

Traditional Chinese Medicine with Data Visualization: Prospects and Difficulties in the Age of Big Data

Poorna Chander Rao

Palmer iSchool of Library, Long Island University, USA

Received on: 21 Oct 2024 Revised on: 20 Nov 2024 Accepted Date: 25 Dec 2024

Published on: 24 Jan 2025

Abstract: With the advent of personal computers, the Internet, portable devices, and big data analytical settings, computer-based medical diagnostic technologies have grown substantially since the 1950s. To address difficult issues related to health and illness, these technologies use the rudiments of information retrieval and representation (IRR). But, these systems have, from the beginning, paid little attention to TCM methods, often because these approaches have failed in randomized controlled studies. There is still a lot of mystery around traditional Chinese medicine (TCM), despite the fact that it is an integral aspect of healthcare systems across the globe, especially in a number of Asian nations. In view of current IRR techniques, it would be beneficial to compare traditional Chinese medicine (TCM) diagnostic and treatment methods with Western medical models in order to find out whether a new kind of translational medicine can be created that improves medical outcomes while lowering health care costs globally. This would be necessary because disease is still prevalent in society. Using bibliometric tools, multiple correspondence analysis, and data visualizations, this study examines author productivity, collaborations, and research trends in TCM and IRR published in SCOPUS from 1985 to 2020. As we embark on a new age of data-intensive scientific discovery in medicine, opportunities and difficulties have been identified that will help us determine the field's future courses.

Keywords: Bibliometrics, data visualization, computer-based medical diagnostic systems, traditional Chinese medicine, data retrieval, herbal pharmaceutical technology, multiple correspondence analysis

Introduction

Numerous resources for biomedical and health science have been contributed by scholars of traditional Chinese medicine (TCM) since its inception thousands of years ago [1]. These resources include published literature, medicinal materials, herbs, diagnostic matrices, clinical records, medical formularies, and more. True information science pioneers in many respects, TCM's founders foresaw societal issues impacting their day and set out to solve them.

and aware of what was available at the time, which consisted mostly of classical medical texts transmitted orally from ancient wise men and women before the invention of written records, as well as botanicals and other naturally occurring compounds collected from the surrounding area [2]. With the goal of containing epidemic diseases, these pioneers in medicine went on to compile and index this information into seminal works (such as crude expert knowledge systems) that would serve as a retrieval mechanism for a variety of medical problems [3].

These old systems mirror mathematical reasoning based on the existence or absence of human biological information [4] and use a binary system of numbers and probabilities that closely resembles the idea of bits and bytes that make up the foundation of our modern computing environment [5]. Their primary sources of information are patterns of natural discord, representations of herbal medicine, and models of treatment. A concept with origins in the field of artificial intelligence (AI) that aims to mimic the deductive process of disease diagnosis is set to become a common reality in the near future, thanks to developments in big data analytics, the Internet of Things (IoT), and society's persistent quest for optimal health and wellness [6]. According to Samsung [7], in the not-too-distant future, we may have ubiquitous sensors that continuously monitor our health. These sensors will be connected to a massive AI network, which will allow them to detect early warning signs of illness, encourage users to make healthier choices, direct medical research, rank us, and possibly even implement a pay-as-you-live model for health and life insurance premiums. Complex traditional Chinese medicine (TCM) algorithms, enabled by information representation and retrieval (IRR) technology, can now aggregate medical data from various sources according to suitable criteria, filling a gap in Western diagnostic practices. These algorithms mimic classical deductive and reasoning procedures to solve medical problems and recommend treatment protocols [8]. Several IRR challenges arise from the complexity of medical knowledge [9], such as the need for standardized data formats, well-formulated inputs, and efficient critical factors for appropriate feature representation. Because traditional Chinese medicine (TCM) is based on ancient practices rather than modern science, there is no universal system that can decipher the myriad medical manuscripts written in different dialects and kept in various places around the world. This further complicates matters [10]. While regional cultural differences and language variations do enrich and diversify the TCM terminology system, they also pose problems with standardization when it comes to modernizing this classical medical modality, particularly when contrasted with Western medical systems, which also face vocabulary issues [11]. This is because medicine has evolved and developed over thousands of years. The Western scientific community also views natural herbal medical practices and other uses like acupuncture as experimental, in part because of the subjective character of these treatments. Relying more on big data, information-rich experiments centered on IRR and mobile technologies as opposed to controlled clinical trials is the way to

go in the future of research.

a method for quantitatively gathering vast amounts of TCM data. A bibliometric analysis with data visualizations related to IRR and TCM, specifically through author, paper, and co-word analysis, will be conducted in this work to better understand the field's conceptual structure and to identify potential research opportunities and challenges in developing and implementing a TCM-based diagnostic and medical recommendation system.

Materials and Methods

Scopus® (<http://www.scopus.com>), considered by some to be the largest abstract and citation database of peer-reviewed literature, including scientific journals, books and conference proceedings, was initially searched on October 9, 2019 and again on November 16, 2019 for all citations with the Boolean string [("Chinese medicine") AND ("information retrieval") OR ("information representation")] located in the article title, abstract, or keywords. The results of the search revealed 192 documents for the period 1985-2020; a BibTeX export file was saved and read into R, a free software environment for statistical computing and graphics (<http://www.r-project.org>), using bibliometrix [12], a tool for comprehensive science mapping analysis. The function `*readFiles*` was initially used to create a single large character vector; this object was then converted into a data frame using the function `*convert2df*`, with cases corresponding to manuscripts and variables to field tags in the original export file, comprising all bibliographic attributes of each document based on Clarivate Analytics WoS Field Tag codified industry standards [13]. During data cleansing, four entries were removed from the data frame due to a lack of author and other document information: three represented conference proceeding introductions and a fourth represented an introductory chapter on semantic grid applications for traditional Chinese medicine.

Results and Discussion

Descriptive analysis

To begin, a descriptive analysis was performed on the bibliographic data frame using the function `*biblioAnalysis*`; a display of the main results are included in Table 1. A total of 188 documents from 105 sources and 899 author appearances were noted, including a collaboration index of 2.49. Figure 1 illustrates the number of publications per year for the collection period 1985-2020; one large spike occurs in 2006 (28 citations) which continues into 2008, followed by a dip and then another, slower increase cumulating in 2017. This trend mirrors publications on IRR in general, which also peaks in 2006, according to Scopus, representing the maturation of computer browsing and the initial transition to mobile smart devices.

Table 2 contains the top 10 most cited papers in the collection, with Kanehisa M, et al. [14] having been cited over 1,500 times for

their work in Japan on computerizing disease information using pathway maps, all Japanese drugs (including every TCM herbal formula), and gene/molecule lists. The second most-cited paper, Tang JL, et al. [15], represents one of the oldest papers in the current collection and is a summary of issues relating to randomized controlled trials in TCM, specifically: lack of blinding; low sample sizes; using another, unproven TCM treatment as the control; not long-term in nature; incompleteness; lack of quantitative data; missing intention to treat; lack of data on baseline characteristics or side effects; short reporting; and presence of publication bias. The third most-cited paper [16], was published in an American Heart Association journal and concludes, in similar fashion, the insufficiency of TCM evidence in using herbal medicinal for stroke patients, due to bias from poor methodology, even though the agents used appeared to be potentially beneficial and nontoxic in nature. The fourth document with the most citations [17] discusses newly published guidelines and technical notes by the European Union, in collaboration with Chinese scientists, to encourage good practice in the collection, assessment, and publication of TCM literature. The fifth most-cited document [18] reviews advances in automated tongue diagnosis, a key requirement for the accurate gathering of quantitative data, while the sixth most-cited document [19] discusses the development of ontology for TCM IRR. Fang YC,

et al. [20] and Qiao X, et al. [21] both discuss the creation of TCM databases, while Wojcikowski K, et al. [22] again point to difficulties with randomized controlled trials in TCM, particularly relating to the use of herbal medicinals in the treatment of kidney disorders. The tenth most-cited document [23] concludes that text mining of TCM literature and clinical data carries with it the potential to clarify misunderstandings, but clear operational definitions are first required.

Table 3 lists total citations by country, along with average article citations; Japan leads this metric due to the Kanehisa M, et al. [24] document noted above, with China positioned strongly behind with 865 total citations. As expected, over 50% of the top 10 countries are located in Asia; the United States remains far behind in this research area with only 10 total citations related to one published article. Table 4 illustrates the top author countries in the collection, with China strongly in the lead with 109 articles (a frequency of 0.76224) – additionally, 89% of these articles (97) are considered single country publications. Given the nature of this data, TCM and IRR research in the East has been mainly conducted as single country publications (China, Hong Kong, Korea, and Japan) while Australia, Canada, and Germany research has been more multi-country in nature.

Table 1: Main information regarding the collection.

Description	
<i>Documents</i>	188
<i>Period</i>	1985 – 2020
<i>Sources</i>	105
<i>Average citations per documents</i>	17.9
<i>Authors</i>	710
<i>Author Appearances</i>	899
<i>Authors of single-authored documents</i>	13
<i>Authors of multi-authored documents</i>	697
<i>Documents per Author</i>	0.265
<i>Authors per Document</i>	3.78
<i>Co-Authors per Documents</i>	4.78
<i>Collaboration Index</i>	2.49

Table 2: Top 10 most cited papers.

Paper	Total Citations (TC)	TC per Year
<i>Nucleic Acids Res</i> [14]	1,506	136.91
<i>Br Med J</i> [15]	203	9.67
<i>Stroke</i> [16]	130	10
<i>J Ethnopharmacol</i> [17]	101	12.62
<i>IEEE Trans Med Imaging</i> [18]	88	5.87
<i>Artif Intell Med</i> [19]	80	5
<i>BMC Complement Altern Med</i> [20]	79	6.58
<i>J Chem Inf Comput Sci</i> [21]	76	4.22
<i>J Lab Clin Med</i> [22]	69	4.93
<i>J Biomed Informatics</i> [23]	59	5.9

Table 3: Top 10 total citations per country.

Country	TC	Average Article Citations
<i>Japan</i>	1,509	754.5
<i>China</i>	865	7.94
<i>Hong Kong</i>	286	47.67
<i>Australia</i>	254	31.75
<i>Taiwan</i>	79	79
<i>Korea</i>	51	10.2
<i>United Kingdom</i>	32	16
<i>Singapore</i>	21	21
<i>Germany</i>	13	6.5
<i>USA</i>	10	10

Table 5 lists the top 10 most productive authors in the collection, based on both number of published articles (full counting) and number of published articles fractionalized (which assigns co-authored publications a fraction of one to each of the co-authors); studies have illustrated that, oftentimes, fractional counting offers a more useful perspective than full counting, especially as a means to avoid misunderstanding or misinterpretation [25]. Fractionalized counting does not affect the most productive author (Zhang Y) but does shift the order of the others slightly and results in the appearance of one new author (Xiong X) in the top 10.

Figure 2 applies the **authorProdOverTime** function on the collection to calculate and visualize the production of these top 10 authors over time, in terms of number of publications and total citations per year, for the period 1985-2020. This illustration clearly depicts the top producing author (Zhang Y) as covering both a wide period (2005-2019) along with more recent proliferation, oftentimes as a co-author, as noted by the number of articles fractionalized (2.09). Other authors with more recent production

include Yu T (8 overall publications), Li J and Wang Y (6 overall publications each), and Liu L (5 overall publications).

Table 6 contains the top 10 most frequent journals, based on number of published articles in the collection – led by the Chinese Journal of Clinical Rehabilitation with 25 articles and followed by Evidence-Based Complementary and Alternative Medicine with 11 publications. However, it is important to also look at this data from the perspective of number of documents published annually; this information for each of the top five sources is visualized in Fig. 3 using the function **sourceGrowth**, which illustrates that the Chinese Journal of Clinical Rehabilitation was only in existence from 2002-2006. Since then, four newer journals have increased their publication rate, particularly Evidence-Based Complementary and Alternative Medicine, which is second in number of articles but clearly the leading publication in this field, particularly as the journal currently holds an h-index of 72 and sits as the sixth ranked journal in complementary and alternative medicine [26].

Table 4: Top 10 corresponding author's countries.

Country	Articles	Frequency	SCP	MCP	MCP Ratio
China	109	0.76224	97	12	0.11
Australia	8	0.05594	4	4	0.5
Hong Kong	6	0.04196	5	1	0.167
Korea	5	0.03497	5	0	0
Canada	2	0.01399	0	2	1
Germany	2	0.01399	1	1	0.5
Japan	2	0.01399	2	0	0
United Kingdom	2	0.01399	2	0	0
Brazil	1	0.00699	1	0	0
Hungary	1	0.00699	1	0	0

Table 5: Top 10 Most productive authors.

Author	No. of Articles	Author	No. of Articles Fractionalized
<i>Zhang Y</i>	10	<i>Zhang Y</i>	2.09
<i>Wu Z</i>	9	<i>Zhou X</i>	2.05

Yu T	8	Wu Z	1.75
Zhou X	8	Wang Y	1.45
Chen H	7	Yu T	1.35
Chen X	7	Li J	1.31
Li J	6	Chen H	1.25
Wang Y	6	Xiong X	1.25
Cui M	5	Cui M	1.19
Liu L	5	Chen X	1.11

Table 6: Top 10 Most frequent journals.

Sources	No. of Articles
<i>Chinese Journal of Clinical Rehabilitation</i>	25
<i>Evidence-Based Complementary and Alternative Medicine</i>	11
<i>Journal of Ethnopharmacology</i>	7
<i>Zhongguo Zhongyao Zazhi</i>	7
<i>Journal of Alternative and Complementary Medicine</i>	6
<i>Chinese Journal of Evidence-Based Medicine</i>	5
<i>Chinese Journal of Integrative Medicine</i>	5
<i>Zhongguo Zhongxiyi Jiehe Zazhi</i>	5
<i>Journal of Chinese Integrative Medicine</i>	4
<i>Journal of Traditional Chinese Medicine</i>	4

Table 7: Top 10 Most frequent keywords.

Author Keywords	No. of Articles	Keywords-Plus	No. of Articles
<i>Traditional Chinese Medicine</i>	27	Information Retrieval	173
<i>Systematic Review</i>	13	Chinese Medicine	151
<i>Chinese Medicine</i>	7	Human	120
<i>Information Retrieval</i>	7	Article	86
<i>Meta Analysis</i>	6	Medicine	77
<i>Information Extraction</i>	5	Humans	67
<i>Ontology</i>	5	Review	66
<i>Review</i>	5	Herbaceous Agent	54
<i>TCM</i>	5	Chinese Traditional	45
<i>Chinese Herbal Medicine</i>	4	Priority Journal	42

Table 8: Historiograph legend.

Year	Reference	Local Citations	Global Citations
1999	TANG JL, 1999, BR MED J	3	203
2001	CHANG IM, 2001, ANN NEW YORK ACAD SCI	1	20
2002	BENSOUSSAN A, 2002, J TOXICOL CLIN TOXICOL	2	36
2002	QIAO X, 2002, J CHEM INF COMPUT SCI	2	76
2004	KA WF, 2004, J ALTERN COMPLEMENT MED	1	7
2004	ZHOU X, 2004, ARTIF INTELL MED	5	80
2005	WANG JF, 2005, CLIN PHARMACOL THER	-	21
2006	LI Y, 2006, ZHONGGUO ZHONG XI YI JIE HE ZA ZHI	1	5
2007	FLOWER A, 2007, J ALTERN COMPLEMENT MED	-	31
2007	CHEN H, 2007, BMC BIOINFORM	2	25
2008	FANG YC, 2008, BMC COMPLEMENT ALTERN MED	5	79

2008	TSE HYG, 2008, J BIOMOL SCREEN	-	2
2008	CHEN H, 2008, BMC BIOINFORM	-	5
2009	MAY BH, 2009, BIOGERONTOLOGY	1	18
2010	ZHANG X, 2010, PROC - INT CONF BIOMED ENG INF, BMEI	1	11
2010	ZHOU X, 2010, J BIOMED INFORMATICS	3	59
2010	BOEHM K, 2010, HEALTH INF LIBR J	-	13
2011	SAMPSON M, 2011, EVID -BASED COMPLEMENT ALTERN MED	-	7
2012	JIANG Z, 2012, IEEE INT CONF E-HEALTH NETWORKING	1	11
2012	MAY BH, 2012, J ALTERN COMPLEMENT MED	1	16
2012	JIANG M, 2012, EVID -BASED COMPLEMENT ALTERN MED	-	43
2014	CHEN X, 2014, COMP MATH METHODS MED	-	3
2014	CHEN H, 2014, BIOMED RES INT	-	3
2015	XIONG X, 2015, NAT REV CARDIOL	-	21
2015	YOU M, 2015, SCI WORLD J	-	1
2015	XIONG X, 2015, BMJ OPEN	1	7
2016	MAY BH, 2016, J ALTERN COMPLEMENT MED	-	9
2016	WAN H, 2016, J AM MED INFORMATICS ASSOC	-	4
2016	YU T, 2016, PROC - INT CONF BIOMED ENG INFORMATICS, BMEI	-	1
2017	LIU YQ, 2017, CHIN J INTEGR MED	1	1
2017	WANG L, 2017, EVID -BASED COMPLEMENT ALTERN MED	-	1
2019	YOON SH, 2019, CHIN J INTEGR MED	-	3

Table 7 contains the top 10 most frequent keywords using two keyword variations: authors' keywords, as specifically selected by each author, and keywords-plus, which are those keywords extracted from the publication by Scopus' database algorithms. The results vary, with keywords-plus identifying many more in common across the collection – for example, information retrieval was only selected by seven authors as a keyword but appears 173 times as a keywords-plus. Overall, while keywords-plus is as effective as authors' keywords in terms of bibliometric analysis investigating the knowledge structure of a particular field, it is often less comprehensive in representing an article's specific content [27]. However, within this collection, the use of keywords-plus may lead to a greater understanding than simply using those keywords identified by the authors, due to the increased volume and commonality of terms; this is particularly evident in Figure 4a and Figure 4b, which clearly illustrate greater and more prolific growth of keyword-plus over time, as compared to authors' keywords.

Network visualizations

Figure 5 utilizes the **threeFieldsPlot** function to create a Sankey diagram. This diagram displays multiple attributes simultaneously, including top authors on the left, top author keywords in the center, and major cited references on the right. It summarizes the activity of the journals and keywords that are presented below. The size of the boxes represents total output, whereas the breadth of the bands is directly related to frequency. The picture in Figure 6 shows the web of partnerships between researchers in and

from one country to another, using the help of the **biblioNetwork** and the **networkPlot** functions. The size of the sphere represents total production, the density of lines represents the strength of the cooperation, and the color represents the nature of the collaboration in a spherical arrangement. The bulk of studies are concentrated on China and its surrounding areas, especially with Western nations; nevertheless, six other nations—Taiwan, Brazil, Japan, Spain, Hungary, and Korea—are shown to be operating autonomously. Germany–United Kingdom–China, USA–Canada–China, and Singapore–China, according to line density, are the primary cooperation networks, but they are feeble. In Figure 7, we can see an example of a word co-occurrence network that groups and maps phrases taken from abstracts of authors. This information was first retrieved from each manuscript's textual abstract field using the **termExtraction** function in conjunction with word stemming. The bigger clusters of information science (red), herbology (green), and biomedicine (blue) create broader categories in this picture, with TCM standing alone at the bottom left. Figure 8 shows the same thing for author keyword co-occurrences; it follows the same pattern as the extracted words from the author abstracts, but because each author knew the work inside and out when they chose their keywords, this visualization is more organized and shows a logical progression from traditional Chinese medicine (TCM), through an information retrieval main section, herbology, and finally Western biomedicine.

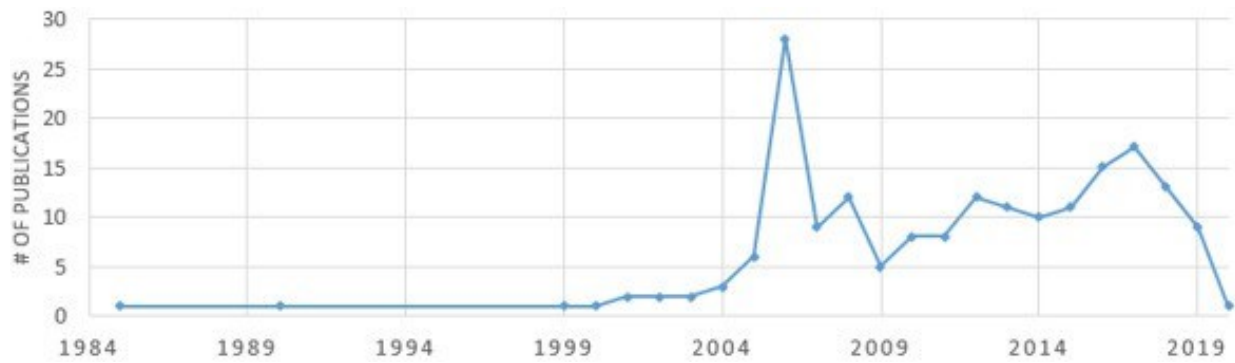


Figure 1: Publications per year 1985-2020 (Scopus).

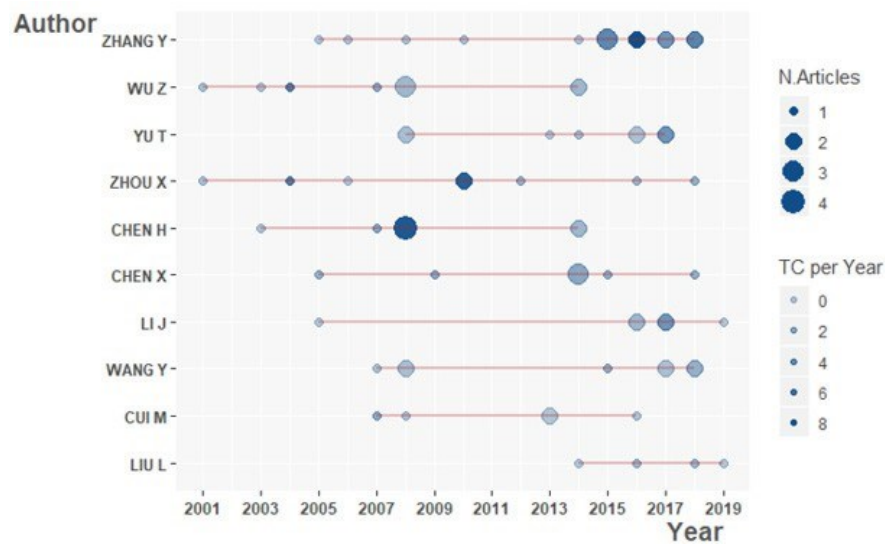


Figure 2: Top 10 author productivity for the period 1985-2020.

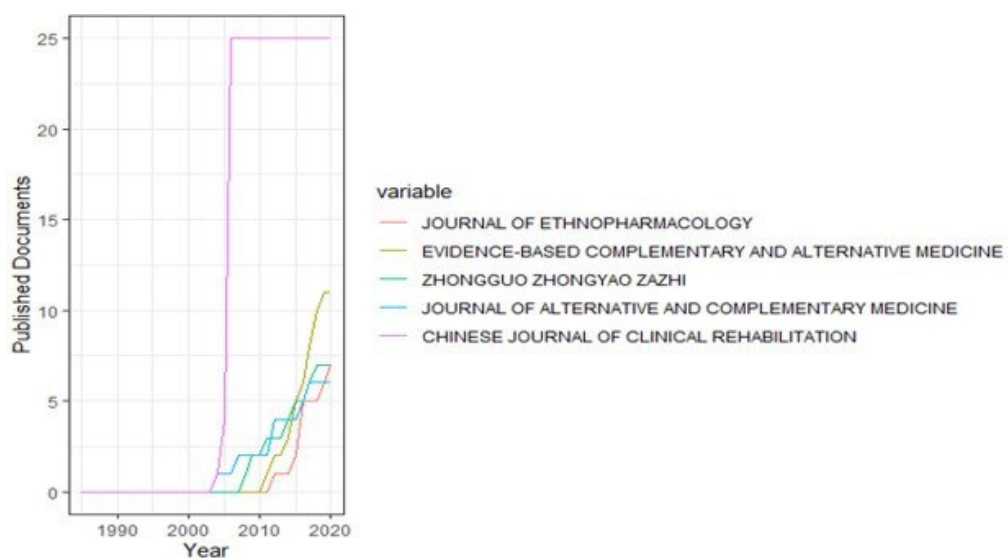


Figure 3: Source Growth, Top Five Journals, 1985-2020.

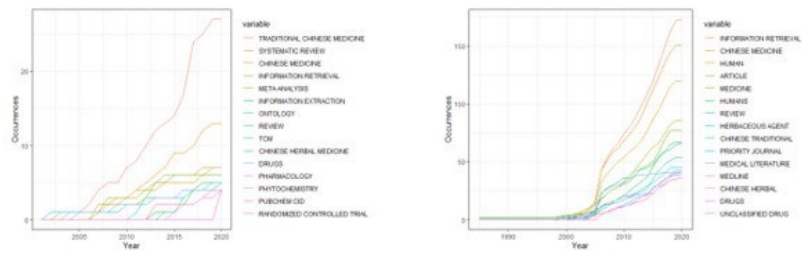


Figure 4a: Author keyword growth per year, 1985-2020. Figure 4b: Keyword-plus growth per year, 1985-2020.

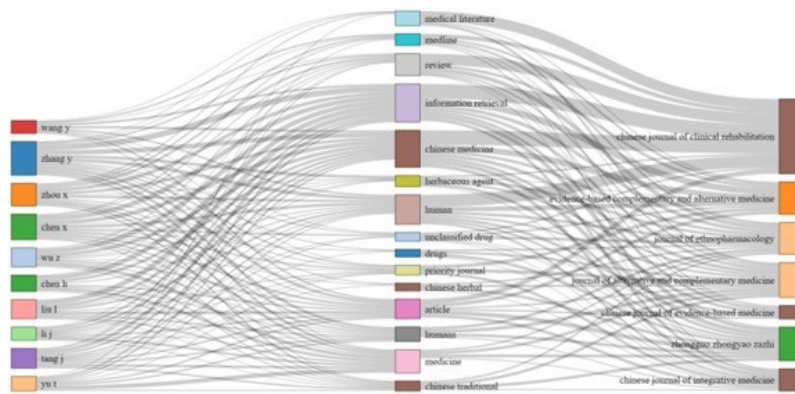


Figure 5: Sankey Diagram of Main Authors, Keywords, and Journals.

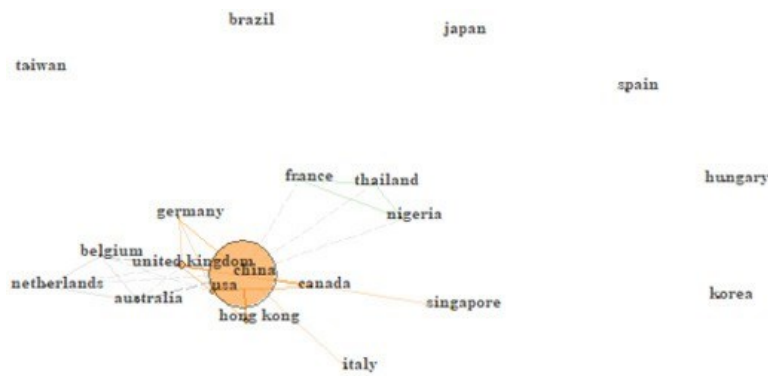


Figure 6: Country collaboration network.

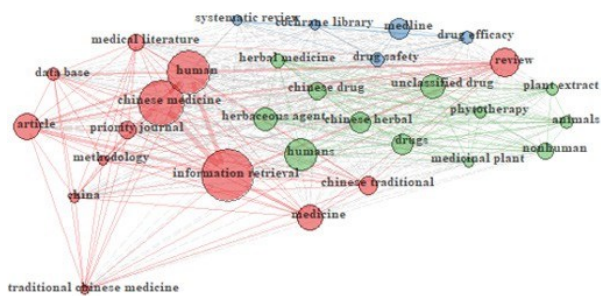


Figure 7: Author abstract co-occurrence network.

In the historiograph for TCM and IRR, the paper by Tang JL, et al. [30] serves as the first work noted and surveyed the efficacy of randomized controlled trials in TCM literature; this spurred a line of research (highlighted in blue) extending to Flower A, et al. [31] who advocated for the Delphi method, Sampson M, et al. [32] who searched for additional databases to identify more successful controlled trials, and Jiang M, et al. [33] who evaluated evidence-based literature for TCM diagnosis and knowledge discovery.

The second and most dominant area of research in Figure 9 (highlighted in red) focuses on TCM information databases and begins with Chang IM [34] who investigated anti-aging and health-promoting elements derived from traditional herbal remedies found in the Traditional Oriental Medicine Database, leading to future research by Boehm K, et al. [35], who provided an overview of 45 published database resources for complementary and alternative medicine. Bensoussan A, et al. [36] established research goals for the search and retrieval of scientific evidence regarding the toxicity of Chinese herbal medicine, which contributed to Wang JF, et al. [37] and the construction of a TCM information database. Qiao X, et al. [38] created a structured database of components extracted from TCM herbs, while Zhou X, et al. [39] wrote an influential paper (5 local citations) that used ontology to construct a unified TCM language system for information retrieval and integration. This led to Zhou X, et al. [40], which investigated research issues regarding TCM text mining, You M, et al. [41], who developed an intelligent system for customized clinical TCM case management and analysis, Wan H, et al. [42], who constructed a heterogeneous factor graph model for extracting relations from TCM literature, and Yu T, et al. [43], who utilized semantic web technologies to build cross-cultural communication between TCM and Western medicine.

Chen H, et al. [44] used semantic and knowledge-based techniques to build e-toolkits that facilitate TCM information sharing; this contributed to Chen H, et al. [45], which introduced state-of-the-art semantic web technologies for biomedicine as a whole, including applications for TCM and translational research. Tse HYG, et al. [46] developed an online TCM herbal medical database built again from the earlier herbal works of Bensoussan A, et al. [47] and Qiao X, et al. [48]. Li Y, et al. [49] focused on utilizing data mining techniques that compared clinical characteristics of TCM and Western medicine in the diagnosis of rheumatoid arthritis, which led to Fang YC, et al. [50], another influential paper, who developed a database to provide information about TCM, genes, diseases, effects, and ingredients from a wide variety of biomedical literature. This was further studied by Zhang X, et al. [51], who created a hierarchical symptom-herb topic model for TCM research in the treatment of diabetes, and Jiang Z, et al. [52], who used link topic models to analyze TCM symptom-herb regularities. Chen X, et al. [53], also used this research to develop a semantic search engine for IRR in modern biology and TCM, along with Chen H, et al. [54], who presented a general web ontology language reasoning

framework to study biological entities across TCM and Western medicine. These works ultimately influenced Wang L, et al. [55], who used topic model and multi-label classifiers to predict the function of TCM herbal prescriptions.

A third, smaller research group in Fig. 9 (highlighted in green) begins with Ka WF [56], who introduced journals and other TCM research materials available online and May BH, et al. [57], who searched English and Chinese databases to review the effectiveness and safety of Chinese herbal medicines for use in the treatment of cognitive and memory impairment. These two works led to additional research by May BH, et al. [58], comparing and evaluating published TCM collections for research and drug discovery searches, and May BH, et al. [59], who searched a database of over 1,000 classical and pre-modern TCM texts for the treatment of memory impairment. A fourth research group was also mapped on the historiograph (highlighted in purple) relating to difficulties in drawing clinical conclusions in the treatment of specific Western medicine disorders with TCM: Xiong X [60] reviewed an article on randomized controlled trials for the treatment of cardiovascular disease with TCM, while Xiong X, et al. [61] researched the clinical effects of a TCM herbal decoction in the treatment of hypertension. A fifth and final research group (highlighted in yellow), albeit small, begins late and relates to the standardization of TCM: Liu YQ, et al. [62] focused on standards and proposals established by the International Organization for Standardization (ISO), which Yoon SH, et al. [63] built from this to investigate the pros and cons of proposing standard terminology for acupotomy, a treatment modality which involves the use of both an acupuncture needle and a surgical scalpel. These five pathways of research in TCM and IRR help illustrate both the research difficulties in the field as well as opportunities in the treatment of specific diseases and the construction of modern databases and ontologies for the future use of this medical modality.

Visualizing the conceptual structure of the field

An exploratory multivariate approach, multiple correspondence analysis enables numerical and graphical examination of patterns in connections of categorical dependent variables, including keywords [64]. To create a field-specific conceptual map, we utilized keywords-plus with no stemming and up to five clusters, all generated using the `*conceptualStructure*` function. The distribution of the dots and their relative positions across the dimensions are used to interpret the results; words with similar distributions are shown closer together in Figure 10. The following is the translation of map data that is proposed by Cuccurullo C. et al. [65]: The size of each point is directly related to the keyword's overall contribution; the closeness of adjacent points indicates the presence or absence of shared substance; and the map's dimensions mirror the topical orientation's characteristic poles within TCM and IRR; the center of the map represents the average position of all the articles, thus the research

field's center. In the upper right corner, you can see Cluster 1 (highlighted in red). This cluster comprises terms from publications discussing traditional Chinese medicine (TCM) herbs and natural substances as potential remedies for both short-term and long-term health problems. In the middle and lower center of the map, you can see the biggest cluster (highlighted in blue). It comprises terms from publications about the creation of databases and retrieval systems for scientific TCM medical literature, including Zhou X, et al. [66]. In the third cluster, which is centered to the left and is highlighted in green, you may find terms related to traditional Chinese medicine (TCM) diagnosis, ontology, and data mining in publications. Cluster four (the purple one) is in the bottom right corner and has terms from publications about the effectiveness (or lack thereof) of evidence-based research and traditional Chinese medicine (TCM), especially in relation to clinical trials. In the upper left-center quadrant, you can see Cluster 5, which is orange-highlighted, and which includes terms from publications about how TCM information is represented in IRR systems. The first dimension of published research stretches horizontally from theoretical to experimental, as shown in this conceptual structure analysis. Vertically expanding as the second dimension, this one specifies published publications and the keywords associated with them throughout a range from clinical to highly technical. Since TCM is an individualized medicine that treats each patient according to their specific pattern of disharmony, based on information mainly obtained through quantitatively-based examinations conducted by humans, there are obvious research gaps in the upper and lower left quadrants. This indicates that more scientific research is needed to support both clinical and technical IRR work in TCM theory. Western biomedicine views much of the research on the right side of the map—namely, Chinese herbal pharmaceutical medicine and its integration with biomedicine—as experimental because there has been no success in controlled trials measuring the efficacy of drugs or the results of TCM treatments. So, this map shows the way forward for TCM and IRR research: using new research paradigms and tools, like big data analytics and internet of things technologies, to combine Eastern and Western medicine in a translational approach, to learn more about how diseases develop and what impacts they have on the body.

Conclusion

In the future of both Samsung [67] and Sinclair DA, et al. [68], a huge, all-encompassing network of medical IRR systems will be behind and underneath our increasingly monitored, emphasized, and refined lives as a way to attain a new state of existence. It may be time for the broad discipline of information studies to emerge from its proverbial shell. Such frameworks have started to materialize in different parts of the universe:

Automakers have started to program themselves, and algorithms that understand our tastes, habits, goals, and bioactivities are being fine-tuned.

in order to get where they're going; drivers are finding out where their passengers are going even before they meet them; and music is continuously being indexed and streamed globally, regardless of language or culture. For the user, device, network, and provider to engage in a digital, behind-the-scenes dance of mobile technology driven by capitalism, all of these systems need underlying IRR architecture. One can only speculate as to whether, in the future, we might be able to solve half of our medical problems by combining TCM and IRR, by applying these conceptual frameworks to problems similar to those ancient TCM practitioners encountered, but with a focus on health and wellbeing rather than profit and market share. This is an admirable goal to strive toward, especially since that the prevalence of many diseases—including chronic ones like obesity and mental illness—seems to be rising. Therefore, further study into TCM is necessary, especially concerning IRR and how it connects to the need to anchor contemporary clinical and technological research in classical theory. Perhaps TCM's future lies in a big data design and deployment methodology with a new medical IRR at its core. This methodology would be translational in nature, reducing health care costs and improving medical outcomes worldwide, much like Jim Gray's vision of scientific discovery, which is driven by the collection, analysis, and comprehension of digital data by an ever-increasing interdisciplinary community of both professional and citizen-like scientists [69].

Acknowledgements

This paper and the research components related to information representation and retrieval would not have been possible without the support of the author's doctoral advisor, Dr. Heting Chu, professor in the Palmer iSchool of Library and Information Science at Long Island University.

References

1. Chen H, Wu Z, Huang C, Xu J (2003) TCM-Grid: weaving a medical grid for traditional Chinese medicine. *International Conference on Computational Science* 2659: 1143-1152.
2. Brashier KE (2014) *Public Memory in Early China*. Harvard University Asia Center, USA, pp. 526.
3. Pregadio F (2008) Zhang Ji (Zhang Zhongjing) In: Selin H (edr), *Encyclopaedia of the History of Science, Technology, and Medicine in Non-Western Cultures*. Springer, Netherlands, pp. 2416.
4. Van Wyk BE, Wink M (2004) *Medicinal plants of the world*. CABI, USA, pp. 520.
5. Yan JF (1991) *DNA and the I Ching: The Tao of Life*. North Atlantic Books, USA, pp. 193.
6. Weiss SM, Kulikowski CA, Amarel S, Safir A (1978) A model-based method for computer-aided medical decision-making. *Artif Intell* 11(1-2): 145-172.
7. Samsung (2019) *KX50: The Future in Focus*.
8. Liang Z, Liu J, Ou A, Zhang H, Li Z, et al. (2019) Deep generative learning for automated EHR diagnosis of traditional Chinese medicine. *Computer Methods and Programs in Biomedicine* 174: 17-23.

9. Yang H, Lyu MR, King I (2013) Efficient online learning for multitask feature selection. *ACM Trans Knowl Discov Data* 7(2): 6.
10. Leydier Y, Ouji A, LeBourgeois F, Emptoz H (2009) Towards an omnilingual word retrieval system for ancient manuscripts. *Pattern Recognit* 42(9): 2089-2105.
11. Long H, Zhu Y, Jia L, Gao B, Liu J, et al. (2019) An ontological framework for the formalization, organization, and usage of TCM-Knowledge. *BMC Med Inform Decis Mak* 19(2): 53.
12. Aria M, Cuccurullo C (2017) Bibliometrix: An R-tool for comprehensive science mapping analysis. *J Informetr* 11(4): 959-975.
13. Aria M, Cuccurullo C (2019) Package bibliometrix: an R-tool for comprehensive science mapping analysis.
14. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M (2009) KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res* 38: D355-D360.
15. Tang JL, Zhan SY, Ernst E (1999) Review of randomised controlled trials of traditional Chinese medicine. *BMJ* 319(7203): 160-161.
16. Wu B, Liu M, Liu H, Li W, Tan S, et al. (2007) Meta-analysis of traditional Chinese patent medicine for ischemic stroke. *Stroke* 38(6): 1973-1979.
17. Chan K, Shaw D, Simmonds MS, Leon CJ, Xu Q, et al. (2012) Good practice in reviewing and publishing studies on herbal medicine, with special emphasis on traditional Chinese medicine and Chinese materia medica. *J Ethnopharmacol* 140(3): 469-475.
18. Pang B, Zhang D, Wang K (2005) The bi-elliptical deformable contour and its application to automated tongue segmentation in Chinese medicine. *IEEE Trans Med Imaging* 24(8): 946-956.
19. Zhou X, Wu Z, Yin A, Wu L, Fan W, et al. (2004) Ontology development for unified traditional Chinese medical language system. *Artif Intell Med* 32(1): 15-27.
20. Fang YC, Huang HC, Chen HH, Juan HF (2008) TCMGeneDIT: a database for associated traditional Chinese medicine, gene and disease information using text mining. *BMC Complement Altern Med* 8(1): 58.
21. Qiao X, Hou T, Zhang W, Guo S, Xu X (2002) A 3D structure database of components from Chinese traditional medicinal herbs. *J Chem Inf Comput Sci* 42(3): 481-489.
22. Wojcikowski K, Johnson DW, Gobe G (2006) Herbs or natural substances as complementary therapies for chronic kidney disease: ideas for future studies. *J Lab Clin Med* 47(4): 160-166.
23. Zhou X, Peng Y, Liu B (2010) Text mining for traditional Chinese medical knowledge discovery: a survey. *J Biomed Inform* 43(4): 650-660.
24. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M (2009) KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res* 38: D355-D360.
25. Perianes Rodriguez A, Waltman L, Van Eck NJ (2016) Constructing bibliometric networks: A comparison between full and fractional counting. *J Informetr* 10(4): 1178-1195.
26. SCImago (2019) SJR-SCImago Journal & Country Rank [Portal].
27. Zhang J, Yu Q, Zheng F, Long C, Lu Z, et al. (2016) Comparing keywords plus of WOS and author keywords: A case study of patient adherence research. *J Assoc Inf Sci Technol* 67(4): 967-972.
28. Berry LL, Parasuraman A (1993) Building a new academic field-The case of services marketing. *J Retail* 69(1): 13-60.
29. Garfield E (2004) Historiographic mapping of knowledge domains literature. *J Inf Sci* 30(2): 119-145.
30. Tang JL, Zhan SY, Ernst E (1999) Review of randomised controlled trials of traditional Chinese medicine. *BMJ* 319(7203): 160-161.
31. Flower A, Lewith GT, Little P (2007) Seeking an oracle: using the Delphi process to develop practice guidelines for the treatment of endometriosis with Chinese herbal medicine. *J Altern Complement Med* 13(9): 969-976.
32. Sampson M, Cogo E, Ajiferuke I, Manheimer E, Campbell K, et al. (2011) Searching for controlled trials of complementary and alternative medicine: a comparison of 15 databases. *Evid Based Complement Alternat Med* 2011: 1-9.
33. Jiang M, Zhang C, Zheng G, Guo H, Li L, et al. (2012) Traditional Chinese medicine Zheng in the era of evidence-based medicine: a literature analysis. *Evid Based Complement Alternat Med* 2012: 1-9.
34. Chang IM (2001) Anti-aging and health-promoting constituents derived from traditional oriental herbal remedies: information retrieval using the TradiMed 2000 DB. *Ann N Y Acad Sci* 928(1): 281-286.
35. Boehm K, Raak C, Vollmar HC, Ostermann T (2010) An overview of 45 published database resources for complementary and alternative medicine. *Health Info Libr J* 27(2): 93-105.
36. Bensoussan A, Myers SP, Drew AK, Whyte IM, Dawson AH (2002) Development of a Chinese herbal medicine toxicology database. *J Toxicol Clin Toxicol* 40(2): 159-167.
37. Wang JF, Zhou H, Han LY, Chen X, Chen YZ, et al. (2005) Traditional Chinese medicine information database. *Clin Pharmacol Ther* 78(1): 92-93.
38. Qiao X, Hou T, Zhang W, Guo S, Xu X (2002) A 3D structure database of components from Chinese traditional medicinal herbs. *J Chem Inf Comput Sci* 42(3): 481-489.
39. Zhou X, Wu Z, Yin A, Wu L, Fan W, et al. (2004) Ontology development for unified traditional Chinese medical language system. *Artif Intell Med* 32(1): 15-27.
40. Zhou X, Peng Y, Liu B (2010) Text mining for traditional Chinese medical knowledge discovery: a survey. *J Biomed Inform* 43(4): 650-660.
41. You M, Chen C, Li GZ, Yan SX, Sun S, et al. (2015) ISMAC: an intelligent system for customized clinical case management and analysis. *Scientific World Journal* 2015: 1-12.
42. Wan H, Moens MF, Luyten W, Zhou X, Mei Q, et al. (2016) Extracting relations from traditional Chinese medicine literature via heterogeneous entity networks. *J Am Med Inform Assoc* 23(2): 356-365.
43. Yu T, Liu J, Yang S, Li J, Jia L (2016) Semantic Web for Knowledge Integration between Traditional Chinese Medicine and Biomedicine. 2015 7th International Conference on Information Technology in Medicine and Education (ITME), China.
44. Chen H, Mao Y, Zheng X, Cui M, Feng Y, et al. (2007) Towards semantic e-science for traditional Chinese medicine. *BMC Bioinformatics* 8(3): S6.
45. Chen H, Wang Y, Wu Z (2008) Introduction to semantic e-Science in biomedicine. *BMC Bioinformatics* 8(3): 1-3.
46. Tse HYG, Li VWT, Hui MN, Kwok Chan P, Han Cheng S (2008) Data mining for chinese materia medica and pharmacological research. *J Biomol Screen* 13(5): 390-395.
47. Bensoussan A, Myers SP, Drew AK, Whyte IM, Dawson AH (2002) Development of a Chinese herbal medicine toxicology database. *J Toxicol Clin Toxicol* 40(2): 159-167.
48. Qiao X, Hou T, Zhang W, Guo S, Xu X (2002) A 3D structure database of components from Chinese traditional medicinal herbs. *J Chem Inf Comput Sci* 42(3): 481-489.
49. Li Y, Li S, Lu AP (2006) Comparative analysis via data mining on the clinical features of Western medicine and Chinese medicine in diagnosing rheumatoid arthritis. *Zhongguo Zhong Xi Yi Jie He Za Zhi* 26(11): 988-991.
50. Fang YC, Huang HC, Chen HH, Juan HF (2008) TCMGeneDIT: a database for associated traditional Chinese medicine, gene and disease information using text mining. *BMC Complement Altern Med* 8(1): 58.
51. Zhang X, Zhou X, Huang H, Chen S, Liu B (2010) A hierarchical symptom-herb topic model for analyzing traditional Chinese medicine clinical diabetic data. 2010 IEEE 3rd International Conference on Biomedical Engineering and Informatics, USA.

52. Jiang Z, Zhou X, Zhang X, Chen S (2012) Using link topic model to analyze traditional Chinese medicine clinical symptom-herb regularities. 2012 IEEE 14th International Conference on e-Health Networking, Applications and Services (Healthcom), China.
53. Chen X, Chen H, Bi X, Gu P, Chen J, et al. (2014) BioTCM-SE: a semantic search engine for the information retrieval of modern biology and traditional Chinese medicine. *Comput Math Methods Med* 2014: 1-13.
54. Chen H, Chen X, Gu P, Wu Z, Yu T (2014) OWL reasoning framework over big biological knowledge network. *Biomed Res Int* 2014: 272915.
55. Wang L, Zhang Y, Zhang Y, Xu X, Cao S (2017) Prescription Function Prediction Using Topic Model and Multilabel Classifiers. *Evid Based Complement Alternat Med* 2017: 8279109.
56. Ka WF (2004) Online research databases and journals of Chinese medicine. *J Altern Complement Med* 10(6): 1123-1128.
57. May BH, Yang AW, Zhang AL, Owens MD, Bennett L, et al. (2009) Chinese herbal medicine for Mild Cognitive Impairment and Age Associated Memory Impairment: a review of randomised controlled trials. *Biogerontology* 10(2): 109-123.
58. May BH, Lu C, Xue CC (2012) Collections of traditional Chinese medical literature as resources for systematic searches. *J Altern Complement Med* 18(12): 1101-1107.
59. May BH, Feng M, Zhou IW, Chang SY, Lu SC, et al. (2016) Memory impairment, dementia, and Alzheimer's disease in classical and contemporary traditional Chinese medicine. *J Altern Complement Med* 22(9): 695-705.
60. Xiong X (2015) Integrating traditional Chinese medicine into Western cardiovascular medicine: an evidence-based approach. *Nat Rev Cardiol* 12(6): 374.
61. Xiong X, Wang P, Li X, Zhang Y (2015) The effect of Chinese herbal medicine Jian Ling Decoction for the treatment of essential hypertension: a systematic review. *BMJ Open* 5(2): e006502.
62. Liu YQ, Wang YX, Shi NN, Han XJ, Lu AP (2017) Current situation of International Organization for Standardization/Technical Committee 249 international standards of traditional Chinese medicine. *Chin J Integr Med* 23(5): 376-380.
63. Yoon SH, Kim YS, Jo HG, Kwon CY (2019) Current usage of terminologies related to acupuncture: a literature research and standardization suggestion. *Chin J Integr Med* 25(2): 147-150.
64. Abdi H, Valentin D (2007) Multiple correspondence analysis. In: Salkind NJ (edr), *Encyclopedia of Measurement and Statistics*. Sage Publishing, USA, pp. 651-657.
65. Cuccurullo C, Aria M, Sarto F (2016) Foundations and trends in performance management. A twenty-five years bibliometric analysis in business and public administration domains. *Scientometrics* 108(2): 595-611.
66. Zhou X, Wu Z, Yin A, Wu L, Fan W, et al. (2004) Ontology development for unified traditional Chinese medical language system. *Artif Intell Med* 32(1): 15-27.
67. Samsung (2019) KX50: The Future in Focus.
68. Sinclair DA, LaPlante MD (2019) *Lifespan: Why we Age-and Why We Don't Have To*. Atria Books, USA, pp. 310.
69. Bourne PE (2013) The reaming of life: based on the 2010 Jim Gray eScience Award Lecture. *Concurr Comput* 25(4): 445-453.