Predatory E-mails Management through Analysis of Non-Academic Scientist Account: International Challenge in Scientific Publishing

Akademik Olmayan Bilim Adamlarının Hesaplarının Analizi Yoluyla Yıkıcı E-posta Yönetimi: Bilimsel Yayıncılıkta Uluslararası Zorluk

Mostafa Essam EISSA
M.Sc., Pharmaceutical Research Facility, Cairo – Egypt,
mostafaessameissa@yahoo.com
https://orcid.org/0000-0003-3562-5935

Engy Refaat RASHED
Assist. Prof., National Centre for Radiation Research and Technology, Cairo – Egypt,
engyrefaat@yahoo.com
https://orcid.org/0000-0002-6593-378X

Dalia Essam EISSA
Prof. Dr., Royal Oldham Hospital, Oldham - United Kingdom,
eissade@yahoo.com
https://orcid.org/0000-0002-6340-8973

ABSTRACT

Scientific distant education and learning have become fast and straightforward in the internet era for information dissemination and sharing through e-publishing. However, this breakthrough has a severe drawback as misinformation and inaccurate scientific pieces of literature might be delivered to the audience. One of the most significant sources of this kind of flaw is the fraudulent publishing that is progressively growing. We aimed in this study to analyze predatory behavior communication through investigation of the e-mail account of a non-academic researcher. The study involved a text analysis of the identified different spam mail types. First, we have transformed e-mails that were considered spam into an Excel database. Then, categorization was executed to identify types of scientific fraud and their rate of delivery in the account during a definite time frame. Common keywords and country names were analyzed to spot the predatory markers. The study showed five categories of scientific predatory e-mails with a remarkable alarming rate of occurrence for invitations received for contribution to journal publishing. More than half of the received e-mails from these types involved country names embracing India, USA and UK. Among the screened 50 words/keywords from suspicious e-mails, only six of them contributed by 60% in frequency of occurrence. The spotting of patterns using text analysis coupled with Pareto charting would be helpful in e-mail management for researchers and scientists, especially non-academicians who do not have a frame of reference in scientific publishing and any qualified mentors or librarian available who could guide the publication process.

ÖZET

Bilimsel uzaktan eğitim ve öğrenme, internet çağında e-yayıncılık yoluyla bilginin yayılması ve paylaşılması için hızlı ve anlaşılır hale geldi. Bununla birlikte, yanlış bilgi ve yanlış bilimsel literatür parçaları izleyiciye iletilebileceğinden, bu atılımın ciddi bir dezavantajı vardır. Bu tür kusurların en önemli kaynaklarından biri, giderek büyüyen hileli yayıncılıktır. Bu çalışmada, akademik olmayan bir araştırmacının e-posta hesabını inceleyerek saldırgan davranış iletişimini analiz etmeyi amaçladık. Çalışma, tanımlanan farklı spam posta türlerinin bir metin analizini içeriyordu. İlk olarak, spam olarak kabul edilen e-postaları bir Excel veritabanına dönüştürdük. Ardından, bilimsel dolandırıcılık türlerini ve bunların belirli bir zaman diliminde hesapta teslim edilme oranlarını belirlemek için kategorizasyon yapıldı. Yırtıcı işaretleri tespit etmek için yaygın anahtar kelimeler ve ülke adları analiz edildi. Çalışma, dergi yayıncılığına katkı için alınan davetler için dikkate değer endişe verici bir oranda ortaya çıkan beş bilimsel yırtıcı e-posta kategorisi gösterdi. Bu türlerden gelen e-postaların yarısından fazlası Hindistan, ABD ve İngiltere'yi kapsayan ülke adlarını içeriyordu. Şüpheli e-postalardan taranan 50 kelime/anahtar kelimeden sadece altı tanesi görülme sıklığına %60 oranında katkıda bulunmuştur. Pareto şemasıyla birleştirilmiş metin analizi kullanılarak kalıpların tespit edilmesi, araştırmacılar ve bilim adamları için, özellikle bilimsel yayıncılıkta bir referans çerçevesi olmayan akademisyen olmayanlar ve mevcut herhangi bir kalifiye danışman veya kütüphaneci için e-posta yönetiminde yardımcı olacaktır.

1. INTRODUCTION

The advantage of scientific publishing in the internet era is undeniable. The dissemination of information through the World Wide Web (WWW) has become instantaneous and in different media forms (Kibbee, 2022). The writers and the audience can communicate effectively in no time and share knowledge and data (UNAB, 2022). E-learning has become a convenient and efficient way to communicate lessons and studies regardless of the distance. However, these pros did not come without cons. Malicious, fake and predatory sites could be found swarming the internet pretending to be legitimate trustable scientific sites harming researchers and readers in different sly ways (Laine and Winker, 2017:285-291).

One of the most dangerous traps for inexperienced researchers or scientists is the affinity to the unprofessional fraud and predatory e-mails that are hiding beneath the science cover and directing them towards stunning and fascinating-looking sites that are deemed scientifically appealing while providing dishonest contents and data (Duc et al., 2020:318). Accordingly, the victim might be engaged either in receiving information, contributing to bogus conferences, embroilment as editorial/review board, writing book/chapter and/or publishing his/her work in this kind of spam activity (Mercier et al., 2017:104-108).

There is a great challenge facing scientific publishing with the growing number of academicians, researchers and scientists around the globe in relation to the available reputable journals (Solms, 2016). This problem in turn put great pressure on the reasonably and highly impacted journals from the number of received works to be considered for publication (Huisman and Smits, 2017:633–650; Jawaid, 2021:616-620). In these circumstances, the journal board might Returning the manuscript or rejection of the article are disappointing and making the authors and researchers vulnerable for manipulation by frauds through spam electronic messages (Agathokleous, 2021).

The present work aimed to analyze the pattern of predatory invitation messages in science through examination of spam e-mails in the account of independent and non-academic research scientists in the healthcare industry. Since Academics usually work within reference frames in scientific publishing backed by mentors and guiding experienced supervisors, independent and non-academic researchers and scientists may be more vulnerable and prone to these predators. Clues and keywords might be helpful in these situations for the recipients to identify the most likely untrustable sources of these messages. The work will include types of spam traps that are covered up under the name of science and their frequency of delivery. Screening of the rate of occurrence of special words or expressions in the text body of the delivered messages where the presence of such pattern in messages might help the young researchers to swiftly blacklist these e-mails, avoid them and classify them as spam e-mails if they are not already in the spam folder.

Almost every day, researchers and scientists in the healthcare field may come by reading electronic messages that look appealing and requesting their contribution to the scientific field. These platforms are using non-professional and inaccurate processes coupled with misinformation for the sack of their own benefits. Apart from the authors who intentionally and knowingly communicate with these unethical sites, there are victims who may fall into the trap of these predatory senders, especially those who have no experience, mentor or guide such as young industrial researchers.

The herein study provides text analysis of suspicious e-mails from untrustworthy sources to detect the most common characteristics and patterns of these malicious messages using an electronic account of a non-academic young scientist in the medical sector of the healthcare field. The study involves a combination of Excel extracted and processed databases along with statistical process analysis software to detect the primary key aspects in the text analysis of the suspect predatory messages pertaining to the spam folder. This work will bring the focus of the e-mail recipient to the detection of the common and pronounced manner of the message text received from the supposedly predatory sources.

3. RESEARCH MATERIALS AND METHODOLOGY

The research design involved short-term monitoring of suspicious e-mails which is limited by the life span of messages in the spam folder. The study design includes combining text analysis with Statistical Process Control (SPC) tools of Pareto diagram and process-behavior (trending or control) charts, in addition to a Pie chart and supportive political map. The study covers a time interval of about 32 days from 18 April 2022 to 20 May 2022.

- *Participants*: The account of the non-academic scientific researcher in the healthcare industry without previous reference bookmark and mentoring guide in the publishing field.

- *Data Collection and Analysis*: An e-mail account of a non-academic scientist in the healthcare industry field was screened for spam messages. The e-mail messages were exported into Excel sheets using a conversion transaction for flexibility for further processing. This could be conveniently achieved using programs such as Yahoo Backup Wizard version 6 (Brown, 2022).

- *Processing through Microsoft Excel*: Data transfer into an Excel sheet was processed, filtered and segregated. Spam e-mails were examined for classification into reasonable groups. World countries were added as a separate reference column range to be used for country name search within the text range under study. Excel functions were used to analyze text through separated message sections in columns (Yap, 2016; MonkeyLearn, 2022). The core Excel function used in this screening was as the following:

fx = SUM (LEN (Database Cells Range) – LEN (SUBSTITUTE (Database Cells Range, Searchable Text Cells Range, "")))/LEN (Searchable Text Cells Range)

- *Implementation of Pareto Principle*: Pareto analysis was used to screen subject and body texts in terms of the involved country names and characteristics and/or repeated words/terms used by the senders. Surveillance of suspect spammers' e-mail addresses was investigated and arranged using the Pareto concept (Parker, 2014; Eissa, 2019a:1-2). Pareto chart was constructed by descending ordering of dataset based on the count and the individual and cumulative contributions fractions or percentages were calculated. A supportive and illustrative map generation was conducted using countries' names with the color intensity proportional to the country's word frequency of repetition. It gives an indication of the most common words that are used in predatory mail to be spotted by the recipient victim.

- *Application of T Control Chart*: Finally, rare event control charts were implemented to demonstrate the frequency and the probability of event occurrence for each category delivery within the recipient account. Weibull fit in the trending chart is described by both shape and scale parameters (Minitab Support, 2022). The charts were drawn using intervals between the reception time of the chronologically arranged e-mails as a whole and as segregated groups based on their types using the filtering function. The control chart shows the window of time as control limits (calculated according to equations used by software manual) and the center line representing 50th centile of the distribution. Tests for assignable causes of variations in time intervals were selected to determine unusual periods of the events frequencies. These alarms are not limited to unusual excursions outside threshold windows (type "1") but also to those showing abnormal pattern within the control limits.

- *The Scale*: The validity and reliability processes in data collection and analyses were ensured using the previous statistical tools (Pareto (for frequency) and rare event chart (for time)) in programs by combining effective data extraction and processing followed by statistical-text analysis and investigation to illuminate the predatory common words, expressions and pattern in the messages related to scientific publishing.

4. THE FINDINGS AND DISCUSSION

The subjection of the processed dataset to the Parto analysis yielded a distinguishable pattern. Spam electronic messages under the science cover embrace more than half of the suspect e-mails in the spam folder within 30 days as could be seen in the Pareto diagram in Figure 1. Currently, five predatory sections could be identified as scams in science: paper publication, events joining, involvement in predatory journal board, book or book chapter publication and proofing, editing and translation services.

Invitation of the publication into predatory and/or fake journals/publishers constituted a significantly large proportion that accounted for almost 80% of the whole predatory act in the science message section and about 40% of the overall spam list. Bogus conferences and fake webinars were the next in the rate of invitation list showing more than one-fourth in the frequency of appearance affecting the scientific section as could be illustrated in the Pie chart of Figure 1. Other invitations for book chapter publishing, editorial and/or reviewer board joining and manuscript processing services demonstrated less than 13% collectively of the science-affected fraud messages. It should be noted that "*Total Spam*" messages referred to all other suspicious e-mails that could be found in the spam folder but are not related to fraud in the scientific field.

Figure 1a. Categories of Suspicious E-Mails Found Within 30 Days of Account Monitoring Activity
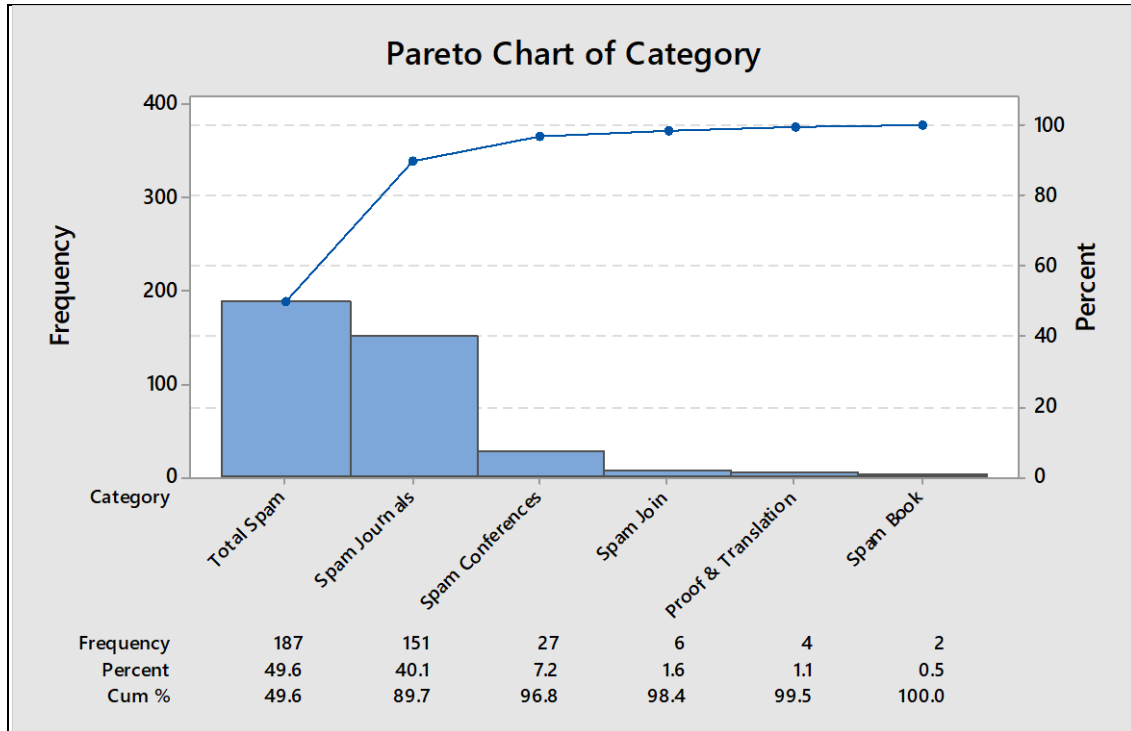
Pareto daigram of science-related and total of other non-scientific suspecious messages (upper graph) and pie plot of the contribution of different spam categories in science-related messages (lower graph)

Figure 1b. Contribution of Different E-Mails in Spam
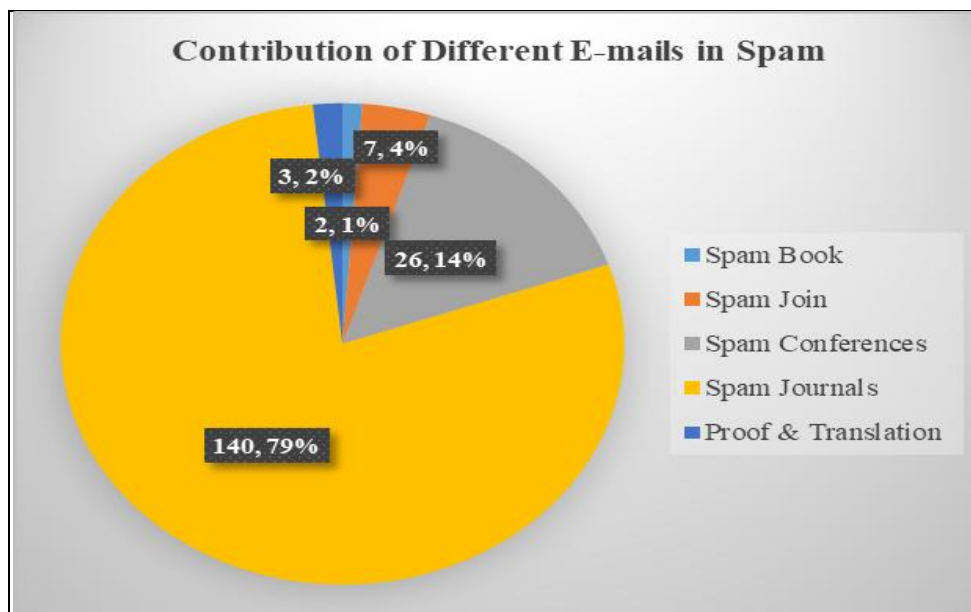
Pie chart demonstrates the contribution of each segement of predatory category in the observed spam messages related to science.
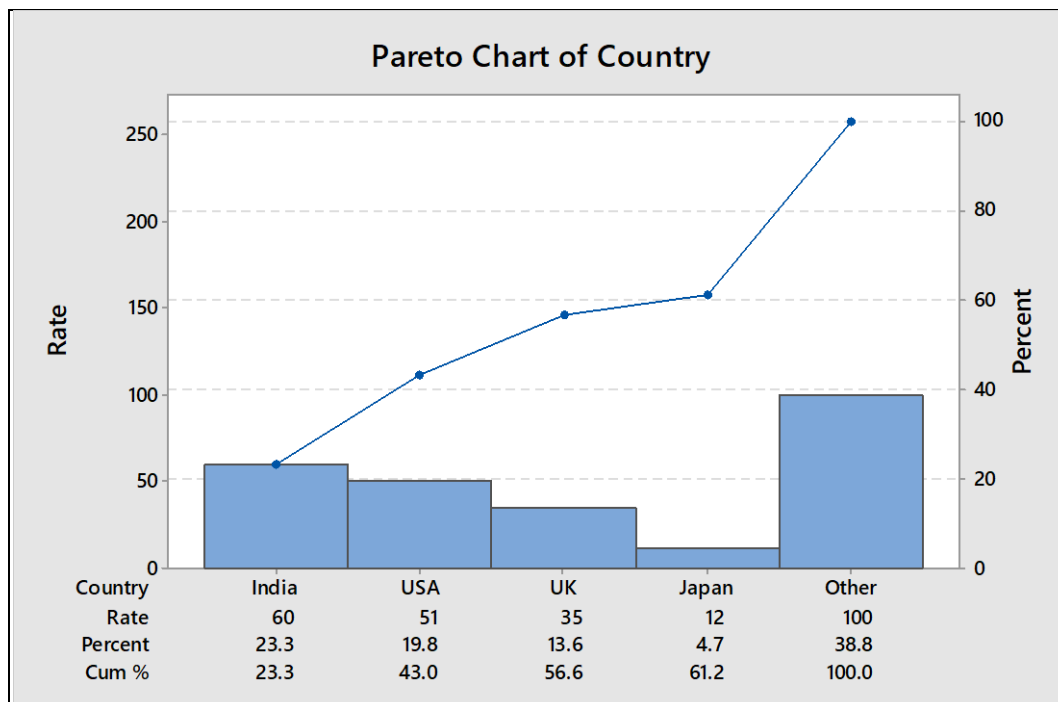
Text Analysis by Country Name: Monitoring of the rate of mentioning of country name within the context of the message parts such as within the subject of the mail body was demonstrated using a world map in Figure 2.

Figure 2a. Map of Counries Involved in the Scientific E-Mail Spams (Upper Graph) and Pareto Diagram of the Main Names of Nations with Highest Frequency of Apprearance in the Messages (Lower Graph)
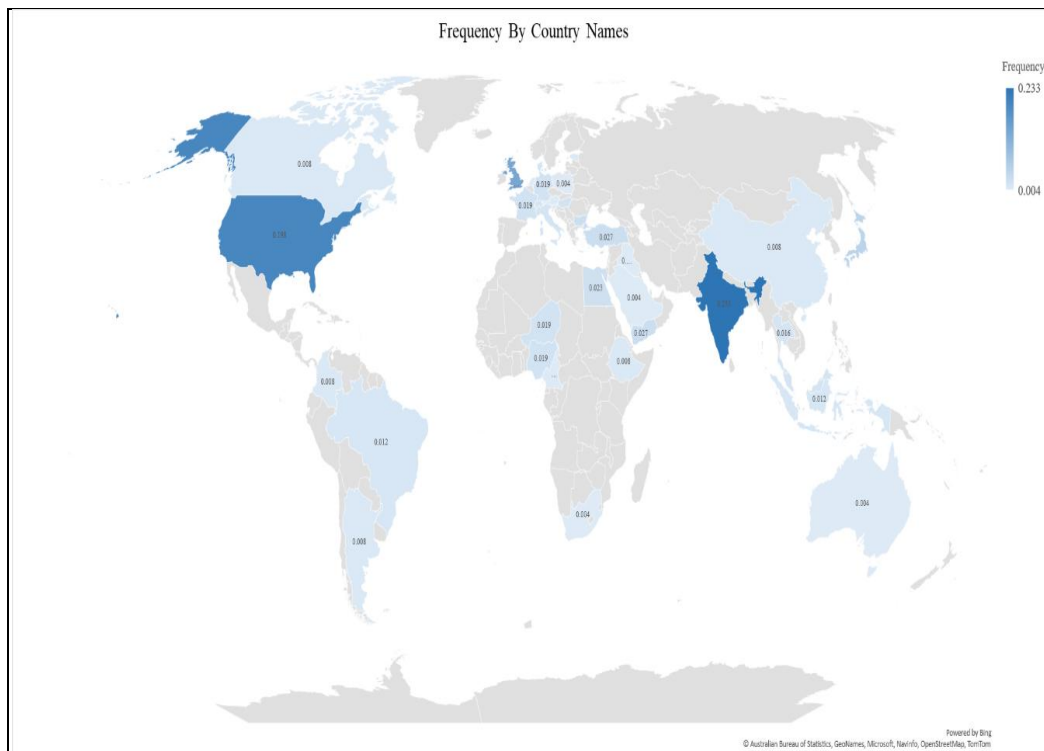
The frequency is proportional to the intensity of the blue color. Screening of the messages showed that they were involved referring to the country of the author(s), editor(s), reviewer(s), journal/publisher name and/or country of origin. They were included also as places for the attendance of the conferences.

Figure 2b. Pareto Chart of Country

| Country	India	USA	UK	Japan	Other
Rate	60	51	35	12	100
Percent	23.3	19.8	13.6	4.7	38.8
Cum %	23.3	43.0	56.6	61.2	100.0

More than 60% of these country texts were spotted to relate to India, USA, UK and Japan as could be arranged in descending order from the Pareto graph.

Figure 3 shows particular words that were common in e-mails under science cover. These words under focus were arranged using a Pareto chart to show 80% contribution by only 15 items. To demonstrate their importance, the senders used words such as "*International Journal*", "*Advances in*" and "*Global*" in their invitation for either conferences and/or journals with a contribution of 24%. Words that lure the victim to fall into the trap by faking "*Impact factor*", "*peer-reviewed*" and "*high quality*" words with a contribution of 21%. Attempts to flatter and courtesy the recipient using words such as "*Call for*", "*welcome you*", "*busy schedule*", "*precious*", "*cordially invite*", "*Looking forward to*", "*Dear Dr*" and "*Greeting*" were found frequently at 34% rate in the messages. At the end of some e-mails (2%) a link to a choice of "want to unsubscribe" is offered despite the recipient did not subscribe in the first place. Focusing on the subject yielded specific pattern as could be found in Figure 4 with only six words and/or expressions were evident by more than 60%, notably the occurrence of words "*submit*" and "*open access*" by a frequency over 51%.

Text Analysis of the Senders' E-mail Addresses: Investigation of the usernames of the senders revealed common words and categories in Figure 5. Square brackets ([,]) were used for broad grouping such as any science field claimed e.g., medicine, pharmacy, engineering, …etc. was denoted as [Field], sender names [Name] and journal name [Journal]. The field of science or subfield, the sender's name and the journal name were 35,4% of all screened usernames. Words such as "*editor*", "*submission*", "*marketing*", "*info*" and "*admin*" were used as usernames for 25,4%.

Figure 3. Common Keywords in the Scientific Spam E-Mails That Might Pinpoint to the High Possibility of the Predatory Act
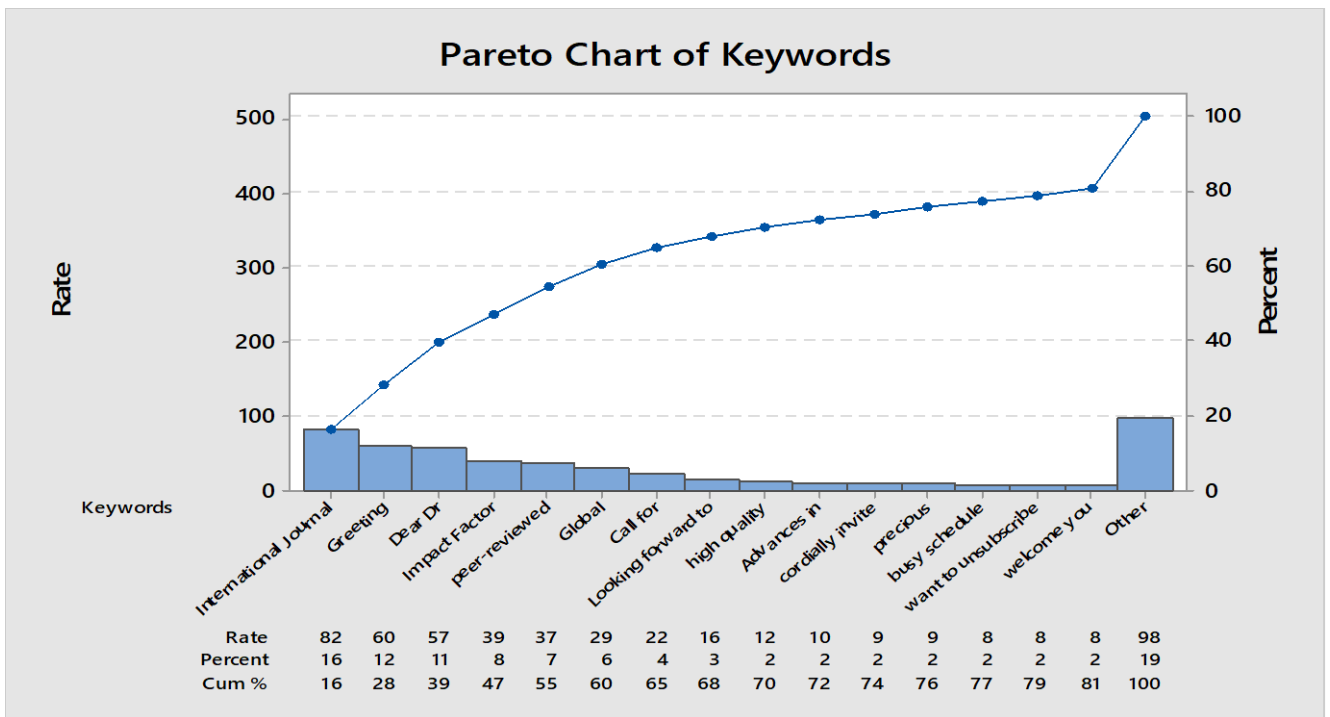
Table 1. Subject Analysis Using Text of the Spam E-Mail Message as Marker

Subject	F	%	C%	Subject	F	%	C%
submit	196	43.27	43	Peer-reviewed Journals	1	0.22	85
open access	39	8.61	52	call from [Journal Name]	1	0.22	85
Manuscript Submission	21	4.64	57	Friendly reminder mail from Editorial Office	1	0.22	85
Article Submission	8	1.77	58	Regarding the article Submission	1	0.22	85
[Journal Name]	7	1.55	60	⏰Final Call for Manuscript/Article Submission: [Edition] - [Date]	1	0.22	85
Come to [Name] Presents [Country] [Field] Leaders Conclave on [Date] in [City]!	6	1.32	61	Manuscript for Publicaton	1	0.22	86
opinion	5	1.10	62	Accepting Articles for the Current Issue	1	0.22	86
Interact with stalwarts of [Field] at The [Edition] of [Country] [Field] leaders conclave.	4	0.88	63	Ready to find the perfect journal?	1	0.22	86

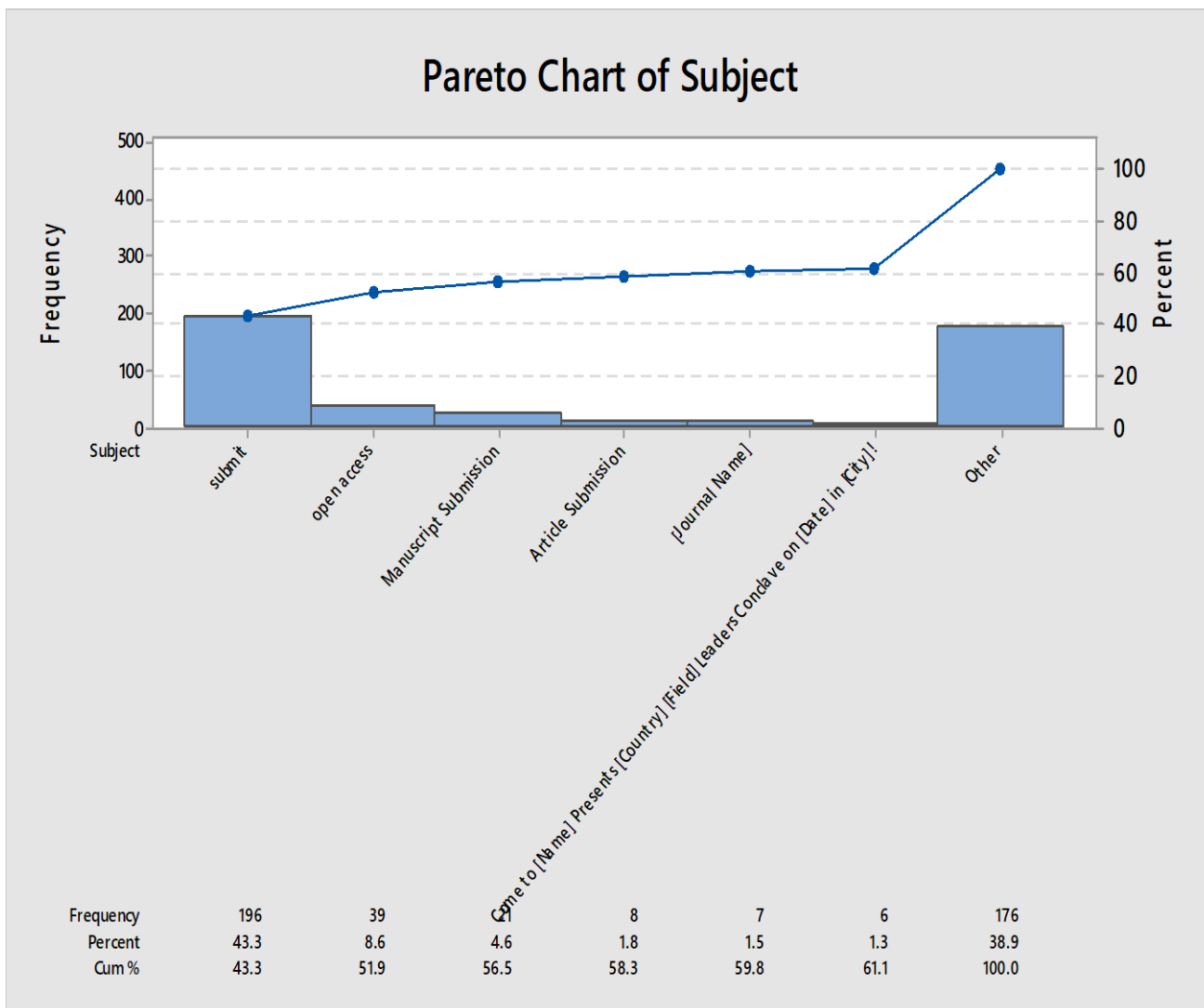|						
Follow Up E-mail: Manuscript Submission Reminder- [Journal Name] [Edition]	5	1.10	64	Affirmative Response: [Journal Name]	1	0.22	86
[Authors Published Work]	3	0.66	65	Endorse your unpublished article in [Field]	1	0.22	87
Academic Journal of Medical Sciences	3	0.66	66	Submit Now: [Journal Name]	1	0.22	87
Selected as Book Chapter [Authors Published Work]	3	0.66	66	Are You Interested to Submit Paper ?	1	0.22	87
Call for paper: [Journal Name]	3	0.66	67	Dear Dr. xx yy ww - Inaugural Issue - [Journal Name]	1	0.22	87
Call for paper: [Journal Name], [Edition]	3	0.66	68	[Journal Name] : Best Article [IF] Best Article Invite for [Date]	1	0.22	87
Your Article	2	0.44	68	Submit Your Valuable Research: [Journal Name]	1	0.22	88
discovery	2	0.44	68	[Free Live Event] JOURNAL SUBMISSION: The SECRETS of SUCCESS!	1	0.22	88
INVITATION TO ISI-INDEXED Int. CONFERENCES, [City], [Date]	2	0.44	69	Invitation For Papers ([Edition], [date], [IF])	1	0.22	88
[Journal Name] Journals | Special Issue | Open Access	2	0.44	69	Call for scholarly articles	1	0.22	88
[Journal Name] Online Submission - Research Paper Publication	2	0.44	70	Dr. [Name] :: Editorial Board invites you	1	0.22	89
Proofreading Service	2	0.44	70	Awaiting for your Submission	1	0.22	89
English Proofreading	2	0.44	71	Re: [Author Published Work]	1	0.22	89
At last, the [Field] Possibilities of [Field] has arrived !!	2	0.44	71	[Journal Name] - [Date] - Reg.	1	0.22	89
Come and explore the possibilities and potential of [Field] at the [Field] Possibilities of [Field] on [Date].	2	0.44	72	[Journal Name]: Best Article [IF] Article Invite for [Date]	1	0.22	89
Dr. xx yy ww: Grab the occupancy of Editorial/Reviewer Board - [Journal Name]	2	0.44	72	Connect with Key [Field] Professionals @ [Conference Name] from [Date]	1	0.22	90
[Conference Name] [Date] | Speaker Invitation	2	0.44	72	Invitation as Committee Member	1	0.22	90
Publish your research work in [Journal Name]	2	0.44	73	xx yy ww - new articles	1	0.22	90
Call for Manuscript	2	0.44	73	Submissions Open for Upcoming Issue- [Journal Name]	1	0.22	90
[Days] to go for the second edition of [Country] [Field] leaders conclave	2	0.44	74	We humbly invite you to submit your interesting article	1	0.22	91
Dr. xx yy ww: Take possession of Editorial/Reviewer Board - [Journal Name]	2	0.44	74	Explore the global trends in [Field] world	1	0.22	91
publish your article	2	0.44	75	xxyyzzww, Avail Speaker Slots at [Conference Name] [Date]	1	0.22	91
Dr. [Name] :: Well reputed Indexing of World.	2	0.44	75	[Author Published Work]	1	0.22	91
Dr. [Name]	2	0.44	75	Call for Original Articles: Invitation to Publish with Us : [Author Published Work]	1	0.22	91
Submit manuscript: [Journal Name] [Edition]	2	0.44	76	Globalize your manuscript	1	0.22	92
submit paper -[Date]	1	0.22	76	Submit Abstract to [Conference Name] [Date]	1	0.22	92
Announcements ([Journal Name] Editorial Board Member)	1	0.22	76	Request your prompt response| [Conference Name] [date]	1	0.22	92
Reg: Article Submission	1	0.22	77	Hi ww xx yy, Call for papers & contributions	1	0.22	92
You are selected | [Publisher] Journals [Field] Domain	1	0.22	77	Dear Dr. xx yy ww - Explore your research work - Most peer-reviewed journal!	1	0.22	92
[Field] submission	1	0.22	77	[Journal Name] - ISI | doi publications for [Date] | [Edition]	1	0.22	93
[Journal]: Submit any unpublished works ✍	1	0.22	77	Call For Papers [Date] [IF]	1	0.22	93
Call for Article-Your mnauescript hasbeen Publish	1	0.22	77	Accepting submissions for Original Articles, Reviews, Short Communications, and Case Reports	1	0.22	93
Article Submission: [Journal Name]	1	0.22	78	Original Research Submission	1	0.22	93
Submit Now- [Journal Name]	1	0.22	78	Dear Dr. xx yy ww - Followup Mail: Explore your research work - Most peer-reviewed journal!	1	0.22	94

Manuscript Publication - [Journal Name]	1	0.22	78	Confirm Your Presentation Slot at [Conference Name] [Date]	1	0.22	94
Publish your manuscript- [Journal Name]	1	0.22	78	[Journal Name]: Invitation Letter.	1	0.22	94
Request for Manuscript Submission	1	0.22	79	DR. [Journal Name] :: Rapid Publication Process	1	0.22	94
Hi ww xy, Are you ready to publish?	1	0.22	79	[Conference Name]	1	0.22	94
[Journal Name]: New Issue [IF] Article Invited for [date]	1	0.22	79	100% waive for publication [Journal Name]	1	0.22	95
Dr. [Name] :: Manuscripts Submitted Journal	1	0.22	79	Greetings from [Publisher Name]	1	0.22	95
High-Quality Proofreading & Translation Services	1	0.22	79	Call for Research Paper	1	0.22	95
[Edition]	1	0.22	80	[Journal Name] - Take Part In Journal As EB Member	1	0.22	95
Panel discussion on [Field]	Register for free	1	0.22	80	RE: Your valuable research	1	0.22	96
Notification for upcoming issue [Date]	1	0.22	80	Famed Article Dear Dr. xx yy zz ww	1	0.22	96
Re: xx y. ww	[Authors published Work]	1	0.22	80	Manuscript Submission Confirmation Request - [Journal Name] [IF] [Edition]	1	0.22	96
Requesting to submit any of your unpublished works	1	0.22	81	Dr. [Name] :: Original Research Articles	1	0.22	96
paper publishing	1	0.22	81	Publishing recent On-going Innovative publications on [Field]	1	0.22	96
Reveal your Novel Study on [Field]	1	0.22	81	Call For Papers - The World's Largest Index	1	0.22	97
Submit your unpublished article on [Journal Name]	1	0.22	81	Current issue of [Journal Name] has been released: Congratulations to All	1	0.22	97
[Field] - Welcoming Manuscript Submissions	1	0.22	81	Dr. [Name] :: Peer Reviewed International Journal	1	0.22	97
Dear Doctor. xx yy ww: Follow-up E-mail: Regarding Your Manuscript Submission: [Journal Name] [IF]	1	0.22	82	Submit Article To [Journal Name]	1	0.22	97
Manuscript Submission-[Journal Name]	1	0.22	82	~!@ @@ Low publication fee Journal @@ ~!@	1	0.22	98
xxyyww@yahoo.com: Expert professionals are needed!	1	0.22	82	Manuscript Submission for [Edition]	1	0.22	98
Make submission [Date]	1	0.22	82	Online manuscript submission - Submit Journal Online	1	0.22	98
Hoping to Help: [Journal Name]	1	0.22	83	Collaborate with us by Submitting your Article : [Journal Name]	1	0.22	98
ww xy - Waiting for Your Valuable Submission [Journal Name]	1	0.22	83	globalize your research	1	0.22	98
Remind: Cordially welcome you to be a Chair/Speaker at [Field]-[Date], [City], [Country]	1	0.22	83	[Free live event] [Field] • [Date]	1	0.22	99
Submit your Scientific Research	1	0.22	83	[Journal Name] [Date]	Speaker Invitation	1	0.22	99
Dr. xx yy ww: Seize the moment of Editorial/Reviewer Board Membership - [Journal Name]	1	0.22	83	Reg: We Looking Forward for Your Article Submission	1	0.22	99
Call for Papers	1	0.22	84	Dear Dr. xx yy ww - Want to Explore your research work in the Most peer-reviewed journal? Here is the best choice for you!	1	0.22	99
Call for Article Submission- [Journal Name]	1	0.22	84	[Journal Name]: New Issue [F] Article Invited for [Date]	1	0.22	100
Avoid Waiting Lists	[Conference Name] [Date]	1	0.22	84	Need Research Article	1	0.22	100
/// [Journal Name] ///	1	0.22	84	Invitation mail for Manuscript Submission	1	0.22	100

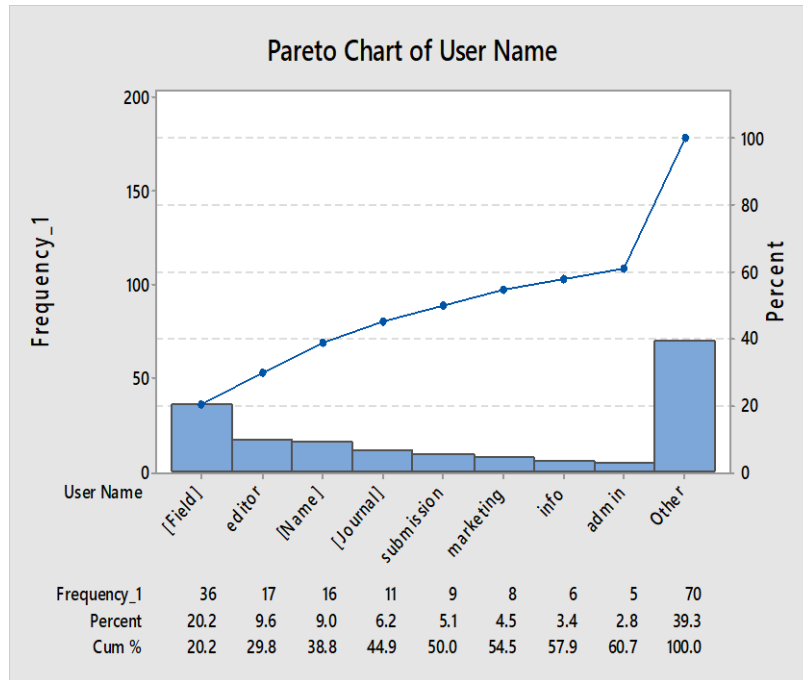|		
Total Rate =		453

Yellow Highlight: Suject title with 60% cumulative frequency   Pink Highlight: Extended subject words reaching 80% cumulative frequency   F: Frequency   C%: Cumulative percent

Implementation of Rare Event Control Chart in Monitoring of Spam E-mails: The T Chart process generates a control chart based on the time intervals between rare event occurrences. It employs the Weibull distribution, which posits that events occur at continuous points in time. This style of chart is commonly used in the life sciences to track the occurrence of incidents. Assessing the rate of receiving suspicious e-mails involving invitations to different science activities would help to determine the probability and magnitude of that event during the survey period. This could be achieved using a rare event control chart of T-type. Only invitation for book or book chapter contribution could not be interpreted due to its low frequency of occurrence (twice/month from the same publisher) with intervals between e-mails of six days, six hours and 34 minutes with an attempt to drag the author(s) through own published work to be incorporated as a book chapter.

Figure 4. Pareto Analysis of Subject Text of Predatory Academic E-Mails

	submit	open access	Manuscript Submission	Article Submission	[Journal Name]	Come to [Name] Presents [Country] [Field] Leaders Conclave on [Date] in [City]!	Other
Frequency	196	39	21	8	7	6	176
Percent	43.3	8.6	4.6	1.8	1.5	1.3	38.9
Cum %	43.3	51.9	56.5	58.3	59.8	61.1	100.0

**Figure 5.** Common Pattern of Usernames in Spam Academic E-Mails

Figure 6. Overall Rate of E-Mails Receiving with Intrusion Nature to the Researcher Account for One Month Period
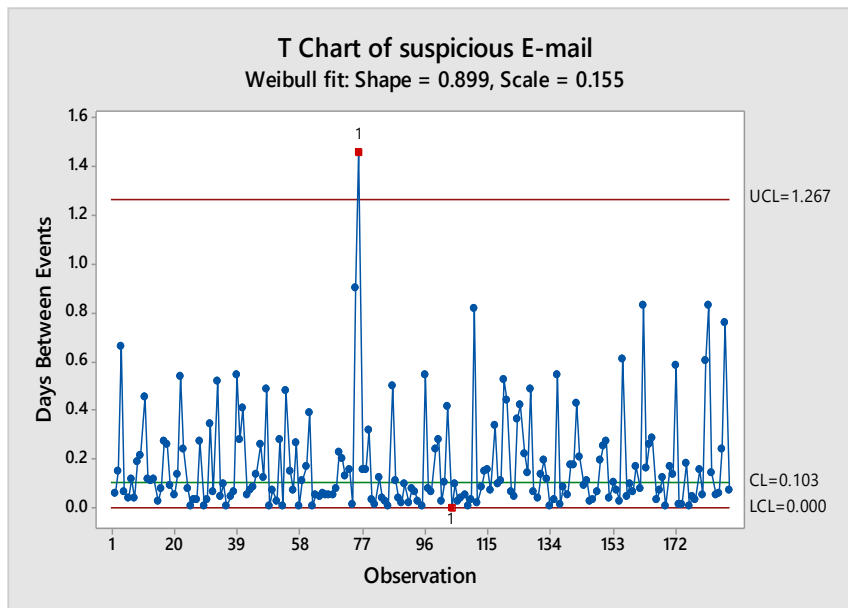

Figures 6 to 10 demonstrated the time interval pattern for the total received spam messages, overall science-related suspicious e-mails, manuscript submission invitations, editorial/reviewer board invitations and invitations to conferences, respectively. Alarming points "red dots" marked unusual intervals between successive events either as abnormally higher or lower times. The days between receiving spam e-mails were expressed using Control Limit (CL) and Upper Control Limit (UCL) and the results from the process behavior charts were (0.10, 1.27), (0.11, 1.50), (0.15, 1.70), (4.82, 14.78) and (0.65, 9.06), respectively.

Figure 7. Rate of Academic E-Mails Receiving with Intrusion Nature to the Researcher Account for One Month Period
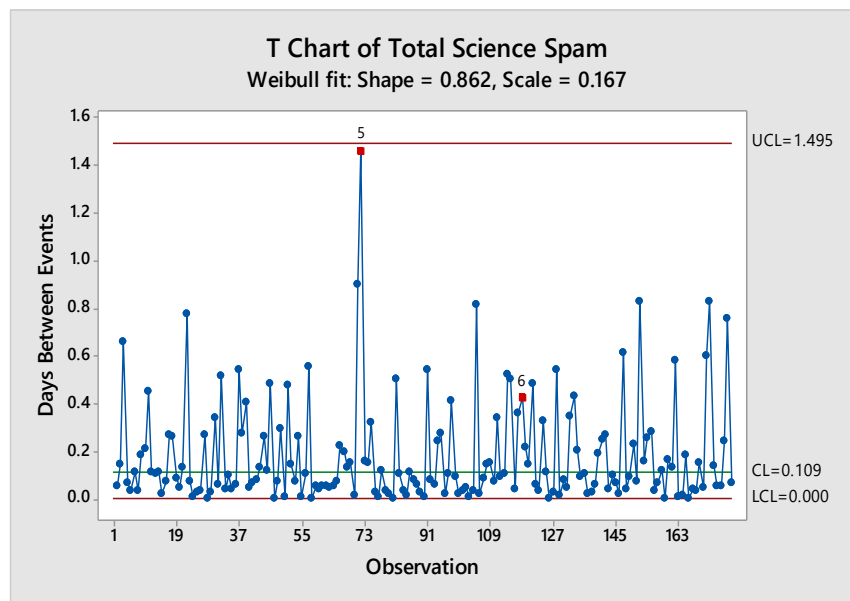

Figure 8. Frequency of Reception of Predatory Journals Invitation for Publication Alone or with Joining the Editorial/Review Board
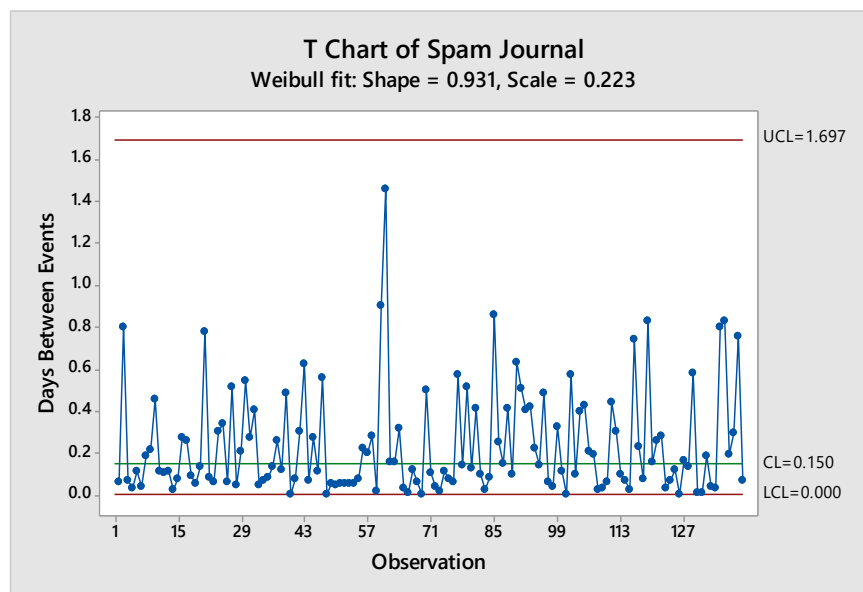

In the internet era, it has become easy to get information and learn huge varieties of topics remotely, notably for researchers and scientists. Scientific papers are rich media for knowledge and innovation. Yet, unethical behavior to ruin this advantage by some immoral individuals has become undeniable. A focused study was conducted on a healthcare professional who is a non-academic researcher in the medical field. A study of the subject study e-mail revealed that it is scarce to find a genuine invitation for scientific publication where that frequency was found at a rate greater than once per six months. Many researchers and scientists may be not aware of predatory and fraudulent behavior and attitude in science in their early careers (Donev, 2020:108-112). Unless there is a guiding reference for scientific publication in their organization, they may not come across this conclusion except after practical experience that might lead to the publishing of their work or contribution through non-trustworthy channels (McGrath, 2022). Librarians and other investigators provide a continuous update on the potential suspicious journals and publishers such as Beall's list (Beall, 2022). In the present situation, when the scientists discovered this list, they tried to withdraw their articles without getting any responses but instead, they found that their work got published.

Figure 9. Editorial/Review Board Invitation Rate from Predatory Publishers Received in E-Mails with One Month Time Frame
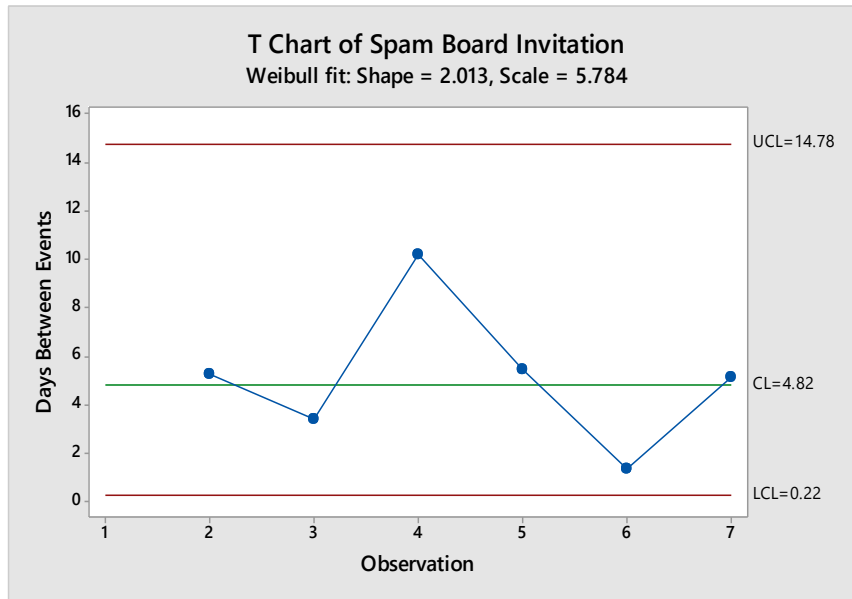
While most investigators propounded that the predatory site owners demonstrated a typical pattern of fraud to gain money without true revision of the work in an unrealistic time, the researchers in the present investigation demonstrated a different experience where the suspect sender did not request fees or charges and even provided a review report with amendments and corrections needed for the work after the reasonable time frame of revision (Eissa, 2018:98-101; Eissa, 2019b:1029; Musick, 2020; Leonard et al., 2021:1009377). Most of the suspicious predatory messages were found in the spam folder, yet several other of these types of e-mails were found also in the inbox. as a means of spoofing the recipient (Mazūra, 2022) mail spoofing is also an addressed Thus, the inbox folder was examined too for the possible predatory issue and the extracted ones were grouped with the spam dataset.

Figure 10. Conference/Webinar Invitations Rate from Suspicious E-Mails Received in the Researcher Account
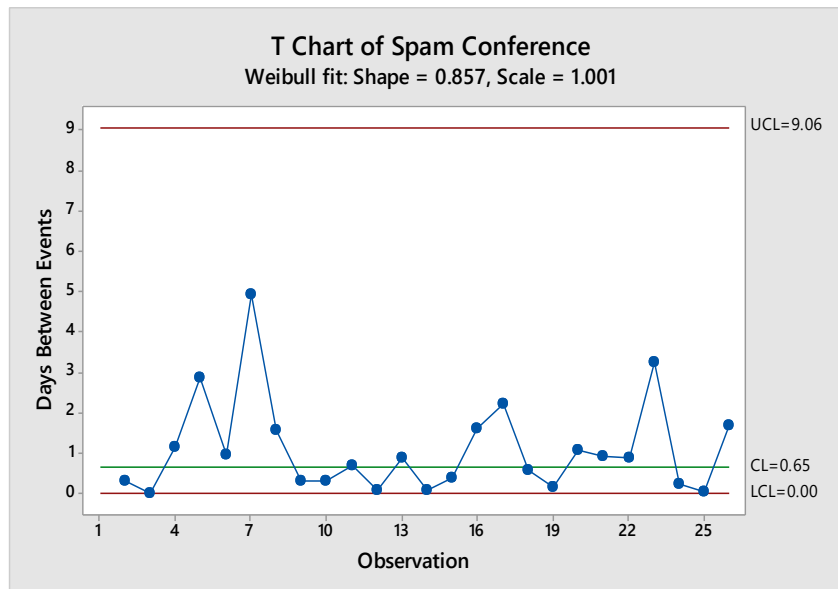
Categorization of spam folder was done to define fraudulent e-mails' theme in science. The life span of the messages in the spam folder is about one month (30 days) which was the time covered in this analysis (Division of Information Technology, 2016; Witman, 2020). It is apparent that this trend of the geographical distribution of the predatory behavior was spotted by other researchers but with some variations (Ibrahim and Saw, 2020:1-6). A strongly evident high rate of spam attacks might be a consequence of the fact that attackers were audacious enough due to their ability to do their actions without any legal prosecution and without bearing any

consequences of their actions (Mallet, 2019; Eissa, 2019c). This could be found obviously from the trending charts as well as Pareto analysis. Screening of keywords in text subject and body of e-mails usually reveals common sense pattern between spam messages senders. Fraud and spam e-mails that have been screened during the account survey showed appealing but clearly non-professional ways to drag young scientists and inexperienced researchers through using some words or expressions (Ken Paxton, 2022). Editorial board invitation alone was the least required segment by the predator activity owners. This might be because they are not primarily concerned with the true and efficient processing of the papers that could be received from the authors. Instead, it could be concluded that they attempt to bring their own benefits, even over their reputation. However, they might attempt to shine their images and credibility by bringing true enthusiastic researchers and scientists into their alleged editorial board on their websites.

5. CONCLUSION

Advancement in the internet era has been accompanied by the growth of predatory publishers and bogus conferences with an attractive-looking appearance and contents that might trap young researchers and inexperienced scientists, especially those who do not have any firm frame of reference or a solid background in the academic e-publishing (USC Libraries, 2022; Vilhelmsson, 2022). This problem is common, notably for those working in developing countries where a lack of a clear, transparent and well-defined scientific publishing regulation is a widespread challenge. In the present study, we have provided a useful means for text analysis to help and guide researchers and scientists, especially those working in the non-academic field to easily spot and identify the suspicious pattern of a predatory act from the senders who also might provide low-quality sites with many broken links. It should be also noted that alternative words and expressions might be used by the spam e-mail senders in their attempts to bring their victims into the trap. However, these attempts usually lead to the same meaning and conclusion that could be drawn from the messages they send to the recipients. The fraud and predatory attitude in the scientific field is a growing challenge and the responsible parties in this action have learned to change their typical actions and responses to accommodate and counteract the new information revealed about them. Apart from willful neglect of the predatory e-publishing avoidance, sincere and diligent science professionals could understand the fraud and bogus behavior and abstain from involvement in this harmful act by realizing the signs of intrusive attitude in the name of science to preserve their knowledge and halt the ruin of the value of the distant scientific e-learning, management and research through the modern era.

YAZAR BEYANI / AUTHORS' DECLARATION:

Bu makale Araştırma ve Yayın Etiğine uygundur. Beyan edilecek herhangi bir çıkar çatışması yoktur. Araştırmanın ortaya konulmasında herhangi bir mali destek alınmamıştır. Makalede kullanılan ölçek için yazar(lar) tarafından ölçeğin orjinal sahibinden izin alındığı beyan edilmiştir. Yazar(lar), dergiye imzalı "*Telif Devir Formu*" belgesi göndermişlerdir. Mevcut çalışma için mevzuat gereği etik izni alınmaya ihtiyaç yoktur. Bu konuda yazarlar tarafından dergiye "*Etik İznine Gerek Olmadığına Dair Beyan Formu*" gönderilmiştir. / This paper complies with Research and Publication Ethics, has no conflict of interest to declare, and has received no financial support. For the scale used in the article, it is declared by the authors that permission was optained from the original owner of the scale. The author(s) sent a signed "*Copyright Transfer Form*" to the journal. There is no need to obtain ethical permission for the current study as per the legislation. The "*Declaration Form Regarding No Ethics Permission Required*" was sent to the journal by the authors on this subject.

YAZAR KATKILARI / AUTHORS' CONTRIBUTIONS:

Kavramsallaştırma, orijinal taslak yazma, düzenleme – **Y1**, **Y2** ve **Y2**, veri toplama, metodoloji, resmi analiz – **Y1**, **Y2** ve **Y2**, Nihai Onay ve Sorumluluk – **Y1**, **Y2** ve **Y2**. / Conceptualization, writing-original draft, editing – **Y1**, **Y2** and **Y2**, data collection, methodology, formal analysis – **Y1**, **Y2** and **Y2**, Final Approval and Accountability – **Y1**, **Y2** and **Y2**.

REFERENCES

AGATHOKLEOUS, Evgenios (2021), "*Mastering the Scientific Peer Review Process: Tips for Young Authors from a Young Senior Editor*", **Journal of Forestry Research**, S.33(1), ss.1-20.

BEALL, Jeffrey (2022), "*Beall's List – of Potential Predatory Journals and Publishers*", **Beallslist.net**, https://beallslist.net (Access Date: 25.07.2022).

BROWN, Chris (2022), "*How to Export Yahoo Emails to Excel? - Using Yahoo to Excel Extractor*", **Turgs Blog**, https://turgs.com/blog/export-yahoo-emails-to-excel/ (Access Date: 27.07.2022).

DIVISION OF INFORMATION TECHNOLOGY (2016), "*Moving a Gmail Message Out of Trash or Spam and Preventing Messages from Being Sent to Spam/Trash | Division of Information Technology*", **it.stonybrook.edu**, https://it.stonybrook.edu/help/kb/moving-misfiled-gmail-messages-out-of-trash-or-spam (Access Date: 29.07.2022).

DONEV, Doncho (2020), "*Predatory in Scientific Publishing - A Burning Issue in Science*", **International Journal on Biomedicine and Healthcare**, S.8(2), ss.108-112.

DUC, Nguyen Minh, HIEP, Dang Vinh, THONG, Pham Minh, ZUNIC, Lejla, ZILDZIC, Muharem, DONEV, Doncho, JANKOVIC, Slobodan M., HOZO, Izet, and MASIC, Izet (2020), "*Predatory Open Access Journals Are Indexed in Reputable Databases: A Revisiting Issue or an Unsolved Problem*", **Medical Archives,** S.74(4), ss.318-322.

EISSA, Mostafa Essam (2018), "*Mobility of Antibiotic Resistance Through PharmaceuticalProducts: A Microbiological Prospective Risk*", **Asian Journal of Biological Sciences**, S.11(2), ss.98-101.

EISSA, Mostafa Essam (2019a), "*Avian Schistosomes Impact on Public Health: A Long-Term Disease Observation*", **Annals of Medical Research**, S.26(8), ss.1723-1724.

EISSA, Mostafa Essam (2019b), "*Food Outbreak: an Overview on Selected Cases over Long-Term Web-Based Monitoring*", **SM Journal of Nutrition and Metabolism**, S.5(1), ss.1029(1-3).

EISSA, Mostafa Essam (2019c), "*Science Predators Should be Eliminated. How to Remove This Proliferating Disease That Attacked Science?*", **Research Gate Corporate Web Page**, https://www.researchgate.net/post/Science_predators_should_be_eliminated_How_to_remove_this_proliferating_disease_that_attacked_science/5df1d231aa1f0931611e42e5/citation/download (Access Date: 26.07.2022).

HUISMAN, Janine and SMITS, Jeroen (2017), "*Duration and Quality of the Peer Review Process: The Author's Perspective*", **Scientometrics**, S.113(1), ss.633-650.

IBRAHIM, Sharaf and SAW, Aik (2020), "*The Perils of Predatory Journals and Conferences*", **Malaysian Orthopaedic Journal**, S.14(2), ss.1-6.

JAWAID, Shaukat Ali (2021), "*Problems Faced by Researchers and Pressure on Impact Factor Journal Editors*", **Pakistan Journal of Medical Sciences,** S.37(3), ss.616-620.

KIBBEE, Jo (2022), "*The World Wide Web as an Information Resource: Pitfalls and Potential*", **Simmons Universiy Corporate Web Page** (E-Article), http://web.simmons.edu/~chen/nit/NIT'96/96-151-Kibbee.html (Access Date: 01.08.2022)

LAINE, Christine and WINKER, Margaret A. (2017), "*Identifying Predatory or Pseudo-Journals*", **Biochemia Medica**, S.27(2), ss.285-291.

LEONARD, Michelle, STAPLETON, Suzanne, COLLINS, Perry, SELFE, Terry Kit and CATALDO, Tara (2021), "*Ten Simple Rules For Avoiding Predatory Publishing Scams*", **PLOS Computational Biology**, S.17(9), ss.(e1009377).

MALLET, James (2019), "*Science Predators Should be Eliminated. How to Remove This Proliferating Disease That Attacked Science?*", **Research Gate Corporate Web Page**, https://www.researchgate.net/post/Science_predators_should_be_eliminated_How_to_remove_this_proliferating_disease_that_attacked_science/5dec1ed53d48b7e74b7aa859/citation/download (Access Date: 02.08.2022).

MAZŪRA, Justinas (2022), "*What is Email Spoofing & How to Protect Yourself?*", **Cybernews**, https://cybernews.com/secure-email-providers/email-spoofing/ (Access Date: 03.08.2022).

McGRATH, Naomi (2022), "*Why and How to Avoid Publishing in a Predatory Journal*", **ifis.org Corporate Web Page**, https://www.ifis.org/en/research-skills-blog/why-and-how-to-avoid-publishing-in-a-predatory-journal (Access Date: 04.08.2022).

MERCIER, Eric, TARDIF, Pier-Alexandre, MOORE, Lynne, SAGE, Natalie Le and CAMERON, Peter A. (2017), "*Invitations Received from Potential Predatory Publishers and Fraudulent Conferences: A 12-Month Early-Career Researcher Experience*", **Postgraduate Medical Journal**, S.94(1108), ss.104-108.

MINITAB (2022), "*Technical Support*", **Minitab Corporate Web Page**, https://www.minitab.com/en-us/support/ (Access Date: 09.08.2022).

MONKEY LEARN (2022), "*How to Use AI in Excel for Automated Text Analysis: Text Analysis with Machine Learning*", **Monkey Learn Blog**, https://monkeylearn.com/blog/how-to-use-ai-in-excel-for-automated-text-analysis/#:~:text=A%20text%20analysis%20model%20can,and%20expressions%20and%20categorize%20them (Access Date: 05.08.2022).

MUSICK, Chad (2020), "*8 Questions and Answers about Predatory Journals: Protecting your Research, Reputation, and Funding from Theft and Fraud*", **Thinkscience.co.jp Corporate Web Page**, https://thinkscience.co.jp/en/articles/predatory-journals (Access Date: 07.08.2022).

PARKER, Michael (2014), "*How to Run a Pareto Chart in Minitab*", **Lean Sigma Corporation**, https://www.leansigmacorporation.com/how-to-run-pareto-chart-in-minitab/ (Access Date: 08.08.2022).

PAXTON, Ken (2022), "*How to Spot and Report Internet and E-mail Scams*", **Office of the Attorney General: texasattorneygeneral.gov Corporate Web Page**, https://www.texasattorneygeneral.gov/consumer-protection/how-spot-and-report-internet-and-e-mail-scams (Access Date: 30.07.2022).

SOLMS, Rossouw von (2016), "*Publish or Perish – But Where?*", **South African Computer Journal**, S.28(1), ss.44-58.

UNAB - UNESCO NETEXPLO ADVISORY BOARD (2022), "*Human Learning in the Digital Era*", **UNAB Corporate Web Page**, https://unesdoc.unesco.org/ark:/48223/pf0000367761?posInSet=1&queryId=N-EXPLORE-8fa8c367-f815-4a6a-a30b-a46391348f26 (Access Date: 10.08.2022).

USC LIBRARIES (2022), "*Research Guides: Evaluating Information Sources - Predatory Publishing*", **Libguides.usc.edu Corporate Web Page**, https://libguides.usc.edu/c.php?g=741391&p=5367265 (Access Date: 10.08.2022).

VILHELMSSON, Andreas (2022), "*Predatory Publishing 2.0: Why it is Still a Thing and What We Can do about it - ECR Community*", **ECR Community Corporate Web Page**, https://ecrcommunity.plos.org/2022/04/14/predatory-publishing-2-0-why-it-is-still-a-thing-and-what-we-can-do-about-it/ (Access Date: 11.08.2022).

WITMAN, Emma (2020), "*How to Retrieve Deleted E-Mails from Yahoo Mail, either by Undeleting Messages or Submitting a 'Restore Request'*", **Business Insider Corporate Web Page**, https://www.businessinsider.com/how-to-retrieve-deleted-emails-from-yahoo (Access Date: 12.08.2022).

YAP, Keith (2016), "*Text Analysis using Excel*", **Keithyap Corporate Web Page**, https://www.keithyap.com.au/text-analysis-using-excel/ (Access Date: 13.08.2022).