, social network analysts, and bibliometricians, among others, and they can be analyzed to study professional and scientific communication patterns along with the topics discussed and the productivity of authors.

KUTUP-L, an electronic discussion list for Turkish librarians, was set up in June 1991 to share information, exchange ideas and discuss professional issues. It currently has some 1,900 registered members. Based on the analysis of messages in KUTUP-L archives, this paper aims to address the following research questions:

- Has the number of messages posted and the subjects discussed in KUTUP-L increased and proliferated over the years?
- Have professional and scientific communication patterns in KUTUP-L changed over the years in librarianship?
- Does the distribution of messages to authors (thus the authors' productivity) conform to Pareto, Price and Lotka laws?

Findings of this study will shed some light on the evolution of KUTUP-L as an electronic discussion list for the last two decades. Intent and content analysis of KUTUP-L postings will provide both quantitative and qualitative information about the level of activity as well as the types of postings, their subjects and authors.

2. Literature Review

Content analysis is a commonly used method of studying messages sent to electronic discussion lists. Wildemuth et al. [1] used content analysis to study 14 different library discussions lists and found that the relatively high percentage of messages were intended to discuss certain issues.

Content analysis of 309 messages posted at PUBYAC, a discussion list created for public librarians working in children's and young adult services, shows that the majority of postings were of reference type, indicating that the list took on the role of an information source for its subscribers [2]. PUBYAC postings were analyzed under six different categories in a different study: programs (27%), finding books (21%), collection (16%), library administration and policy (9%), professional issues (9%), and announcements (7%). Half the messages were responses to earlier requests and (37%) were inquiries while the rest (13%) were announcements and general comments. Authors of messages were generally thankful and complimented the list and its subscribers [3].

A survey of MEDLIB-L (Medical Library Association's discussion list) users showed that about 90% of them read MEDLIB-L at work and spend less than three hours a week for this purpose. They used the list to comment on various issues and answer questions more often than to ask questions or start discussions [4]. The former types of messages are called "reflexive" ones while the latter are "initiatory" [5, 6]. Similarly, almost three quarters of messages sent to EVALTALK, a listserv for evaluation professionals, were comments/responses on requests while the rest were questions or requests, indicating that subscribers used EVALTALK discussion list as an informational tool [7]. We see the same pattern in the messages of the trombone users' discussion list, Trombone-L and a listserv used as a journal. Some 72% of the Trombone-L messages were comments/answers and 28% questions, although percentages varied by topics discussed (One-third of the messages were off-topic.) [8]. Reflexive messages constituted 65% of listserv messages while the rest were initiatory messages [9]. More than half (51%) the messages posted to HUMANIST discussion list were made up of responses while 25% questions, 19% announcements and 5% administrative ones [10]. Some 56% of the messages sent to ABUSE-L, a discussion list on social work, were classified under "discussion" (i.e., reflexive messages) [11].

The communication patterns of authors posting at discussion lists tend to conform to some bibliometric laws such as Pareto (80% of messages get posted by 20% of all authors), Price ("the square root of all authors would post half the messages") and Lotka (60% of authors send one message to the list while decreasing percentages of authors send more, i.e., 15% two, 6.6% three, 3.75% four and so on) [12,13,14,15]. Messages posted at two discussion

lists (LINGUIST and HEL-L) seemed to conform to Lotka's Law, although the correlation was not high [16].

3. Data and Method

To address research questions, we first obtained access to KUTUP-L logs hosted by the Middle East Technical University in Ankara, Turkey. Logs archived between June 1991 and September 1994 were not available due to technical reasons. We imported the contents of all messages along with associated metadata to a spreadsheet package and cleaned the data before analysis. Descriptive statistics and bibliometric analysis are based on a total of 19,827 messages posted on KUTUP-L between 1994 and 2008. We used Pareto, Price and Lotka's laws to find out if the author productivity in KUTUP-L conforms to decreasing power laws. As given in the previous section, the first two laws are relatively easy to explain. To test if data fit Lotka's Law, we used $f(n) = C / n^\alpha$ formula wherein $f(n)$ is a function of frequency, C and α are constants ($C > 0$ and $\alpha \geq 0$). Thus, the number of authors posting n messages is proportional to decreasing power law [14].

We then selected a stratified sample of 977 messages for content analysis (sample size 5%). Using Bellack's communication model, we classified each message either as "initiatory" (i.e., asking a question or initiating a discussion) or "reflexive" (i.e., answering a question or commenting on an issue) [5,6]. Based on Berman's "intent analysis", we also categorized each message according to its purpose (or "intent") under "information transfer" (IT), "information request" (IR) or "discussion of an issue" (IS) [11]. Next, we carried out content analysis to identify the subject(s) of each message using Jarvelin and Vakkari's subject classification [17, 18].

We presented the descriptive statistics and the results of bibliometric and content analysis using tables and figures. We grouped the findings of content analysis in five-year intervals to detect the changes of patterns in professional and scientific communication over the years.

4. Findings and Discussion

A total of 19,827 messages were posted on KUTUP-L between 1994 and 2008, half of which belong to the last five years (2004-2008). The average number of postings per month in recent years is about 175. Findings indicate that communication in the list has increased continuously over the years and KUTUP-L has become a major venue of communication and discussion among library and information professionals (Figure 1). The heaviest

The anatomy of an electronic discussion list for librarians, KUTUP-L

message traffic was observed in March as the Turkish Library Week gets celebrated in the last week of March every year while the list was less busy during summer months. About one third of 1,900 list members never posted a message on KUTUP-L. The gender of 1,232 unique authors contributing to the list is evenly distributed (52% female, 48% male), although males posted more (59%) messages than females.

The distribution of messages to authors conform to both Price and Pareto (80/20 rule) laws: Almost half (49%) the messages were posted by 35 out of 1,232 authors while 20% of all authors posted 83% of all messages. One third (369) of all authors posted only one message, half the percentage predicted by Lotka's Law (60%) (Table 1). More than half the authors (52.60%) posted three messages at most, constituting a mere 3% of all messages. The great majority of authors contributed to the list very little, thereby making them primarily "lurkers". Some 40% of all authors posted five or more messages to the list.

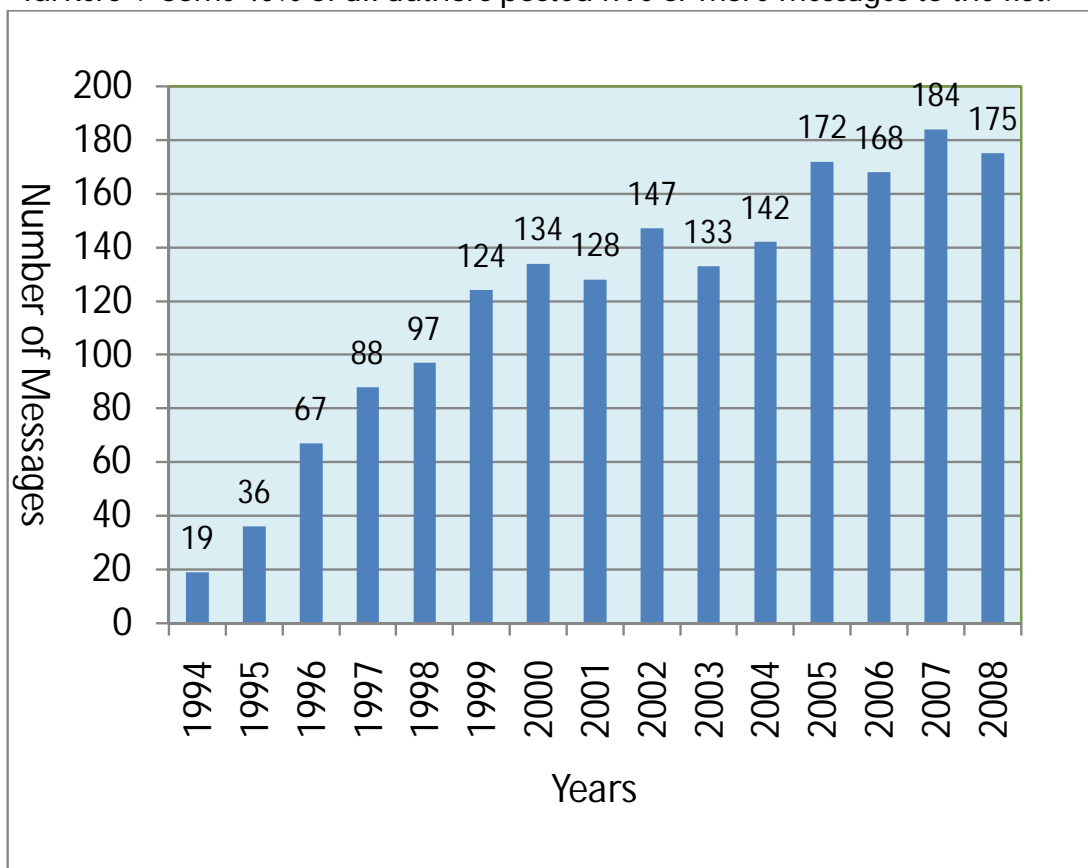


Figure 1: Average number of messages sent to KUTUP-L on a monthly basis (1994-2008)

The authors' productivity data for KUTUP-L posters seem to fit Pareto and Price laws fairly well. Yet, the distribution of messages to authors does not conform to Lotka's Law, which is due to the fact that KUTUP-L has a relatively stable base of contributors (much more than what Lotka's Law predicts) who send messages to the list from time to time. It could be that

The anatomy of an electronic discussion list for librarians, KUTUP-L

characteristics of authorship of a journal article differ from that of a post to a discussion list such as KUTUP-L.

Table 1: Test of Lotka’s Law on KUTUP-L authorship data

# of messages	Expected percentage of authors according to Lotka’s Law (%)	Expected number of authors according to Lotka’s Law	Observed percentage of authors (%)	Observed number of authors
1	60.00	739	30.00	369
2	15.00	184	14.40	178
3	6.60	82	8.20	101
4	3.75	46	7.47	92
5	2.40	30	4.71	58
6	1.60	20	3.33	41
7	1.20	15	3.08	38
8 or more			28.71	355
Total	100.00	1,232	100.00	1,232

Note: Percentage and numbers of authors contributing more than 7 messages to KUTUP-L according to Lotka’s Law are not given in the table.

Intent analysis

Three quarters (76% to be exact) of all postings were “initiatory” (i.e., asking a question or initiating a discussion) while the rest were reflexive (i.e., answering a question or commenting on earlier postings). Table 2 provides descriptive statistics, at five-year intervals, about findings of intent analysis based on a stratified sample of 977 KUTUP-L messages. The percentage of reflexive messages tended to decrease over the years, suggesting that more list members seemed to be indifferent towards KUTUP-L postings. The intention was transferring information in two thirds (67%) of all messages, followed by starting a discussion (23%) and asking for information (10%). The percentage of postings aiming to transfer information increased over the years while the percentage of postings with discussion topics decreased. Forwarded postings made up 16% of all messages, although the percentage is decreasing. The percentage of postings containing links to other web sites is on the rise (20%).

Table 2: Intent analysis of KUTUP-L postings

Type of messages	Years			Total
	1995-1999	2000-2004	2005-2008	
Initiatory	137 (63%)	279 (80%)	325 (79%)	741 (76)
Reflexive	82 (37)	70 (20)	84 (21)	236 (24)
Total	219 (100)	349 (100)	409 (100)	977 (100)

Note: Figures in brackets refer to percentages.

Content analysis

Content analysis of on-topic messages shows that about one third (32%) were related with information management services, 23% with library and information science, and 19% with professional and scientific communication (Table 3). (Each on-topic message was classified under the main topic). Topics discussed on KUTUP-L varied over the years. For instance, postings on information management and professional/scientific communication issues became more prominent in recent years while the percentage of postings on cataloging issues went down drastically (from 50% to 21%) over the years.

The percentages of reflexive postings were well over 50% for some topics (e.g., professional issues, professional training, library management and library automation), indicating that some topics drew more attention and generated more discussion on KUTUP-L. The percentage of postings intended to generate discussion has also increased and the topics of such postings were in line with those of reflexive ones, further reinforcing the willingness of KUTUP-L members to make it a more dynamic electronic discussion list.

Out of 977 KUTUP-L messages in our stratified sample, 393 (or 40% of all messages) were off-topic. Some of those messages were irrelevant while others consisted of postings of trial messages, virus warnings, announcements of social activities, deaths, and so on. The percentage of off-topic messages rose to 46% in recent years. Table 4 provides descriptive data on off-topic messages. In general, more (53%) than half of off-topic messages were irrelevant (i.e., unrelated with the purpose of the discussion list). Announcements of promotions (16%), deaths (9%), and job ads (9%) consisted of one-third of all off-topic messages.

The anatomy of an electronic discussion list for librarians, KUTUP-L

Table 3: KUTUP-L messages by topics

Subjects	Subject 1	Subject 2	Total	%
100 Professional issues	39	4	43	7
101 Library Association's activities	30	0	30	5
The Professions total (100)	69	4	73	12
300 Publishing	18	6	24	4
400 Education in LIS	12	2	14	2
600 Analysis of LIS	1	0	1	0
701 Inter-library loan activities	59	20	79	12
702 Collections	57	6	63	10
703 Information or Reference Services	2	0	2	0
704 User education	1	0	1	0
705 Library Buildings or Facilities	12	0	12	2
706 Library Administration or Planning	10	0	10	2
707 Library Automation	22	2	24	4
708 Other Library or Information Service	14	1	15	2
LIS Service Activities total (700)	177	29	206	32
801 Cataloguing	14	3	17	3
802 Classification or Indexing	3	1	4	1
803 Information Retrieval	4	5	9	1
804 Bibliographic Databases	5	1	6	1
805 Databases	9	2	11	2
Information Retrieval total (800)	35	12	47	7
901 Information Dissemination	7	1	8	1
905 Information Use	2	0	2	0
906 Information Management	4	0	4	1
Information Seeking total (900)	13	1	14	2
1001 Scientific or Professional Publishing	10	0	10	2
1003 Other Aspects of Scientific or Professional Communication	101	12	113	17
Scientific and Professional Communication total (1000)	111	12	123	19
1100 Other LIS Aspects	148	0	148	23
Total	584	66	650	100

Table 4: Off-topic KUTUP-L messages

Types of messages	1995-99	2000-04	2005-08	Total	Avg (%)
Irrelevant messages	47 (23)	78 (37)	84 (40)	209 (100)	53
Announcements of promotions, congrats, etc.	6 (10)	18 (29)	39 (62)	63 (101)	16
Announcements of deaths, etc.	0 (0)	9 (24)	28 (76)	37 (100)	9
Job Ads	1 (3)	15 (44)	18 (53)	34 (100)	9
Trial messages	3 (21)	6 (43)	5 (36)	14 (100)	4
Messages on social activities	3 (21)	6 (43)	5 (36)	14 (100)	4
KUTUP-L	5 (42)	6 (50)	1 (8)	12 (100)	3
Virus warnings	2 (20)	7 (70)	1 (10)	10 (100)	3
Total	67	145	181	393	101

Note: Figures in brackets refer to percentages. Some totals are not equal to 100% due to rounding errors.

We calculated the productivity of 34 authors in our sample who posted six or more messages by dividing the number of off-topic messages by the total number of messages (both off- and on-topic) sent by each author. The average productivity was 60%, perfectly in line with the percentage of on-topic messages.

5. Conclusion

As an unmoderated discussion list since its inception in 1991, KUTUP-L seems to have an impact on professional lives of many Turkish librarians in that they use it as a venue to ask questions, share news and information with their colleagues, follow up current developments and discuss professional as well as social issues. The number of messages and unique authors contributing to the list has increased considerably, indicating that KUTUP-L has become a popular and dynamic discussion list.

Although a wide variety of subjects have been discussed on KUTUP-L, the percentage of reflexive messages aimed at discussing a subject or commenting on a professional issue is comparatively low (24%). A more comprehensive study on the subjects of KUTUP-L postings is in order. However, the relatively high (40%) percentage of off-topic KUTUP-L messages might have discouraged its further use, as library and information professionals may not wish to spend their precious time sifting through irrelevant postings. Screening off-topic messages before distribution may help in this respect but this would tax the list owner's time and resources further.

The anatomy of an electronic discussion list for librarians, KUTUP-L

Most of the KUTUP-L postings were authored by a relatively few list members. KUTUP-L authors' productivity data conform to Pareto and Price laws but not in accordance with Lotka's Law. KUTUP-L has more list members contributing two or more messages to the list.

We hope that the results of bibliometric and content analysis of KUTUP-L postings will be helpful in studying professional and scientific communication patterns of library and information professionals in a larger context.

References

- [1] WILDEMUTH, B. et al. What's everybody talking about: message functions and topics on electronic lists and newsgroups in information and library science. *Journal of Education for Library and Information Science*, 38 (2), 1997, p. 137-156.
- [2] BAR-ILAN, J; ASSOULINE, B. A content analysis of PUBYAC- A preliminary study. *Information Technology and Libraries*, 16, 1997, p. 165-174.
- [3] EDWARDS, MM. "A content analysis of the PUBYAC discussion list". (Unpublished Master's Thesis). University of North Carolina at Chapel Hill, Chapel Hill, NC, 1999.
- [4] SCHOCH, NA; SHOOSHAN, SE. Communication on a listserv for health information professionals: uses and users of MEDLIB-L. *Bulletin of the Medical Library Association*, 85 (1), 1997, p. 23-32. Available at: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC226219/> (April 2010).
- [5] BELLACK, AA. et al. *The language of the classroom*. New York: Teachers College Press, 1966.
- [6] BELLACK, AA. Methods for observing classroom behavior of teachers and students. Presented at a conference sponsored by Paedagogisches Zentrum, Berlin, November 12-15, 1968.
- [7] CHRISTIE, CA; AZZAM, T. What's all the talk about? Examining EVALTALK, an evaluation listserv. *American Journal of Evaluation*, 25 (2), 2004, p. 219-234. Available at: doi: 10.1177/109821400402500206 (April 2010).
- [8] NICKLAS, EW. "The Trombone-L e-mail discussion list : an analysis of its users and Content." (Unpublished PhD Thesis). The University Of Missouri, 2002.
- [9] PIBURN, MD; MIDDLETON, JA. Listserv as journal: computer-based reflection in a program for pre-service mathematics and science teachers. Paper presented at the *International Conference on Science, Mathematics and Technology Education (Hanoi, Vietnam, January 6-9, 1997)*.

The anatomy of an electronic discussion list for librarians, KUTUP-L

- [10] MAY, A. Automatic classification of e-mail messages by message type. *Journal of the American Society for Information Science*, 48, 1997, p. 32-39.
- [11] BERMAN, Y. Discussion groups in the Internet as sources of information: the case of social work. *Aslib Proceedings*, 48 (2), 1996, p. 31-36. Available at: <http://www.aslib.co.uk/proceedings/1996/feb/1.html> (April 2010).
- [12] TRUESWELL, RL. Some behavioral patterns of library users: the 80/20 rule. *Wilson Library Bulletin*, 43, 1969, p. 458-461.
- [13] DE SOLLA PRICE, D. A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27, 1976, p. 292-306.
- [14] EGGHE, L. *Power laws in the information production process: Lotkian informetrics*. Amsterdam: Elsevier, 2005.
- [15] EGGHE, L; ROUSSEAU, R. *Introduction to informetrics: Quantitative methods in library, documentation and information science*. Amsterdam: Elsevier Science Publishers, 1990. Available at: <http://uhdspace.uhasselt.be/dspace/handle/1942/587> (April 2010).
- [16] KUPERMAN, V. Productivity in the Internet mailing lists: A bibliometric analysis, *Journal of the American Society for Information Science and Technology*, 57, 2006, p. 51-59.
- [17] JARVELIN, K; Vakkari, P. Content analysis of research articles in library and information science., *Library and Information Science Research*, 12, 1990, p. 395-421.
- [18] JARVELIN, K; VAKKARI, P. Evolution of library and information science 1965-1985: Content analysis of journal articles. *Information Processing Management*, 29, 1993, p. 129-144.