Dissemination of research in virtual reality-based rehabilitation: Journal publication profiles

Patrice L. (Tamar) Weiss
Dept. of Occupational Therapy
University of Haifa
Haifa, Israel
tamar@research.haifa.ac.il

Daphne R. Raban
Department of Business Administration
University of Haifa
Haifa, Israel
ORCID 0000-0003-1791-0310

Dorit Geifman
Department of Business Administration
University of Haifa
Haifa, Israel
dorit.geifman@gmail.com

Emily A. Keshner
Dept. of Physical Therapy
Temple University
Philadelphia, PA, USA
emily.keshner@temple.edu

***Abstract*— The application of virtual reality (VR) to rehabilitation is a relatively young, interdisciplinary field where clinical implementation very rapidly follows scientific discovery and technological advancement. Recently, computational tools in the area of data and information science have provided researchers with an ability to examine trends in their field. These tools have been used in all areas of science, from the humanities and social sciences to the health and biological sciences. The objective of the current paper is to study patterns of academic publications in the VR-based rehabilitation literature over the past 22 years. We conclude that exploring the distribution of topics and journals in the field of VR-based rehabilitation reveals a science that is evolving from one with a predominantly technology development focus, to one that focuses on how technology can support rehabilitation principles and outcomes. This shift in purpose has been demonstrated by the journals selected for submissions from the scientific community as well as the modification in mission and focus of particular journals.**

Keywords—Virtual reality, rehabilitation journal publications, topic modeling, Bradford distribution

I. INTRODUCTION

The application of virtual reality (VR) to rehabilitation is a relatively young, interdisciplinary field where clinical implementation very rapidly follows scientific discovery and technological advancement. In general, the field views itself as multidisciplinary, with many meetings comprised of those involved with technology development, software design, and clinical application. There has been considerable growth in the number and type of applications of VR to rehabilitation over the past 20 years as presented at conferences and depicted in reviews that periodically appear in the literature e.g., [1, 2, 3, 4]. These reviews are instructive when focusing on applications of VR technology to specific disability or impairment and often provide research insights and future directions. As the amount of evidence increases, a series of meta-analyses have been published e.g., [5, 6, 7]. Despite some noted improvements in the quality of research designs, the conclusions of such analyses remain equivocal with calls for additional studies with greater numbers of participants and more rigorous outcome measures.

In order to progress the science while maintaining the multidisciplinary nature of the field, it is necessary for results to be disseminated across disciplines. Only in this way, can we share current information and identify strengths and weaknesses in methodology. More recently, computational tools in the area of data and information science have provided researchers with an ability to examine trends in their fields. These tools have been employed in all areas of science, from the humanities and social sciences to the health and biological sciences [8, 9, 10, 11].

A recent example of interest to attendees of the International Conference on Virtual Rehabilitation is the analysis reported by Cipresso et al. [12] who applied tools of network analytics to study the broad field of virtual reality and augmented reality beginning from 1970. Although they did not include the term "rehabilitation" in their search, they found that the field was constantly expanding as well as moving toward increasing clinical usage.

We have also employed computational tools in the area of data and information science that provide the ability to examine trends in a field of study [13]. We focused on whether the field of VR in rehabilitation has succeeded in sharing information across its many disciplines over the course of the last 22 years. This window of exploration (1996-2018) was selected as the time that VR started to appear in the literature as an application to medicine or special education. Our analysis revealed that publication rates have continuously increased across three time periods: 1996 to 2005 (231 publications), 2006 to 2014 (807 publications), and 2015 to mid-2018 (776 publications) and that the principal topics have shifted from a focus on computer science and psychology to rehabilitation and public health.

We further found that a range of terms are used to represent the main corpus of VR-based rehabilitation, without an agreed upon set of terms specific to the field of VR-based rehabilitation. Rather, an assortment of central terms and concepts including "virtual reality", "virtual gaming", "virtual environments", "simulated environments" are used. Moreover, these terms are not used consistently across the many research directions. This means that there are many different and distinguishable areas of research and clinical foci (e.g., Tele-rehabilitation, Gait & Balance, Cognitive Rehabilitation, Gaming) that collectively define the field of VR-based rehabilitation.

The principal conclusion arising from that initial analysis was that VR-based rehabilitation consists of a network of scientific communities with a shared interest in overlapping

Partial support for this research was provided by the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation 1716/12.

methodologies rather than a directed and focused research field. By exploring the cross-citations for each paper included in our original analysis, we determined that the communities engaged in research or clinical applications of VR formed assemblages that could be distinguished by a focus on physical/motor or psychological/cognitive rehabilitation. These assemblages appear to thrive as a network of traditional disciplines utilizing a common technology rather than as a new multidisciplinary field.

A challenge to future progress in the field of VR-based rehabilitation is that a sharing of methods, results and conclusions across investigatory domains only occurs when there is an identified need or scientific rationale. Without some link between the disciplines, it is more difficult to educate ourselves about scientific rigor and technological validity across this diverse group.

One pathway for shared knowledge is the scientific journal. Publishing results in peer-reviewed journals is the main recognized medium for the long-term impact of scientific communications [14]. The major goal of this means of communication is dissemination of scientific data that has been screened for quality and relevance from the viewpoint a given journal's objectives and scope.

Selection of a relevant journal to publish research findings is not a trivial matter for the researcher. Its impact may be profound both in terms of the extent of dissemination of a researcher's work and the judgement placed on its value. In fact, the impact of specific research may be affected by the choice of journal in which it is published. The selection process, however, is very difficult in cases of interdisciplinary research using novel technologies in a field of study that is dynamically evolving. Researchers having an allegiance to more than one discipline may struggle with decisions regarding which data and analysis procedures are suited to which scientific publications. Will their target readers readily find these data, and will citation records attest to the research's impact?

The objective of the current paper is to study patterns of academic publications in the VR-based rehabilitation literature over the past 22 years. Using text analytic tools and the Bradford distribution (see below), we identified the key scientific journal titles in our field of study, divided them into zones of importance and examined changes in the numbers of publications in these journals and zones over three selected periods: 1996-2004, 2005-2013 and 2014-mid-2018. We conclude with a brief interpretation of how this analysis may guide current researchers in their journal selection considerations.

II. METHODS

The study involved four stages: data retrieval from the ISI Web of Science, data cleanup by implementing topic modeling, final topic modeling analysis and the analysis of the Bradford distribution.

A. Stage 1: Data retrieval

Data included items retrieved from the Web of Science (WoS) database spanning 1996-mid-2018. Search terms included "rehabilitation" both as a topic and as a WoS category. The search terms "simulation" and "virtual" were included. A post-hoc bi-gram analysis revealed that "virtual" appeared with other commonly used terms such as: world, walking, training, rehabilitation, reality, objects, mirror, hand, environment(s), and classroom. Papers were then limited to human-centered research by excluding the following areas considered to be irrelevant to the field of interest: water resources, environmental sciences, ecology, automation control systems, construction building technology, materials science, mathematical computational biology, agriculture, energy fuels, mechanics, operations research management science, mathematics, physical geography, physics, cell biology, business economics, fisheries, forestry, geography, thermodynamics, metallurgy metallurgical engineering, mining mineral processing, nuclear science technology, criminology penology, meteorology atmospheric sciences, archaeology, film radio television, geochemistry geophysics, oceanography.

The resultant bibliographic search generated a corpus of 3131 papers across the three periods. Experts' inspection of the papers retrieved indicated that the set included non-relevant papers and required further cleanup as described next.

B. Stage 2: Data clean-up

Topic modeling was used to identify and remove topics containing papers that were only marginally relevant to the core domain of VR-related rehabilitation. Topic modeling, Latent Dirichlet Allocation (LDA) in this case, is a text mining method used to identify latent topics and themes in a large corpus of text that consists of many documents [15]. It is used to navigate through large archives [16], enhance information retrieval methods [17] and identify latent topics within repositories of documents [18]. In order to avoid a bias towards research areas that dominated the later and more prolific years of the search, we divided the initial corpus into three periods: Period 1 (P1) extended from 1996 to 2005, Period 2 (P2) extended from 2006 to 2014 and Period 3 (P3), from 2015 to mid-2018 and applied the algorithm separately on each period. The cutoffs between P1 and P2 and between P2 and P3 were based on key milestones in technology development such as the advent of low cost, off-the-shelf devices such as the Nintendo Wii and Sony PlayStation 2 EyeToy whose major clinical impact occurred at the start of P2 and the availability of HMDs whose cost and technical parameters made highly suitable for clinical usage at the start of P3 [13, Table 1].

For each period we created a corpus of text that consisted of the publication title, abstract and all keywords available in WoS. The outcome of the topic modeling algorithm was two-fold: (1) a set of topics, each assigned with a distribution of raw terms (Figure 1a), and (2) for each publication, the distribution of the topics running through them (Figure 1b). A publication is regarded representative of a topic if the topic is ranked highest on its distribution list (Figure 1b). The review of the 20 highest ranking terms per topic as well as the 20 highest ranking publications (Figure 1a) were used by the domain experts to annotate each topic with its thematic name and to identify relevant and non-relevant topics and individual papers. Subject experts screened excluded papers to make sure we did not omit any relevant work, and 39 papers were manually returned.

The resultant 1814 relevant papers between the years 1996 to mid-2018, which were taken to the next stage of analysis, consisted of 231 publications in Period 1 (1996-2005), 807 publications in Period 2 (2006-2014) and 776 publications in Period 3 (2015 – mid-2018).

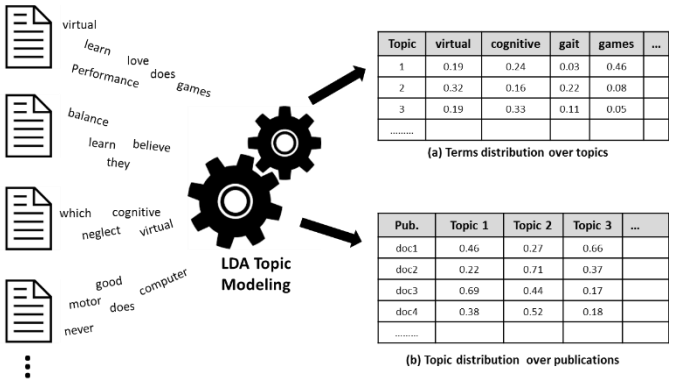

Fig 1: A schematic demonstration of the LDA topic modeling process. (a) illustrrates the distribution of terms over topics and (b) illustrrates the distribution of topics over publications.

C. Stage 3: Final topic selection

The purpose of Stage 2 was to refine and filter the data set, in Stage 3 the purpose was to study the emergent research field from the most relevant collection of publications. At this stage, the topic modeling algorithm was applied to the corpus as a whole without partitioning it into the three separate periods. This was done in order to generate a set of topics common across all periods. Following several iterations, the domain experts determined that 10 topics represent an optimal partitioning of the corpus into thematic structure. As a result, each of the publications in the entire corpus could be assigned to one of the topics, based on its highest ranked probabilities (cf. Figure 1b). Domain experts then assigned a thematic name to each topic based on the 20 terms and verified that each term was relevant to the 20 highest ranked publications for each of the 10 topics (Table 1).

TABLE I. TOPIC THEMES BASED ON TERMS, TITLES, KEYWORDS AND ABSTRACTS OF THE 20 MOST RELEVANT PAPERS. N IS THE NUMBER OF PUBLICATIONS PER TOPIC.

Topic	***# Publications***
Cognitive issues	230
Tele-Rehabilitation	226
Simulation	140
Psychological issues	105
Neural impact	226
Gait & Balance	190
Perception/ Navigation	162
Gaming	169
Neurological conditions	133
Interventions	233

D. Stage 4:Bradford's Distribution of Journal Productivity

In order to identify the distribution of journals containing scientific papers about VR-based rehabilitation, we applied the Bradford distribution which determines "how information on a subject is distributed among the resources where such information may be expected to be found." [20, pg. 138). This distribution assumes that information is neither randomly scattered, nor concentrated in a single location. Instead, information (in our case, scientific publications) scatters in a characteristic pattern, a pattern that should have obvious implications for how that information can most successfully and efficiently be sought. In other words, research has shown that the distribution of journals covering specific topics or broad disciplines is not random; it is a power law distribution. Some journals are central to a discipline and contain the majority of its publications, while other journals are relevant but not central, still other journals are only remotely related [21].

Bradford's distribution is used to partition the journals in a field to 'core', 'zone II', and 'periphery'. It is based on the observation that the core for any discipline carries a similar number of articles as the second and third zones of articles; however, these articles are concentrated in a small number of journals which constitute the core, usually up to ten. The second zone contains a comparable number of articles spread over more journals (about 50), and in the periphery, the number of journals that deal with the subject multiplies (about 250).

The Bradford distribution was implemented by: (1) summing the number of papers per journal title from the clean data obtained in Stage 2; (2) ranking the journal titles based on the number of papers per journal title; and (3) dividing the journals into three groups, each containing a similar number of papers (1814 divided into three, so about 600 papers per group) but a different number of journal titles. This process was performed four times: once for the entire dataset and once for each of the three time periods (P1, P2 and P3). Each group of journal titles obtained belongs to a different zone: the core of highly relevant journals, the second zone of journals with moderate relevance to the field, and the periphery with journals which publish in our field of interest only occasionally (in the periphery).

III. RESULTS

The Bradford distribution was applied to the 1814 publications obtained during Stage 2 across the entire time (1996 to mid-2018) as well as for each of the three periods (P1, P2 and P3) examined (Table II). Fourteen journals were in the core zone across the entire period; journals comprising the core zone varied from four during Period 1, ten during Period 2 and 15 during Period 3.

TABLE II. BRADFORD'S DISTRIBUTION OF JOURNALS OVER THE ENTIRE PERIOD AND FOR EACH OF P1, P2 AND P3.

	Total	***P1***	***P2***	***P3***
Core	14	4	10	15
Zone II	72	21	39	62
Periphery	409	61	209	226

Table III presents the titles of the journals in the core for each of the three periods in ranked order. Marked in grey are titles, which constitute the core of all publications in the field during the study period, 1996-2018 (14 journal titles, as indicated in Table II).

TABLE III. TITLES OF THE JOURNALS IN THE CORE FOR EACH OF THE THREE PERIODS IN RANKED ORDER.

	Core Journal Titles
Period 1	Cyberpsychology & Behavior
	Archives of Physical Medicine and Rehabilitation
	Disability and Rehabilitation
	Presence-Teleoperators and Virtual Environments
Period 2	Journal of Neuroengineering and Rehabilitation
	IEEE Transactions on Neural Systems and Rehabilitation Engineering
	Archives of Physical Medicine and Rehabilitation
	Neurorehabilitation
	Cyberpsychology & Behavior
	Neurorehabilitation and Neural Repair
	Disability and Rehabilitation
	Brain Injury
	Games for Health Journal
	Journal of Rehabilitation Research and Development
Period 3	Journal of Neuroengineering and Rehabilitation
	Journal of Physical Therapy Science
	Games for Health Journal
	Disability and Rehabilitation -Assistive Technology
	IEEE Transactions on Neural Systems and Rehabilitation Engineering
	Physical Therapy
	Archives of Physical Medicine and Rehabilitation
	Disability and Rehabilitation
	Clinical Rehabilitation
	PLOS ONE
	Neurorehabilitation
	Topics in Stroke Rehabilitation
	Neurorehabilitation and Neural Repair
	European Journal of Physical and Rehabilitation Medicine
	Methods of Information in Medicine

Of the initial four core journals from Period 1, two journals have persisted in the core over subsequent periods: Archives of Physical Medicine and Rehabilitation and Disability and Rehabilitation. Cyberpsychology & Behavior persisted into Period 2, but not into Period 3 when it was replaced by Cyberpsychology, Behavior and Social Networking and adopted a different scope and mandate. Presence-Teleoperators and Virtual Environments moved from the core to Zone 2 in Period 2, and was replaced by PLOS ONE in Period 3. Interestingly, although Period 3 is shorter than the other periods, no less than four journals have emerged within recent years as core to our field of study.

Period 2 core journals appear to have a strong focus on neurological aspects of VR-based rehabilitation. In Period 3, researchers seem to have chosen to publish in a wider range of clinically-oriented journals in addition to the previous choices. The ten core journals for P2 included three from P1 but not the fourth journal, Presence-Teleoperators and Virtual Environments. Two key engineering-oriented journals (Journal of NeuroEngineering and Rehabilitation and IEEE Transactions on Neural Systems and Rehabilitation) assumed a predominant role in reporting how technology is applied to rehabilitation. The Games for Health Journal made its first appearance in Period 2.

Table IV maps the results of the topic modeling onto the core journals. As shown in this table, during Period 1, the main contribution to the core of VR-based rehabilitation came from research about tele-rehabilitation, psychological issues, and simulation. During Period 2, we observe a somewhat more even distribution with a substantial contribution from neural impact articles, and a marked decline in the topics from Period 1. Lastly, the topic Intervention focused on research that tested or described VR-based treatment as opposed to studies that focused on assessment or technology development. Intervention is the leading topic for core publications in the most recent years in Period 3 indicating an emerging need to provide evidenced-based demonstrations of clinical effectiveness.

TABLE IV. TOPIC CONTRIBUTIONS TO CORE JOURNALS PER PERIOD

Topics	***Percent articles in the core***		
	P1	***P2***	***P3***
Cognitive issues	15	11	7
Gait & Balance	0	11	11
Gaming	5	8	13
Interventions	8	13	20
Neural Impact	11	16	10
Neurological Conditions	0	5	11
Perception/Navigation	5	10	9
Psychological Issues	19	5	4
Simulation	18	7	4
Tele-rehabilitation	20	13	10

IV. DISCUSSION & CONCLUSIONS

In this paper we used the Bradford Distribution to identify publication profiles for the scientific journals that publish articles in the field of virtual reality as applied to rehabilitation for the period ranging from 1995 to mid-2018. In order to deal with the substantial number of publications, we applied a computational technique called topic modeling which helped us clean the data and identify the main topical themes in our field.

This analytic approach is based on computational tools in the area of data and information science. These results provide researchers with an ability to examine trends in their field which can facilitate decision-making regarding the selection of the most suitable journals for publicizing keeping current in the field and disseminating their results.

Scientists are not usually concerned with the history of their science, which can lead to forgotten facts and redundant discovery. As described in their recent paper, Casadevall and Fang [20, page 4460] explain that "A better understanding of history can illuminate social influences on the scientific process, allow scientists to learn from previous errors, and provide a greater appreciation for the importance of serendipity in scientific discovery. Moreover, history can help to assign credit where it is due and call attention to evolving ethical standards in science. History can make science better."

With this in mind, it is instructive to note the journals that comprise the core for each period of research in VR-based rehabilitation. Rather, the trend is for those that used to be part of the core to shift to the middle or periphery zones as new journals emerge in the core. In some cases, this may be due to a change in the journal's mandate. For example, CyberPsychology & Behavior originally published articles on a wide range of VR applications to rehabilitation. Since 2010, this publication has changed both its name (CyberPsychology, Behavior & Social Networking) and its mandate to disseminate research on understanding the social, behavioral, and psychological impact of today's social networking practices, including Twitter, Facebook, and internet gaming and commerce. Another explanation is that a journal has divided its focus (e.g., Disability and Rehabilitation and Disability and Rehabilitation-Assistive Technology), with one of the journals offering a higher acceptance rate in a specific research area than the other.

In order to attract readership to their journals, several journals have facilitated advancement of VR applications to rehabilitation by hosting special issues, sometimes based on the proceedings of relevant conferences such as the International Conference on Virtual Rehabilitation and the International Conference on Disability, Virtual Reality and Associated Technologies. In other cases, special issues have been proposed by leading researchers in the field and accepted by the journal editors. Both Disability and Rehabilitation and Presence-Teleoperators and Virtual Environments hosted special issues related to VR-based rehabilitation during P1, serving as key supporters of the incipient field. This approach may appear self-serving, but it does present opportunities for dissemination among a broad audience which would strengthen the interdisciplinary mission of the field.

We conclude that exploring the history of publications in the field of VR-based rehabilitation reveals a science that is evolving from one with a predominantly technology development focus, to one that focuses on how technology can support rehabilitation principles and outcomes. This shift in purpose has been demonstrated by the journals selected for submissions from the scientific community as well as the modification in mission and focus of particular journals. This partnership between stakeholders suggests a healthy, dynamic evolution that will keep the science and application of VR relevant and valid. A challenge to future progress in the field of VR-based rehabilitation is that a sharing of methods, results and conclusions across investigatory domains only occurs when there is an identified need or scientific rationale.

Our reliance on only one database could be viewed as a limitation of this study, but WoS is generally accepted in the scientific community as a source of top-quality scientific publications that provides a good estimation of the significant research in the field. Future studies may consider employing additional databases to examine how these interdisciplinary communities individually define themselves with the goals of gathering information and working collectively toward disseminating information essential to associated communities.

REFERENCES

[1] A. A. Rizzo and G. J. Kim, "A SWOT analysis of the field of virtual reality rehabilitation and therapy," Presence: Teleoperators and VEs, vol. 14, no. 2, pp. 119–146, Apr. 2005.

[2] M. F. Levin, P. L. Weiss, and E. A. Keshner, "Emergence of virtual reality as a tool for upper limb rehabilitation: Incorporation of motor control and motor learning principles," Phys Ther, vol. 95, no. 3, pp. 415–425, Mar. 2014.

[3] Z. Li, X.-G. Han, J. Sheng, and S.-J. Ma, "Virtual reality for improving balance in patients after stroke: A systematic review and meta-analysis," Clin Rehabil, vol. 30, no. 5, pp. 432–440, May 2016.

[4] D. Ravi, N. Kumar, and P. Singhi, "Effectiveness of virtual reality rehabilitation for children and adolescents with cerebral palsy: an updated evidence-based systematic review," Physiotherapy, vol. 103, no. 3, pp. 245–258, Sep. 2017.

[5] K. E. Laver, S. George, S. Thomas, J. E. Deutsch, and M. Crotty, "Virtual reality for stroke rehabilitation: an abridged version of a Cochrane Review," Eur J Phys Rehabil Med, pp. 497–506, Aug. 2015.

[6] D. Corbetta, F. Imeri, and R. Gatti, "Rehabilitation that incorporates virtual reality is more effective than standard rehabilitation for improving walking speed, balance and mobility after stroke: A systematic review," J Physiother, vol. 61, no. 3, pp. 117–124, Jul. 2015.

[7] M. C. Howard, "A meta-analysis and systematic literature review of virtual reality rehabilitation programs," Comput Human Behav, vol. 70, pp. 317–327, May 2017.

[8] E. M. Talley, D. Newman, D. Mimno, B. W. Herr, H. M. Wallach, G. A. P. C. Burns, A. G. M. Leenders, and A. Mccallum, "Database of NIH grants using machine-learned categories and graphical clustering," Nat Methods, vol. 8, no. 6, pp. 443–444, Jun. 2011.

[9] D. Yu, Z. Xu, W. Pedrycz, and W. Wang, "Information sciences 1968–2016: A retrospective analysis with text mining and bibliometric," Inform Sci, vol. 418-419, pp. 619–634, Aug. 2017.

[10] B. Chen, S. Tsutsui, Y. Ding, and F. Ma, "Understanding the topic evolution in a scientific domain: An exploratory study for the field of information retrieval," J Informetr, vol. 11, no. 4, pp. 1175–1189, Nov. 2017.

[11] A. Goldstone and T. Underwood, "What can topic models of PMLA teach us about the history of literary scholarship?," J Digit Humanit, vol. 2, no.1, pp. 39-48.

[12] P. Cipresso, I. A. C. Giglioli, M. A. Raya, and G. Riva, "The past, present, and future of virtual and augmented reality research: A network and cluster analysis of the literature," Front Psychol, vol. 9, pp. 1–20, Nov. 2018.

[13] E. A. Keshner, P. L. Weiss, D. Geifman, and D. Raban, "Tracking the evolution of virtual reality applications to rehabilitation as a field of study," J. NeuroEng. Rehabil. Unpublished.

[14] D. R. Raban and D. Geifman, “Forming socio-intellectual capital: The case of LINKS. ,” in Learning in a NetworKed Society (LINKS), Y. Kali and A. Baram-Tsabari, Eds. Springer, 2018.

[15] D. M. Blei, “Probabilistic topic models,” Communications of the ACM, vol. 55, no. 4, pp. 77–84, Apr. 2012.

[16] D. J. Newman and S. Block, “Probabilistic topic decomposition of an eighteenth-century American newspaper,” J Am Soc Inf Sci Technol, vol. 57, no. 6, pp. 753–767, Feb. 2006.

[17] F. Jian, “A simple enhancement for ad-hoc information retrieval via topic modelling,” in Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 733–736, July 2016.

[18] M. J. Bates, “Speculations on browsing, directed searching, and linking in relation to the Bradford distribution,” in Emerging frameworks and methods: Proceedings of the Fourth International Conference on Conceptions of Library and Information Science (CoLIS 4), 2002, pp. 137–150.

[19] V. P. Diodato, Dictionary of Bibliometrics. New York, NY: Routledge, 1994.

[20] A. Casadevall and F. C. Fang, “(A) Historical science,” Infect Immun, vol. 83. pp. 4460–4464, 14-Sep-2015.

.