DEVELOPMENT OF RESEARCH TRENDS EVOLUTION MODEL FOR COMPUTER SCIENCE FOR MALAYSIAN PUBLICATION

NURUL MARDHIAH AZURA BINTI MD NADZAR

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy

School of Computer Science
Faculty of Engineering
Universiti Teknologi Malaysia

APRIL 2022

DEDICATION

Alhamdullilah. Thanks to Allah for giving me such an opportunity, bless me to learn a lit bit of His knowledge and courage, strength and braveness to continue until the end of this journey. I dedicate this thesis, definitely to myself, for not giving up in the middle of the journey and keeping on pushing until it is finished. Also, to My husband, Norhady M Amin. thanks a lot for letting me explore another degree, your patience, love, support, for always being on my side and being the greatest man. To my parents and siblings, Md Nadzar Ab Rahman, Norhan Jaafar, Mohd Azrul Nadziadin, Nurul Hani Azrina & Mohd Kamil Azran. Thanks for all your doa and support for all my success.

PhD's journey was super hard.

Indeed.

ACKNOWLEDGEMENT

In preparing this thesis, I was in contact with many people, researchers, academicians, and practitioners. They have contributed to my understanding and thoughts. In particular, I wish to express my sincere appreciation to my main supervisor, Dr Aryati Bakri, for her encouragement, guidance, critics and friendship. I am forever grateful and thankful for her support from the beginning until the end of this research and study. I am also very thankful to my co-supervisor Professor Madya Dr Roliana Ibrahim for their guidance, advice and motivation. Without their continued support and interest, this thesis would not have been the same as presented here.I am also indebted to Kementerian Pengajian Tinggi Malaysia for funding my PhD study. Librarians at UTM also deserve special thanks for their assistance in supplying the relevant literature. My fellow postgraduate student should also be recognised for their support. My sincere appreciation also extends to all my colleagues and others who have assisted. Their views and tips are useful indeed. Unfortunately, it is impossible to list all of them in this limited space. I am grateful to all my family members.

ABSTRACT

Nowadays, there seem to be research trends done on studies that manipulate publications that utilise the text mining approach. However, most of these studies only investigated the gaps faced by existing research trends models, and the execution of text mining of bibliometric elements and the timeline windows representing the "trends" was not clarified. Thus, this study aimed to develop the conceptual model for research trends in Malaysian publications, specifically, to incorporate the text element of bibliometrics and the execution of timeline windows to identify research trends. In the context of research trends, the evolution or growth of some research area from one period to another is important. This included what has happened, what is currently happening, and predicting potential research trends that will happen in the near future for others to continue the research development. The element in the newly developed model was extracted from the literature review and adapted from one of the selected models. The new model consisted of three stages which is the first stage consisted of three elements - selecting document collection; the second stage was the selection of the bibliometric element; and the third stage was the execution of text mining, co-word analysis from the selected textual bibliometric element, the implementation of two timeline windows (fixed time and sliding time windows-timeline). Also, the execution of the third stage required aid from tools - CiteSpace. The newly developed model was tested, and data were downloaded from two databases, Scopus (10,052 publications) and Web of Science (WoS) (22,088 publications), for a duration between 1995 and 2019. This study identified that the research trend pattern became more active from 2002 onwards. Besides that, the research topic became fresher and more unconventional throughout the timelines. Research topics on artificial intelligence, network communication, and wireless sensor networks are the hottest topics and timeless. Besides that, knowledge management, internet banking, online shopping, and eCommerce were the alternative options for computer science researchers. Each timeline's evolution and blooming shows that researchers are investigating each topic thoroughly. In addition, some small topics do not appear in fixed timeline windows but instead emerge from sliding timeline windows, such as system development, shared banking service, virtual team collaboration, and internet policy. This study also captured the highlighted keywords that could give hints or appear as an initial idea for the next research journey. Experts' evaluation and validation were executed as the interpretation of experimental results require experts' expertise, experience, and views. A semi-structured interview was done with thirteen experts who have remarkable expertise in research and development. From the discussion, most experts agreed that the model could help others identify the research trends and potential new research topics emerging for future research journeys. The newly developed model could be beneficial to those who need hints for their next exploration and help those keen to understand how to execute the text mining within the bibliometric elements.

ABSTRAK

Pada masa kini, terdapat trend penyelidikan yang dilakukan terhadap kajian yang memanipulasi penerbitan yang menggunakan pendekatan penyelidikan teks. Walau bagaimanapun, kebanyakan kajian ini hanya mengkaji jurang yang dihadapi oleh model aliran penyelidikan sedia ada, dan pelaksanaan penyelidikan teks unsur bibliometrik dan tetingkap garis masa yang mewakili "trend" tidak dijelaskan. Oleh itu, kajian ini bertujuan untuk membangunkan model konsep untuk aliran mendedahkan penyelidikan dalam penerbitan Malaysia, khususnya, untuk menggabungkan elemen teks bibliometrik dan pelaksanaan tetingkap garis masa untuk mengenal pasti trend penyelidikan. Dalam konteks trend penyelidikan, evolusi atau pertumbuhan beberapa bidang penyelidikan dari satu tempoh ke tempoh yang lain adalah penting. Ini termasuk apa yang telah berlaku, apa yang sedang berlaku, dan meramalkan potensi trend penyelidikan yang akan berlaku dalam masa terdekat untuk orang lain meneruskan pembangunan penyelidikan. Elemen dalam model yang baru dibangunkan telah diekstrak daripada kajian literatur dan disesuaikan daripada salah satu model yang dipilih. Model baharu terdiri daripada tiga peringkat iaitu peringkat pertama terdiri daripada tiga elemen - memilih koleksi dokumen; manakal peringkat kedua ialah pemilihan elemen bibliometrik; dan peringkat ketiga ialah pelaksanaan penyelidikan teks, analisis kata bersama dari elemen bibliometrik teks yang dipilih, pelaksanaan dua tetingkap garis masa (masa tetap dan masa gelongsor tingkap-garis masa). Selain itu, pelaksanaan peringkat ketiga memerlukan bantuan daripada alatan - CiteSpace. Model yang baru dibangunkan telah diuji, dan data telah dimuat turun daripada dua pangkalan data, Scopus (10,052 penerbitan) dan Web of Science (WoS) (22,088 penerbitan), untuk tempoh antara 1995 dan 2019. Kajian ini mengenal pasti bahawa corak trend penyelidikan menjadi lebih aktif dari tahun 2002 dan seterusnya. Selain itu, topik penyelidikan menjadi lebih segar dan tidak konvensional sepanjang garis masa. Topik penyelidikan tentang kecerdasan buatan, komunikasi rangkaian dan rangkaian penderia tanpa wayar ialah topik paling hangat dan abadi. Selain itu, pengurusan pengetahuan, perbankan internet, membeli-belah dalam talian, dan e-Dagang merupakan pilihan alternatif untuk penyelidik sains komputer. Evolusi dan perkembangan setiap garis masa menunjukkan bahawa penyelidik sedang mengkaji setiap topik dengan teliti. Selain itu, beberapa topik kecil tidak muncul dalam tetingkap garis masa tetap tetapi sebaliknya muncul daripada tetingkap garis masa gelongsor, seperti pembangunan sistem, perkhidmatan perbankan kongsi, kerjasama pasukan maya dan dasar internet. Kajian ini berjaya menangkap kata kunci yang diserlahkan yang berpotensi memberi petunjuk atau idea awal untuk perjalanan penyelidikan seterusnya. Penilaian dan pengesahan pakar telah dilaksanakan kerana tafsiran keputusan eksperimen memerlukan kepakaran, pengalaman dan pandangan pakar. Temu bual separa berstruktur telah dilakukan dengan 13 pakar yang mempunyai pengalaman dalam penyelidikan dan pembangunan penyelidikan. Daripada temu bual itu, kebanyakan pakar bersetuju bahawa model itu boleh membantu orang lain mengenal pasti arah aliran penyelidikan dan potensi penyelidikan baharu yang muncul untuk perjalanan penyelidikan masa depan. Model yang baru dibangunkan boleh memberi manfaat kepada mereka yang memerlukan petunjuk untuk penerokaan seterusnya dan membantu mereka yang berminat untuk memahami cara melaksanakan penyelidikan teks dalam elemen bibliometrik

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ERTM	-	Evolution of research trends model
BM	-	Bibliometrics Model

LIST OF SYMBOLS

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Overview

Research into research trends is not new in a research environment. There are various purposes of executing the study in research trends. Among all the aims was to have an idea of the movement and evolution of research in a certain period and get an idea of a new direction for new potential research. Some research aims to acknowledge the evolution of research in certain disciplines or get an idea of the movement of research happening in certain countries to grab some new picture for the next potential study or seek any collaboration opportunities between disciplines or research areas. Whatever the purpose or aim of executing research in research trends, it is all about getting the idea of current trends and what the next research area or theme they can explore and make new discoveries.

This chapter is dedicated to discussing the problem background to address the gaps of the recent study, problem statement that this study intends to solve, and research questions and objectives related to both the problem background and problem statement.

1.2 Problem Background

There are three implementations when it comes to doing the study of research trends with either reviewing the previous paper (Saha, 2018), statistical analysis (Erfanmanesh, Amin, Didegah, & Omidvar., 2017; Lei & Liu, 2019), or manipulation towards previous publication (Amado, Cortez, Rita, & Moro, 2018). Between these three implementations, this study has identified two methods used ages ago by previous researchers called citation-based (Gurzki & M.Woisetschläger, 2017; D. Yu

& Xu, 2017), and another method known as text-based (Kasemsap, 2017; Nie & Sun, 2017). For the citation-based method, normally, the execution focuses on citation activity, such as co-citation (D. Yu & Xu, 2017), bibliographic coupling (Boyack & Klavans, 2010), and also direct citation relation (Gurzki & M.Woisetschläger, 2017). Between these two methods, this study has identified that the citation-based method has some influence or bias towards identifying research trends that are related to citation activities in a publication which has been questioned by some other researchers regarding the validity of several citations as well as the intention of the source (Tyler, 2018; Zhou, Amadi, & Zhang, 2020). Moreover, numerous previous experiments and models have been developed using the citation-based method (Jeong, 2020; Tang, Chang, & Hwang, 2021; Wang. & Waltman., 2016). However, the latter process, which is named the text-based method, receives less attention from researchers (Ramírez, Sánchez-Cañizares, & Fuentes-García, 2019), and also this study identified a smaller number of models in this method. Therefore, this study intends to experiment and develop a model concerning the text-based approach.

As mentioned in the above section, the text-based method gets less attention and a smaller number of models developed by previous researchers. The current text-based model that a group of researchers created in 2019, named as 'Evolution Research Trends Model' (ERTM) (Ramírez, Sánchez-Cañizares, & Fuentes-García, 2019) and the Bibliometric Model by Qi Wang (Wang. & Waltman., 2016). Both models aimed to reveal research trends or the evolution of research trends in publications by using a different approach, as mentioned earlier. However, this study identified some limitations that can be improved in identifying research trends, such as for the Bibliometric Model, whether the input data used citation information or direct citation information (Wang. & Waltman., 2016). These types of input data have some biases in which the information influences the research group, which will affect the exposure of research trends output (Tyler, 2018; Zhou et al., 2020).

Moreover, the total number of texts can be limited from the citation information, which normally provides only the title of publication and journal's name (Ioannidis, 2019). Sometimes, it can be simple and compact, limiting the extraction of information (Aksnes, 2019; Ioannidis, 2019). Besides, other bibliometric elements can

extract insight into research trends, such as title, sub-title, authors keywords, index journal, and abstract (Aksnes, 2019; Rons, 2018). This input can potentially give an even broader coverage and exposure that will extract the truth from the words themselves and do not have any influences or inspiration from other factors in the citation information (Faust, 2018; Santonen & Conn, 2015; Zhao et al., 2018)

Next, it is related to the study which the previous research has done in research trends identification. This study identified that most of the previous studies intended to have one summary for the whole timeline (Fergnani, 2019; Lei & Liu, 2019), and it is implemented in selected models in this study too (Ramírez, Sánchez-Cañizares, & Fuentes-García., 2019; Wang. & Waltman., 2016). With the name of 'evolution', if another researcher has the intention to execute the ERTM with the illustration provided by the previous founder, certainly, it can be misunderstood. Consequences with the result of execution will be limited or not broad enough to identify the evolution or movement of potential research. Evolution can be defined as a process from one stage to another or the development of something (Boggs, 2019). Therefore, it is important to show clearly in the model which part or phase, or element needs to execute the 'evolution' experiment. It can be unfavourable for the executor to explore the research trends and any potential emerging research trends if the model is not clearly shown. This is not to disclaim that previous research did wrong to identify research trends in a specific domain. However, it will be beneficial if the research trends are done in an evolutionary way (Jeyaraj & Zadeh, 2019). Clear movement of research trends from the start to the end of the timeline is identified transparently and detailed (Jeyaraj & Zadeh, 2019; V.Mäntylä, Graziotin, & Kuutila, 2018). With this execution, it can be acknowledged by others interested in the domain or from other areas, and subsequently, will contribute to new knowledge. However, some previous research executed the experiment for research trends that were directed by each year (Akhavan, Ebrahim, Fetrati, & Pezeshkan, 2016b; Najmi, Rashidi, Abbasi, & Travis Waller, 2016). One needs to remember that research sometimes needs to be discovered and matured. Therefore, extracting the research by year without any time frame can be risky as it will be overlooked while executing the experiment. Hence, the aim for execution results needs to be clear before experimenting.

Research trends evolution supposedly happens periodically to capture the trend's movement and which trends emerge in a certain period (Kashani & Roshani., 2019; V.Mäntylä et al., 2018). Research trends move periodically depending on trends or events in a certain period (Rossetto, Bernardes, Borini, & Gattaz, 2018; V.Mäntylä et al., 2018). Therefore, if only one timeline is being done as shown in the existing model, we might lose the details on insight information of research trends. The keywords being extracted from the experiment will also be fewer. Moreover, the differentiation between these periods can give some knowledge regarding the progress, publication activity, and productivity of research and development in a certain dedicated area (Madani & Weber, 2016; Montoro-Sanchez, 2017; Rossetto et al., 2018). Certainly, it is beneficial to all related bodies concerned about productivity for research and development as a considerable amount of grants have been constantly given to researchers.

To develop the new model, it can potentially become a reference for others in executing new experiments. Therefore, this study believes that a comprehensive model is needed to guide others to ensure the investigation is conducted appropriately and gives appropriate output at the end of the experiment. Moreover, in the evolution of research trends, details of what has happened, what is currently happening, and what potentially will happen in the future are among the questions raised by researchers (Kashani & Roshani., 2019). This is to make sure that research and development are continuously happening and the idea to make another new research can be identified with some initial execution before going into depth exploration in a certain research area. Nowadays, it becomes a demand in the research environment to search or explore any possible angle with any potential problem raised in the nation and research world (Subramanian, 2017). With the evolution of the research trend model, this study believes that it has the potential to help researchers identify and acknowledge current research trends that can support their future research and development (Nie & Sun, 2017). Furthermore, from the evolution of research trends with the implementation of co-occurrences keywords, the list of appropriate keywords and the cluster of research trends are expectedly emerging from the experiment (Varga, 2018). The listed keywords that arise from the investigation can at least be supported and hints to understand more about the research trends in a certain period.

Lastly, the implementation of research trends studies in computer science is reviewed (Xiuwen Chen, Chen, Wu, Xie, & Lic, 2016). The computer science domain is also a focal point to identify research trends that either review the previous paper, statistical analysis, or manipulation towards previous publication (Akşit, 2018; Huynh, Luong, & Hoang, 2012). Each execution gives different input and impact on research trends identification (Amjad & Ali, 2019; Huynh et al., 2012; Sánchez-Prieto, 2020). All these studies that have already been implemented and experimented with have given huge benefits to interested researchers in the computer science domain (Amjad & Ali, 2019; Huynh et al., 2012; Ruiz-Real, Luis, Uribe-Toril, Torres, & Pablo, 2021; Sánchez-Prieto, 2020). It has given other researchers some idea of what had happened in the domain and the initial idea for the next research concept (Ruiz-Real et al., 2021). However, to the best knowledge of this study, most of the previous studies only utilise the citation-based method and less in the text-based approach. Therefore, this study intends to use the newly developed model in the domain and identify any possibilities of what have already been done by this domain's researchers and the chances of new topics of research and the opportunities of emerging research trends manipulated from this domain's publication.

This study intends to enhance and develop a new model that includes all possible measures and criteria to identify the research trends in a certain period and the listed keywords that are beneficial to others in their research and development

1.3 Problem Statement

Research on research trends has become a staple experiment to identify what had happened previously, but it also becomes an aid tool to identify potential new research for future discoveries. Also, it is a tool or medium to get some new idea or initial hint of keywords extracted from the experiment. Identifying research trends is proven when several developed models have the same aim and objective but with a different approach and potentially different outputs. There are two models: a citation-based model and a text-based model. This study identifies the text-based model between these two bases and has less bias than the citation-based model. However,

there is some room for improvements that can potentially be applied to text-based models, and the selected model for the text-based method is 'the model of the evolution of research trends'.

The execution of bibliometric analysis in this model was too broad; therefore, the specific element to execute the text mining approach was not included. To implement a text mining method specifically for co-word analysis, which was claimed in the previous model, some indicator or element for a total number of occurrences and keywords need to be excluded while executing' the experiment and percentage to be included in the cluster output. These are the important components for managing the co-word analysis. The details do not exist in the previous model, and it is worried that others interested in executing the model can mistranslate it. The element and component to perform the co-word analysis must be included and illustrated clearly in the model. However, these were not seen in the previous model.

Another problem identified by this study is the timeline execution as the previous model claimed about 'evolution' stated in the model's title. However, the timeline illustrated and executed from the model does not represent the 'evolution' of research trends. Evolution is the movement of an item from one stage to another stage or from one form to another form. Therefore, the timeline stated in the previous model can be misunderstood by other interested researchers. This study identified another method in determining the comprehensive research trends, which involve two different timelines. Although the study regarding these two timelines was discovered earlier, it has not been utilised in the previous model. Therefore, to ensure that other researchers are interested in identifying research trends comprehensively, this study infers that these timelines need to be included and illustrated clearly in the model.

In addition, research trends in computer science in Malaysia's scope have various ways to identify research trends executed by researchers. However, fewer studies are determining the overall evolution in computer science. Most of the studies generated one summary result, which does not involve the broad movement of research trends.

Another point is the implementation of text mining, specifically co-occurrences keywords, in identifying research trends in computer science. Undeniably, some studies were done, but the execution between these studies differed from each other. This study identified a lack of a model that can act as a guide to determine and identify research trends in computer science. Lastly, fewer studies present the keywords, which potentially gives some idea of the next embarkation journey for future research. Although it only extracts simple and short keywords, it can help some researchers have a hint to combine the research area and determine which area can be investigated further or which area needs to be revised with new technologies.

With these problems, this study believes that improvements to the selected development model mentioned earlier in this section can significantly improve in determining and identifying research trends. The new development model can perhaps act as guidance or a template for those who do not really understand how to execute the co-occurrences keywords experiment as it has some needed criteria that gives a significant output.

1.4 Research Importance

The importance of this research can be divided into two aspects. One of the aspects is developing a new evolution model of research trends to enhance the current model. The enhancement of the current model is necessary as, to the best knowledge of this study, the existence of an evolution model to identify the research trends that consist of appropriate measures and possible criteria that implement the co-occurrences keywords approach to ensure the output from the execution is close to none. In the co-occurrences keywords approach, there are several important measures and criteria to focus on to avoid any incorrect output at the end of the experiment. The comprehensive standards and criteria to enhance the current model will benefit others to better understand how to execute the model and identify research trends in publication.

Besides, to the best knowledge of this study, several other models for identifying research trends exist but with the involvement of a mathematical model. The intention is to enhance the current model without involving any mathematical model because not many people know and understand the challenges of a mathematical model. If the mathematical model is wrongly executed, the chances of getting a wrong output are high. If an incorrect work is obtained, the interpretation of the result can also be affected. Therefore, this study needs to explore and enhance the current model to make it convenient to understand and execute by others who do not know how to manage mathematical models.

Another one is executing a text-based model with detailed criteria to execute. This study identified that the text-based model is less biased than the citation-based model. However, the text-based model needs some complex elements and standards to guide others to execute and generate significant output that can help future usage. Moreover, the advantage of the text-based model is that the work is not influenced by outliers, such as the citation activity within the research group. It is a more versatile implementation, and the extraction of research trends is the truth from the published text.

By the end of the experiment, the output from the investigation will be in a cluster form. In the cluster, there are several keywords that represent the research trends. These keywords are expected to be used several times in a publication; therefore, they emerge in clusters. The appearance of these used keywords is another important matter. If the measures or criteria to implement the co-occurrences keywords are incorrectly executed, the possibility of getting inappropriate keywords as output is high. Therefore, the intention to enhance the current model with appropriate measures and criteria is important to make sure the production in identifying the evolution of research trends is beneficial for the next exploration of research and development.

Also, to the best knowledge of this study, this study's scope is to manipulate Malaysian publications in computer science to identify evolution in research trends which have never been explored before by previous researchers. Therefore, this is the first study to examine the insight of research trends in dedicated timelines by

implementing the approach of co-words analysis. The research trends from dedicated time intervals within the dedicated timeline can be acknowledged from this exploration. It can be beneficial to our researchers for future research and investigation. The new development model, which is enhanced from the current model, can act as a guide for others and is not only limited to acknowledging the evolution of research trends, but it can benefit to help to get an idea of how to fully utilise any possible keywords that may potentially emerge in the experiment's output.

1.5 Research Contribution

This research is expected to have two main contributions. One of the contributions is the evolution of research trends executed year-to-year to acknowledge the research establishment, pattern and trends, and other potential emerging research hidden in the existing Malaysian publication. This contribution of the evolution of research trends possibly can shape our current research and development and future planning. As we know, future research certainly depends on existing research. The explorations from the existing study will give huge benefits to others to design their exploration and discoveries.

Another contribution from this study is the new conceptual model for the evolution of research trends. The focus and scope for this new development model are for computing. This new model is developed based on what is currently happening in the chosen content. The model is expected to guide and maneuver those keen to know the research trends in a specific research area. This study already grouped and detailed each element to ensure that others correctly understand how to execute this model. It is important to know how this model works to get a good impact for the next interpretation stage. This study also looks forward to enhancing if there is any lack in aspect or element that will strengthen this new model.

1.6 Research Question

To solve the mentioned problems, several research questions were raised to assist this study in developing the new model:

- RQ 1: What are the criteria and existing models in identifying research trends?
- RQ 2: What is the research trend that emerged from the selected publication in Computer Science in 25 years for Malaysian publication?

1.7 Research Objective

The development of the research question leads to the following set of research objectives:

Aim: To reveal the research trends that potentially emerge from Malaysian publications.

- RO 1: To develop a conceptual model for research trends within Malaysian publication.
- RO 2: To identify the research trends in the Computer Science discipline in Malaysian publication.

1.8 Research Hypothesis

Several hypotheses need to be noticed from the new development conceptual model to evolve research trends throughout this study.

- H1: The accumulation of the text element in bibliometrics can give insightful knowledge with the execution of specific techniques.

- H2: Implementing two window timelines in the model can give more insight information regarding research trends instead of executing one window timeline only.

1.9 Scope of Research

With the development of a conceptual model for the evolution of research trends, several scopes are required to set the boundary or framework for this study.

First and foremost, the scope is the study. This study will apply and manipulate data related to this country – Malaysia. The model will be developed according to Malaysian publication behaviour and pattern. Certainly, each country has its approach to measuring its productivity, and here in Malaysia, the higher learning institution has its indicator to measure the movement of research. Therefore, all related publications which inter-related the development of the new model will be selected. All journals that publish within Malaysia's scope with the utilisation of Malaysia's data will be included.

The new conceptual model for the evolution of research trends is mainly developed to identify research movement and publication progress. The publications used in this study primarily came from journal databases, namely Scopus and Web of Science. Most research publications from higher learning education in Malaysia were published in both databases. Therefore, the number of publications related to Malaysia is expectedly more elevated than other databases. Besides that, the type of publication consists of 'articles', 'conference presiding' and 'review paper'.

The selected discipline is the last scope for developing a new conceptual model. Since research trends research has quite huge coverage and is versatile enough to be implemented in numerous fields, this study has decided to focus on a small scale. Computer science has become the selected discipline for this study. The utmost reason to choose this computing area is the relationship and interconnection with another

field. The computer science area was claimed to have two views, one of them is to focus human perception towards the technologies, and the other one is to utilise the technologies to help human's daily tasks. Therefore, this study reckoned to get the interesting output of research trends in computer science from the experiment and the analysis.

1.10 Structure of Thesis

This thesis comprises seven chapters. The first three chapters deal with the introduction, literature review and research methodology of the study. The next three chapters are concerned with the analysis, findings and results related to research objectives. The last chapter presents the conclusions, research contributions, and future research directions. An overview of all these chapters is given below.

Chapter 1

Chapter one provides an overall view of the present study. The introduction section presents a general discussion on the topic of interest and its importance and motivations for the study. Based on the literature search in the international and national context, it raises issues related to the stated research problem and provides a comprehensive background of research issues. A set of research objectives, related questions, significance, expected contributions and scope of the study are also presented here.

Chapter 2

This chapter discusses the relevant literature in a few sections. The first section is regarding the bibliometric analysis as the scope of this study is mainly for publication. This section will review indicators and elements consisting of bibliographic information. The review is vital to understand how other previous studies identify the evolution of research trends and which features they used to manipulate the publication to emerge the research trends. The piece plays a part of the main role

in the development of a new conceptual model in chapter 3. The second section is a review of research trends. The aim was to understand how other researchers execute their experiments in identifying research trends in publication. Malaysia's journal scope for research trends was also reviewed in this section. The third section is regarding theories that can be adapted in developing the model. One chosen approach potentially becomes a base for the new development model. The new model will use-dependent and independent factors for the idea. And as for the third section is regarding any other model in identifying research trends already developed by other researchers. The chosen model becomes a base model. This study will adapt the selected model and enhance it according to the scope and environment of this study.

Chapter 3

This chapter is mostly regarding methodology on how this study will be conducted. The first section in this chapter is about research design, paradigm and method. This study implements the interpretive paradigm to understand the social phenomena within the mentioned scope. The second section is the operational framework as guidance to achieve the goals and objectives in this study. The illustration of the working framework is also included in this chapter. The next section is mapping the research objectives and questions from chapter 1 with the strategies for achieving and executing the experiment. Database, year selection, data collection criteria, type of document, and data type will also be reported in chapter 3. Later, all downloaded data is presented accordingly. And as for the last section, the development of the new model will also be explained in this chapter. The phases started with the element for the new model until how to execute the model. The latest development model will be illustrated after all steps are explained. Lastly, the hypothesis from the new model acts as an aim to answer the research question and research objectives.

Chapter 4

The initial study and analysis of productivity of research performance are fully described here. This is the preliminary study and result that drives the intention to do more investigations in this research. This preliminary study aims to identify the

performance and productivity of research performance in the Malaysian context. The conclusion for the study will be at the end of the chapter. This chapter also answers one of the research objectives: identifying research trends in Malaysian publications. The analysis's result shows what has been done by our research within 20 years and the emerging research trends that might be able to be implemented in Malaysia's research and development activities. In this chapter, the most crucial analysis, such as the cluster formed in the experiment, a list of potentially beneficial keywords for our researchers to use for their next exploration, and the timeline of each research trend.

Chapter 5

This chapter presents the evaluation and validation for the newly developed model, obtained result, and interpretation from the effect that the researcher did throughout the research. Validation and views from several experts were also explained as well. This is one of the important chapters in this research to make sure that the model is understandable by experts and the interpretation result from the experiment is accepted by experts. A series of questionnaires for semi-structured interviews are also documented in this chapter too. The questionnaires were built from the literature review, which the researcher identified throughout this research.

Chapter 6

Finally, the conclusions, research contributions, future research directions, limitations of the study and concluding remarks are presented in the last chapter.

1.11 Conclusion

As for conclusion in this chapter, the background of the study, problem statement, research objectives and question are reviewed. It is important to grab the attention of what has been done by other researchers and compare the study with other possible bodies that monitor productivity for research and development. From the background of the study, then only comes the problem statement. In this study, two

problem statements are raised, as mentioned in the above section. With the expanded problem statement, then come the objectives of this study. The objectives of this study are to act as guidance for this study to achieve and solve the problem statement. Later, the research question is mentioned to guide the research objectives in this chapter and an overview of each chapter that consists of this study.

REFERENCES

Abramo., G., Cicero., T., & D'Angelo., C. A. (2011). Assessing the varying level of impact measurement accuracy as a function of the citation window length. *Journal of Informetrics, 5*(4), 659-667. doi:10.1016/j.joi.2011.06.004

Abrizah, & Mee., C. W. (2017). Malaysia's Computer Science research productivity based on publications in the Web of Science, 2000-2010. *Malaysian Journal of Library & Information Science, 16*(1), 109-124.

Adie, E., & Roe, W. (2013). Altmetric enriching scholarly content with article level discussion & metrics.pdf. *Learned Publishing, 26*(1), 11-17. doi:http://dx.doi.org/10.1087/20130103

Afify, Y. M., Moawad, I. F., Badr, N. L., & Tolba, M. F. (2017). Cloud services publication and discovery. *Handbook of research on machine learning innovations and trends*, 204-228.

Aggarwal, C. C., & Zhai, C. X. (2012). An Introduction to Text Mining. *Mining text data*, 1-10. doi:10.1007/978-1-4614-3223-4_1

Akhavan, P., Ebrahim, N. A., Fetrati, M. A., & Pezeshkan, A. (2016a). Major trends in knowledge management research: a bibliometric study. *Scientometrics, 107*(3), 1249-1264.

Akhavan, P., Ebrahim, N. A., Fetrati, M. A., & Pezeshkan, A. (2016b). Major trends in knowledge management research: a bibliometric study. *Scientometrics, 107*(3), 1249–1264. doi:10.1007/s11192-016-1938-x

Akoum, I. (2016). Research, development and innovation in Malaysia: Elements of an effective growth model. *Asian Economic and Financial Review, 6*(7), 390.

Akşit, M. (2018). The role of computer science and software technology in organizing universities for industry 4.0 and beyond. *Federated Conference on Computer Science and Information Systems (FedCSIS)*, 5-11.

Aksnes, D. W., Liv Langfeldt, and Paul Wouters. (2019). Citations, citation indicators, and research quality: An overview of basic concepts and theories. *SAGE Open, 9*(1), 2158244019829575.

Al-Jabi, S. W. (2019). Global research trends in West Nile virus from 1943 to 2016: a bibliometric analysis. *Globalization and Health, 13*(1), 1-9.

Albert, N., & Thomson, M. (2018). A Synthesis of the Consumer-Brand Relationship Domain: Using Text Mining to Track Research Streams, Describe Their Emotional Associations, and Identify Future Research Priorities. *Journal of the Association for Consumer Research, 3*(2), 130-146.

Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques. *arXiv preprint arXiv*, 1707.02919

Amado, A., Cortez, P., Rita, P., & Moro, S. (2018). Research trends on Big Data in Marketing: A text mining and topic modeling based literature analysis. *European Research on Management and Business Economics, 24*(1), 1-7. doi:10.1016/j.iedeen.2017.06.002

Amjad, T., & Ali, A. (2019). Uncovering diffusion trends in computer science and physics publications. *Library Hi Tech.*

Amran, F. H., Rahman, I. h. K. A., Salleh, K., Ahmad, S. N. S., & Haron, N. H. (2014). Funding trends of research universities in Malaysia. *Procedia - Social and Behavioral Sciences 164*, 126 – 134. doi:10.1016/j.sbspro.2014.11.060

Apte, C., Damerau, F., Weiss, S. M., Apte, C., Damerau, F., & Weiss, S. (1998). Text mining with decision trees and decision rules. *In In Proceedings of the Conference on Automated Learning and Discorery, Workshop 6: Learning from Text and the Web.*

Aryadoust, V., & Ang, B. H. (2019). Exploring the frontiers of eye tracking research in language studies: a novel co-citation scientometric review. *Computer Assisted Language Learning*, 1-36. doi:10.1080/09588221.2019.1647251

Aryadoust, V., Tan, H. A. H., & Ng, L. Y. (2019). A Scientometric review of Rasch measurement: The rise and progress of a specialty. *Frontiers in Psychology, 20*, 2197. doi:10.3389/fpsyg.2019.02197

Azam, N., & Yao, J. T. (2012). Comparison of term frequency and document frequency based feature selection metrics in text categorization. *Expert Systems with Applications, 39*(5), 4760-4768. doi:10.1016/j.eswa.2011.09.160

Bache, K., Newman, D., & Smyth, P. (2013). Text-based measures of document diversity. *KDD '13 Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining* 23-31.

Bailey, B., Mark Budden, Michael Dorff, and Urmi Ghosh-Dastidar. (2009). Undergraduate research: How do we begin. *MAA Focus, 29*(1), 14-16.

Bakri, A., & Zakaria, A. M. U. (2013). Malaysian Publication Productivity in the Field of Engineering using Scopus A Bibliometric Study. *International journal of Innovative Computing*.

Balili, C., Segev, A., & Lee, U. (2017). Tracking and Predicting the Evolution of Research Topics in Scientific Literature. *IEEE International Conference on Big Data (Big Data)*, 1694-1697.

Bamidis, P. D. (2017). Internet of things in health trends through bibliometrics and text mining. *Informatics for health: Connected citizen-led wellness and population health, 235*(73).

Bell, J. (2014). Doing Your Research Project: A guide for first-time researchers. *McGraw-Hill Education (UK)*.

Besselaar, P. V. D., & Heimeriks, G. (2006). Mapping research topics using word-reference co-occurrences: A method and an exploratory case study. *Scientometrics, 68*(3), 377–393. doi:10.1007/s11192-006-0118-9

Bhattacharjee, P., & Mitra., P. (2021). A survey of density based clustering algorithms. *Frontiers of Computer Science, 15*(1), 1-27.

Biljecki, F. (2016). A scientometric analysis of selected GIScience journals. *International Journal of Geographical Information Science, 30*(7), 1302-1335. doi:10.1080/13658816.2015.1130831

Bloch, C., & Sørensen., M. P. (2014). The size of research funding: Trends and implications. *Science and Public Policy, 42*(1), 30-43. doi:10.1093/scipol/scu019

Boggs, G. L. (2019). Revolution and evolution. *NYU Press*.

Bojović, S., Matić, R., Popović, Z., Smiljanić, M., Stefanović, M., & Vidaković, V. (2014). An overview of forestry journals in the period 2006-2010 as basis for ascertaining research trends. *Scientometrics, 98*(2), 1331-1346. doi:10.1007/s11192-013-1171-9

Bornmann, L. (2014). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics, 8*(4), 895-903. doi:10.1016/j.joi.2014.09.005

Bornmann, L., Mutz, R., Hug, S. E., & Daniel, H.-D. (2011). A multilevel meta-analysis of studies reporting correlations between the h index and 37 different h index variants. *Journal of Informetrics, 5*(3), 346-359. doi:10.1016/j.joi.2011.01.006

Boyack, K. W., & Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology, 61*(12), 2389-2404. doi:10.1002/asi.21419

Brik, B., Bettayeb, B., & M'hammed Sahnoun, a. F. D. (2019). Towards predicting system disruption in industry 4.0: Machine learning-based approach. *Procedia Computer Science, 151*, 667-674.

Campello, R. J., Kröger, P., Sander, J., & Zimek, A. (2020). Density-based clustering. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10*(2), e1343.

Cartes-Velásquez, R., & Delgado, C. M. (2014). Bibliometric analysis of articles published in ISI dental journals, 2007-2011. *Scientometrics, 98*(3), 2223-2233. doi:10.1007/s11192-013-1173-7

Chadegani, A. A., Salehi, H., Yunus, M. M., Farhadi, H., Fooladi, M., Farhadi, M., & Ebrahim, N. A. (2013). A Comparison between Two Main Academic Literature Collections: Web of Science and Scopus Databases. *Asian Social Science, 9*(5). doi:10.5539/ass.v9n5p18

ChaomeiChen. (2005). CiteSpace Quick Guide version 1.2.

Charidimou, A., Fox, Z., Werring, D. J., & Song, M. (2016). Mapping the landscape of cerebral amyloid angiopathy research: an informetric analysis perspective. *J Neurol Neurosurg Psychiatry, 87*(3), 252-259. doi:10.1136/jnnp-2015-310690

Chen, C. (2005). Measuring the Quality of Network Visualization. *In Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL'05)*, 405-405.

Chen, G., & Xiao, L. (2016). Selecting publication keywords for domain analysis in bibliometrics: A comparison of three methods. *Journal of Informetrics, 10*(1), 212-223. doi:10.1016/j.joi.2016.01.006

Chen, R.-C., Liang, J.-Y., & Ren-HaoPan. (2000). Using recursive ART network to construction domain ontology based on term frequency and inverse document frequency. *Expert Systems with Applications, 34*(1), 488-501. doi:10.1016/j.eswa.2006.09.019

Chen, X., Chen, B., Zhang, C., & Hao, T. (2017). Discovering the recent research in natural language processing field based on a statistical approach. *Springer,*

Cham(In International Symposium on Emerging Technologies for Education), 507-517.

Chen, X., Chen, J., Wu, D., Xie, Y., & Li, J. (2016). Mapping the Research Trends by Co-word Analysis Based on Keywords from Funded Project. *Procedia Computer Science, 91*, 547-555. doi:10.1016/j.procs.2016.07.140

Chen, X., Chen, J., Wu, D., Xie, Y., & Lic, J. (2016). Mapping the Research Trends by Co-word Analysis Based on Keywords from Funded Project. *Procedia Computer Science, 91*, 547-555. doi:10.1016/j.procs.2016.07.140

Choi, S., & Seo, J. (2021). Trends in Healthcare Research on Visual Impairment and Blindness: Use of Bibliometrics and Hierarchical Cluster Analysis. *Ophthalmic Epidemiology* 1-8. doi:10.1080/09286586.2020.1863993

Clarivate. (2020). Reclassification of Papers in Multidisciplinary Journals for Creating Field Baselines. *Web of Science*. Retrieved from https://incites.help.clarivate.com/Content/Research-Areas/wos-reclass-papers-multidiscipline-journals.htm

Cobo, M. J., Wang, W., Laengle, S., Merigó, J. M., Yu, D., & Herrera-Viedma, E. (2018). Co-words Analysis of the Last Ten Years of the International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems. *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, 667-677. doi:10.1007/978-3-319-91479-4_55

Conlan, K., Baggili, I., & Breitinger, F. (2016). Anti-forensics: Furthering digital forensic science through a new extended, granular taxonomy. *Digital investigation, 18*, S66-S75.

Corrales-Garay, D., Ortiz-de-Urbina-Criado, M., & Mora-Valentín, E.-M. (2019). Knowledge areas, themes and future research on open data: A co-word analysis. *Government information quarterly, 36*(1), 77-87.

Costa, C., Schurr, U., Loreto, F., Menesatti, P., & Carpentier, S. (2019). Plant phenotyping research trends, a science mapping approach. *Frontiers in Plant Science, 9*(1933).

Davoudian, A., Chen, L., & Liu, M. (2018). A survey on NoSQL stores. *ACM computing surveys (CSUR), 51*(2), 1-43.

Dev, C. S., Parsa, H. G., Parsa, R. A., & Bujisic, M. (2015). Assessing Faculty Productivity by Research Impact: Introducing Dp2 Index. *Journal of Teaching in Travel and Tourism, 15*(2), 93-124. doi:10.1080/15313220.2015.1026471

Dhillon, I. S. (2001). Co-clustering documents and words using bipartite spectral graph partitioning. *Proceeding KDD '01 Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining* 269-274 doi:10.1145/502512.502550

Di Leo, G., & Sardanelli, F. (2020). Statistical significance: p value, 0.05 threshold, and applications to radiomics—reasons for a conservative approach. *European radiology experimental, 4*, 1-8.

Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research, 133*, 285-296. doi:10.1016/j.jbusres.2021.04.070

Dorta-Gonzalez, P., & Dorta-González, M.-I. (2013). Impact maturity times and citation time windows: The 2-year maximum journal impact factor. *Journal of Informetrics, 7*(3), 593-602.

Eassom, H. (2017). How to Choose Effective Keywords for Your Article. *Discover the Future of Research*. Retrieved from https://hub.wiley.com/community/exchanges/discover/blog/2017/06/07/how-to-choose-effective-keywords-for-your-article

Eck, N. J. v., & Waltman, L. (2011). Text mining and visualization using VOSviewer. *Centre for Science and Technology Studies, Leiden University, The Netherlands*, 1-5.

Elsbach, K. D., & Stigliani, I. (2019). New Information Technology and Implicit Bias. *Academy of Management Perspectives, 33*(2). doi:10.5465/amp.2017.0079

EPU, E. P. U. (2015). Eleventh Malaysia plan, 2016-2020: Anchoring growth on people. *Putrajaya: Prime Minister's Department*.

Erfanmanesh, Amin, M., Didegah, F., & Omidvar., S. (2017). Research productivity and impact of Library and Information Science in the Web of Science. *Malaysian Journal of Library & Information Science, 15*(3), 85-95.

Erfanmanesh, M., & Abdullah, A. (2018). Mapping worldwide research on the Internet of Things during 2011-2016. *The Electronic Library*. doi:/10.1108/EL-09-2017-0196

Faust, O. (2018). Documenting and predicting topic changes in Computers in Biology and Medicine: A bibliometric keyword analysis from 1990 to 2017. *Informatics in Medicine Unlocked, 11*, 15-27. doi:10.1016/j.imu.2018.03.002

Feng, J., Zhang, Y. Q., & Zhang, H. (2017). Improving the co-word analysis method based on semantic distance. *Scientometrics, 111*(3), 1521–1531. doi:10.1007/s11192-017-2286-1

Fergnani, A. (2019). Mapping futures studies scholarship from 1968 to present: A bibliometric review of thematic clusters, research trends, and research gaps. *Futures, 105*, 104-123. doi:10.1016/j.futures.2018.09.007

Filardo, G., Graca, B. d., Sass, D. M., Pollock, B. D., Smith, E. B., & Martinez, M. A.-M. (2016). Trends and comparison of female first authorship in high impact medical journals: observational study (1994-2014). *bmj, 352*, i847. doi:10.1136/bmj.i847

Fop, M., & Murphy, T. B. (2018). Variable selection methods for model-based clustering. *Statistics Surveys, 12*, 18-65.

Fu, X., Niu, Z., & Yeh, M.-K. (2016). Research trends in sustainable operation: a bibliographic coupling clustering analysis from 1988 to 2016. *Cluster Computing, 19*(4), 2211-2223.

Garousi, V., & Mäntylä, M. V. (2016). Citations, research topics and active countries in software engineering: A bibliometrics study. *Computer Science Review, 19*, 56-77. doi:10.1016/j.cosrev.2015.12.002

Geminiani, A., Ercoli, C., Feng, C., & Caton., J. G. (2014). Bibliometrics study on authorship trends in periodontal literature from 1995 to 2010. *Journal of Periodontology, 85*(5), e136-e143. doi:10.1902/jop.2013.130354

Gerdsri, N., Kongthon, A., & Puengrusme, S. (2017). Profiling the Research Landscape in Emerging Areas Using Bibliometrics and Text Mining: A Case Study of Biomedical Engineering (BME) in Thailand. *International Journal of Innovation and Technology Management, 14*(2), 1740011. doi:10.1142/S0219877017400119

Ghouila, A., Siwo, G. H., Entfellner, J.-B. D., Panji6, S., Button-Simons, K. A., Davis, S. Z., . . . Mulder, N. (2018). Hackathons as a means of accelerating scientific discoveries and knowledge transfer. *Genome research, 28*(5), 759-765. doi:10.1101/gr.228460.117

Glanzel, W. (2003). Bibliometrics as a research field a course on theory and application of bibliometric indicators. *Course Handouts.*

Glenisson, P., Glänzel, W., Janssens, F., & Moor, B. D. (2005). Combining full text and bibliometric information in mapping scientific disciplines. *Information Processing & Management, 41*(6), 1548-1572. doi:10.1016/j.ipm.2005.03.021

González-Alcaide, G., Llorente, P., & Ramos, J. M. (2016). Bibliometric indicators to identify emerging research fields: publications on mass gatherings. *Scientometrics, 109*(2), 1283-1298. doi:10.1007/s11192-016-2083-2

González, M. I. D., & González, P. D. (2016). Do fixed citation windows affect the impact maturation rates of scientific journals? *Investigacion Bibliotecologica: Archivonomia, Bibliotecologia e Información 30*(68), 73-89. doi:10.1016/j.ibbai.2016.06.004

Grix, J. (2018). The foundations of research. *Macmillan International Higher Education.*

Gudanowska, A. E. (2017). Transformation towards Industry 4.0-identification of research trends and aspect of necessary competences in the light of selected publications. *Research in Logistics & Production, 7.*

Guenther, L., & Joubert, M. (2017). Science communication as a field of research : identifying trends, challenges and gaps by analysing research papers. *JCOM: Journal of Science Communication, 16*(2), 1-19. doi:10.22323/2.16020202

Gurzki, H., & M.Woisetschläger, D. (2017). Mapping the luxury research landscape: A bibliometric citation analysis. *Journal of Business Research, 77*, 147-166. doi:10.1016/j.jbusres.2016.11.009

Gutiérrez-Salcedo, M., Martínez, M. Á., Moral-Muñoz, J. A., Herrera-Viedma, E., & Cobo., M. J. (2018). Some bibliometric procedures for analyzing and evaluating research fields. *Applied Intelligence, 48*, 1275–1287.

Haa, J. H., & Riffe, D. (2015). Crisis-related research in communication and business journals: An interdisciplinary review from 1992 to 2011. *Public Relations Review, 41*(4), 569-578. doi:10.1016/j.pubrev.2015.06.019

Hao, T., Chen, X., & Li, G. (2018). A bibliometric analysis of text mining in medical research. *Soft Computing, 22*, 7875–7892. doi:10.1007/s00500-018-3511-4

Haunschild, R., Bornmann, L., & Marx, W. (2016). Climate change research in view of bibliometrics. *Plos One, 11*(7), e0160393. doi:10.1371/journal.pone.0160393

He, Q. (1999). Knowledge discovery through co-word analysis *Library Trends, 48*(1), 133-159.

He, Q., Chen, B., Pei, J., Qiu, B., Mitra, P., & Giles, L. (2009). Detecting topic evolution in scientific literature: how can citations help? *CIKM '09: Proceedings of the 18th ACM conference on Information and knowledge management*, 957–966. doi:10.1145/1645953.1646076

Hearst, M. (2003). What is text mining. *SIMS, UC Berkeley 5*.

Hilmi, M. F., & Mustapha, Y. (2013). Eleven Years (1999–2009) of The Malaysian Journal of Distance Education: A Bibliometric Study. *Malaysian Journal of Distance Education, 15*(1), 1-12.

Hood, W. W., & Wilson, C. S. (2011). The literature of bibliometrics, scientometrics,and informetrics.pdf. *Scientometrics, 52*(2), 291–314. doi:10.1023/A:1017919924342

Hoz-Correa, A. l., Muñoz-Leiva, F., & Bakucz, M. (2018). Past themes and future trends in medical tourism research: A co-word analysis. *Tourism Management, 65*, 200-211. doi:10.1016/j.tourman.2017.10.001

Hu, J., & Zhang, Y. (2015). Research patterns and trends of Recommendation System in China using co-word analysis. *Information Processing & Management, 51*(4), 329-339. doi:10.1016/j.ipm.2015.02.002

Huang, M.-H., & Chang, C.-P. (2013). Detecting research fronts in OLED field using bibliographic coupling with sliding window. *Scientometrics, 98*(3), 1721-1744. doi:10.1007/s11192-013-1126-1

Huang, M.-H., & Chang, C.-P. (2016a). A comparative study on three citation windows for detecting research fronts. *Scientometrics, 109*(3), 1835-1853.

Huang, M.-H., & Chang, C.-P. (2016b). A comparative study on three citation windows for detecting research fronts. *Scientometrics, 109*, 1835–1853. doi:10.1007/s11192-016-2133-9

Huang, Z.-X., Tian, H.-Y., Hu, Z.-F., Zhou, Y.-B., Zhao, J., & Yao, K.-T. (2008). GenCLiP: a software program for clustering gene lists by literature profiling and constructing gene co-occurrence networks related to custom keywords. *BMC Bioinformatics, 9*(308). doi:10.1186/1471-2105-9-308

Hung, J. l. (2012). Trends of e-learning research from 2000 to 2008: Use of text mining and bibliometrics. *British Journal of Educational Technology, 43*(1), 5-16. doi:10.1111/j.1467-8535.2010.01144.x

Huynh, T., Luong, H., & Hoang, K. (2012). Integrating bibliographical data of computer science publications from online digital libraries. *Intelligent Information and Database Systems, 7198*, 226-235.

Ioannidis, J. P., Jeroen Baas, Richard Klavans, and Kevin W. Boyack. (2019). A standardized citation metrics author database annotated for scientific field. *PLoS biology, 17*(8), e3000384.

Iwami, S., Mori, J., Sakata, I., & Kajikawa, Y. (2014). Detection method of emerging leading papers using time transition. *Scientometrics, 101*(2), 1515-1533. doi:10.1007/s11192-014-1380-x

Iwami, S., & Sakata, I. (2015). Growing Topics and Emerging Researches in Science. *The International Society for Professional Innovation Management.*

J, P.-D., L, C.-G., D, C.-M., & R, B.-M. (2014). Text Mining of Scientific Big Data for Decision Making in Conservation of Mediterranean Marine Biodiversity. *Journées d'Intelligence Économique-BIG DATA MININ*

J.McLachlan, G. (2015). Computation: Expectation-Maximization Algorithm. *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*, 469-474.

Jacso, P. (2005). As we may search – Comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced database. *Current Science, 89*(9), 1537.

Jae Yun, L., Kim, H., & Pan Jun, K. (2010). Domain analysis with text mining: Analysis of digital library research trends using profiling methods. *Journal of Information Science, 36*(2), 144-161. doi:10.1177/0165551509353251

Jalali, M. S., Ashouri, A., & Zhang., O. H.-R. (2016). Information diffusion through social networks: . *Expert Systems with Applications, 44*, 187-197. doi:10.1016/j.eswa.2015.09.014

Jankowski, N. W. (2006). Creating community with media: History, theories and scientific investigations. *The handbook of new media, 55-74.*

Jeong, C., Sion Jang, Eunjeong Park, and Sungchul Choi. (2020). A context-aware citation recommendation model with BERT and graph convolutional networks. *Scientometrics, 124*(3), 1907-1922.

Jeyaraj, A., & Zadeh, A. H. (2019). Evolution of information systems research: Insights from topic modeling. *Information & Management*, 103207. doi:10.1016/j.im.2019.103207

Jia, W., Peng, J., & Cai, N. (2020). An approach to improving the analysis of literature data in Chinese through an improved use of Citespace. *Knowledge Management & E-Learning, 12*(2). doi:10.34105/j.kmel.2020.12.013

Jiang, S., Prasad, A., Kan, M.-Y., & Sugiyama, K. (2018). Identifying Emergent Research Trends by Key Authors and Phrases. *Proceedings of the 27th International Conference on Computational Linguistics*(259–269).

Jiang, Y., Ritchie, B. W., & Benckendorff, P. (2019). Bibliometric visualisation: An application in tourism crisis and disaster management research. *Current Issues in Tourism, 22*(16), 1925-1957. doi:10.1080/13683500.2017.1408574

Jing, L., Ng, M. K., Xu, J., & Huang, J. Z. (2005). Subspace Clustering of Text Documents with Feature Weighting K-Means Algorithm. *PAKDD 2005: Advances in Knowledge Discovery and Data Mining* 802-812. doi:10.1007/11430919_94

Jo, Y., Hopcroft, J. E., & Lagoze., C. (2011). The web of topics: discovering the topology of topic evolution in a corpus. *In Proceedings of the 20th international conference on World wide web*, 257-266.

Kalantari, A., Kamsin, A., Kamaruddin, H. S., Ebrahim, N. A., Gani, A., Ebrahimi, A., & Shamshirband, S. (2017). A bibliometric approach to tracking big data research trends. *Journal of Big Data, 4*(1), 30. doi:10.1186/s40537-017-0088-1

Kamdem, J. P., Duarte, A. E., Lima, K. R. R., Rocha, J. B. T., Hassan, W., Barros, L. M., . . . Tsopmo, A. (2019). Research trends in food chemistry: A bibliometric review of its 40 years anniversary (1976–2016). *Food Chemistry, 294*, 448-457.

Kameshwaran, K., & Malarvizhi, K. (2014). Survey on clustering techniques in data mining. *International Journal of Computer Science and Information Technologies, 5*(2), 2272-2276.

Kasemsap, K. (2017). Text mining: Current trends and applications. *Web data mining and the development of knowledge-based decision support systems*, 338-358.

Kashani, E. S., & Roshani., S. (2019). Evolution of innovation system literature: Intellectual bases and emerging trends. *Technological Forecasting and Social Change, 146*, 68-80. doi:10.1016/j.techfore.2019.05.010

Kaur, A., Singh, V. P., & Gill, S. S. (2018). The future of cloud computing: opportunities, challenges and research trends. *IEEE*, 213-219.

Kawalec, A. (2017). Research trends in library and information science based on Spanish scientific publication 2000 to 2010. *Malaysian Journal of Library & Information Science, 18*(2), 1-13.

Kayser, V., & Blindb, K. (2017). Extending the knowledge base of foresight: The contribution of text mining. *Technological Forecasting and Social Change, 116*. doi:10.1016/j.techfore.2016.10.017

Kazakis, N. A. (2014). Bibliometric evaluation of the research performance of the Greek civil engineering departments in National and European context. *Scientometrics, 101*(1), 505-525. doi:10.1007/s11192-014-1326-3

Kenchakkanavar, A. Y. (2014). Types of e-resources and its utilities in library. *International journal of information sources and services, 1*(2), 97-104.

Khan, G. F., & Wood., J. (2015). Information technology management domain: emerging themes and keyword analysis. *Scientometrics 105, 105*(2), 959-972. doi:10.1007/s11192-015-1712-5

Khasseh, A. A., Soheili, F., Moghaddam, H. S., & MousaviChelak, A. (2017). Intellectual structure of knowledge in iMetrics: A co-word analysis. *Information Processing & Management, 53*(3). doi:10.1016/j.ipm.2017.02.001

Kim, B.-Y., & Delen, D. (2018). Medical informatics research trend analysis: A text mining approach. *Health informatics journal, 24*(4), 432-452.

Kim, D., & Oh, A. (2011). Topic Chains for Understanding a News Corpus. *International Conference on Intelligent Text Processing and Computational Linguistics*, 163-176.

Kim, J. H. (2017). A Review of Cyber-Physical System Research Relevant to the Emerging IT Trends: Industry 4.0, IoT, Big Data, and Cloud Computing. *Journal of Industrial Integration and Management, 2*(3). doi:10.1142/S2424862217500117

Kolb, D. A., Boyatzis, R. E., & Mainemelis, C. (2001). Experiential learning theory: Previous research and new directions. *Perspectives on thinking, learning, and cognitive styles, 1*(8), 227-247.

Kostoff, R. N. (1993). Co-Word Analysis. *Evaluating R&D Impacts: Methods and Practice*, 63-78.

Kostoff, R. N., Toothman, D. R., Eberhart, H. J., & Humenik, J. A. (2001). Text mining using database tomography and bibliometrics: A review. *Technological*

Forecasting and Social Change, 68(3), 223-253. doi:10.1016/S0040-1625(01)00133-0

Kriegel, H. P., Kröger, P., Sander, J., & Zimek, A. (2011). Density-based clustering. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 1*(3), 231-240.

Kulkarni, D., and Herbert A. Simon. (1988). The processes of scientific discovery: The strategy of experimentation. *Cognitive science, 12*(2), 139-175.

Kwona, S., Liub, X., Portera, A. L., & Youtiea, J. (2019). Research Addressing Emerging Technological Ideas Has Greater Scientific Impact. *Research Policy, 48*(9), 103834.

Lacasse, J. R., Hodge, D. R., & Bean, K. F. (2011). Evaluating the productivity of social work scholars using the h-index. *Research on Social Work Practice, 21*(5), 599-607. doi:10.1177/1049731511405069

Lahiri, S., Choudhury, S. R., & Caragea, C. (2014). Keyword and Keyphrase Extraction Using Centrality Measures on Collocation Networks. *arXiv:1401.6571 [cs.CL]*.

Lamberton, C., & Stephen, A. T. (2016). A thematic exploration of digital, social media, and mobile marketing: Research evolution from 2000 to 2015 and an agenda for future inquiry." *Journal of Marketing, 80*(6), 146-172. doi:10.1509/jm.15.0415

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., & Hoffmann, M. (2014). Industry 4.0. *Business & Information Systems Engineering, 6*(4), 239-242.

Le, M.-H., Ho, T.-B., & Nakamori, Y. (2005). Detecting Emerging Trends from Scientific Corpora. *International Journal of Knowledge and Systems Sciences, 2*(2).

Leahey, E. (2016). From sole investigator to team scientist: Trends in the practice and study of research collaboration. *Annual review of sociology, 42*, 81-100.

Lee, J. H., & Segev, A. (2012). Knowledge maps for e-learning. *Computers & Education, 59*(2), 353-364. doi:10.1016/j.compedu.2012.01.017

Lei, L., & Liu, D. (2019). The research trends and contributions of System's publications over the past four decades (1973–2017): A bibliometric analysis. *System, 80*, 1-13. doi:10.1016/j.system.2018.10.003

Li, C., Ji, X., & Luo, X. (2019). Phytoremediation of Heavy Metal Pollution: A Bibliometric and Scientometric Analysis from 1989 to 2018. *International*

Journal of Environmental Research and Public Health, 16(23), 4755. doi:10.3390/ijerph16234755

Li, J., & Chen, C. (2016). CiteSpace: Text mining and visualization in scientific literature. *Capital University of Economics and Business Press: Beijing, China* 149-152.

Li, M., & Chu, Y. (2016). Explore the research front of a specific research theme based on a novel technique of enhanced co-word analysis. *Journal of Information Science, 43*(6), 725-741. doi:10.1177/0165551516661914

Li, M., & Chu, Y. (2017). Explore the research front of a specific research theme based on a novel technique of enhanced co-word analysis. *Journal of Information Science, 43*(6), 725-741. doi:10.1177/0165551516661914

Li, M., Yang Qin, Bing Liu, and Xiaowen Chu. (2021). Enhancing the efficiency and scalability of blockchain through probabilistic verification and clustering. *Information Processing & Management, 58*(5), 102650.

Li, W., & Zhao, Y. (2015). Bibliometric analysis of global environmental assessment research in a 20-year period. *Environmental Impact Assessment Review, 50*, 158-166. doi:10.1016/j.eiar.2014.09.012

Liang, T.-P., & Liu, Y.-H. (2018). Research Landscape of Business Intelligence and Big Data analytics: A bibliometrics study. *Expert Systems with Applications, 111*, 2-10. doi:doi.org/10.1016/j.eswa.2018.05.018

Lin, T.-J., Lin, T.-C., Potvin, P., & Tsai, C.-C. (2018). Research trends in science education from 2013 to 2017: a systematic content analysis of publications in selected journals. *International Journal of Science Education* 1-21. doi:10.1080/09500693.2018.1550274

Lin, W., Wu, Z., Lin, L., Wen, A., & Li, J. (2017). An ensemble random forest algorithm for insurance big data analysis. *IEEE Access, 5*, 16568-16575.

Lin, Y.-S., Jiang, J.-Y., & Lee, S.-J. (2014). A Similarity Measure for Text Classification and Clustering. *IEEE Transactions on Knowledge and data Engineering, 26*(7), 1575 - 1590. doi:10.1109/TKDE.2013.19

Liu, B., Chen, H., & Huang, X. (2018). Map Changes and Theme Evolution in Work Hours: A Co-Word Analysis. *International journal of environmental research and public health 15*(5), 1039. doi:10.3390/ijerph15051039

Liu, R., Huang, W., Fei, Z., Wang, K., & Liang, J. (2019). Constraint-based clustering by fast search and find of density peaks. *Neurocomputing, 330* 223-237.

Liu, S., Sun, Y.-P., Gao, X.-L., & Sui, Y. (2019). Knowledge domain and emerging trends in Alzheimer's disease: a scientometric review based on CiteSpace analysis. *Neural regeneration research, 14*(9), 1643–1650. doi:10.4103/1673-5374.255995

Liu, Y., Goncalves, J., Ferreira, D., Xiao, B., Hosio, S., & Kostakos, V. (2014). Chi 1994-2013 Mapping Two Decades of Intellectual Progress through Co-word Analysis. *Proceeding CHI '14 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* 3553-3562. doi:10.1145/2556288.2556969

Lopez-Munoz, F., Povedano-Montero, F. J., Chee, K.-Y., Shen, W. W., Fernandez-Martin, P., Garcia-Pacios, J., . . . Alamo, C. (2018). A Bibliometric Analysis of Scientific Production on Second-Generation Anti-Psychotic Drugs in Malaysia. *Malaysian Journal of Medical Sciences, 25*(3), 40-55.

López-Robles, J. R., Jose Ramón Otegi-Olaso, Igone Porto-Gómez, Hamurabi Gamboa-Rosales, and Nadia Karina Gamboa-Rosales. . (2020). Understanding the intellectual structure and evolution of Competitive Intelligence: A bibliometric analysis from 1984 to 2017. *Technology Analysis & Strategic Management, 32*(5), 604-619. doi:10.1080/09537325.2019.1686136

MacDermid, J. C., Fung, E. H., & Law, M. (2015). Bibliometric analyses of physical and occupational therapy faculty across Canada indicate productivity and impact of rehabilitation research. *Physiotherapy Canada, 67*(1), 76-84. doi:10.3138/ptc.2013-71BC

MacDonald, K. I., & Dressler, V. (2018). Using Citation Analysisto Identify Research Fronts: A Case Study with the Internet of Things. *Science & Technology Libraries, 37*(2), 171-186. doi:10.1080/0194262X.2017.1415183

Macnamara, J. (2010). Emergent communication practices. *The 21st century media (r) evolution:*.

Madani, F., & Weber, C. (2016). The evolution of patent mining: Applying bibliometrics analysis and keyword network analysis. *World Patent Information, 46*, 32-48. doi:10.1016/j.wpi.2016.05.008

Martí-Parreño, J., Méndez-Ibáñez, E., & Alonso-Arroyo, A. (2016). The use of gamification in education: a bibliometric and text mining analysis. *Journal of computer assisted learning, 32*(6), 663-676. doi:10.1111/jcal.12161

Martín-Martín, A., Orduna-Malea, E., Ayllón, J. M., & López-Cózar, E. D. (2016). The counting house, measuring those who count : Presence of Bibliometrics, Scientometrics, Informetrics, Webometrics and Altmetrics in Google Scholar Citations, ResearcherID, ResearchGate, Mendeley,& Twitter. *arXiv:1602.02412v1, 21*.

Md Khudzari, J., Kurian, J., Tartakovsky, B., & Raghavan, G. S. V. (2018). Bibliometric analysis of global research trends on microbial fuel cells using Scopus database. *Biochemical Engineering Journal, 136*, 51-60. doi:10.1016/j.bej.2018.05.002

Meaney, C., Moineddin, R., Voruganti, T., O'Brien, M. A., Krueger, P., & Sullivan., F. (2016). Text mining describes the use of statistical and epidemiological methods in published medical research. *Journal of Clinical Epidemiology, 74*, 124-132. doi:10.1016/j.jclinepi.2015.10.020

Meilă, M., & Heckerman, D. (2001). An experimental comparison of model-based clustering methods." *Machine learning, 42*(1), 9-29.

Meystre, S., and Peter J. Haug. . (2005). Automation of a problem list using natural language processing. *Bmc Medical Informatics and Decision Making, 5*(1), 1-14.

Mike, T., & Pardeep, S. (2012). Webometric research with the Bing Search API 2.0. *Journal of Informetrics, 6*(1), 44-52. doi:10.1016/j.joi.2011.10.002

Montoro-Sanchez, I. D.-V. A. (2017). Research evolution in science parks and incubators: foundations and new trends. *Scientometrics, 110*, 1243–1272. doi:10.1007/s11192-016-2218-5

Moon, T. K. (1996). The expectation-maximization algorithm. *IEEE Signal processing magazine 13*(6), 47-60.

Morris, A. (2002). From Idea to Impact: A Guide to the Research Process. Building Effective Research, 1. Research Report. *Learning and Skills Development Agency, Regent Arcade House*.

MOSTI. (2015). Ministry of Science Technology And Innovation - Science and Technology Knowledge Productivity in Malaysia - Bibliometric 2015. Retrieved from http://www.data.gov.my/data/ms_MY/dataset/number-of-articles-and-proceedings-and-publication-growth-2001-2014-2

MOSTI. (2020). National survey of research and development (R&D) in Malaysia 2019. *Ministry of Science, Technology And Innovation*.

Murtagh, F., & Contreras, P. (2012). Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2*(1), 86-97.

Murtagh, F., & Contreras, P. (2017). Algorithms for hierarchical clustering: an overview, II. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 7*(6), e1219.

Nadzar, N. M. A. M., Bakri, A., & Ibrahim, R. (2017). A Bibliometric Mapping of Malaysian Publication using Co-Word Analysis. *International Journal of Advances in Soft Computing & Its Applications, 9*(3), 90-113.

Najmi, A., Rashidi, T. H., Abbasi, A., & Travis Waller, S. (2016). Reviewing the transport domain: an evolutionary bibliometrics and network analysis. *Scientometrics, 110*(2), 843-865. doi:10.1007/s11192-016-2171-3

Nejad, A. M., Qaracholloo, M., & Rezaei, S. (2020). Iranian doctoral students' shared experience of English-medium publication: the case of humanities and social sciences. *Higher Education, 80*, 255–271

Newell, W. H., Wentworth, J., & Sebberson, D. (2001). A theory of interdisciplinary studies. *Issues in Interdisciplinary Studies*.

Nguyen, V. T., Kravets, A. G., & Duong, T. Q. (2021). Predicting Research Trend Based on Bibliometric Analysis and Paper Ranking Algorithm." *Cyber-Physical Systems: Digital Technologies and Applications, 109*.

Nicholson, S. (2006). The basis for bibliomining: Frameworks for bringing together usage-based data mining and bibliometrics through data warehousing in digital library services. *Information Processing & Management, 42*(3), 785-804. doi:10.1016/j.ipm.2005.05.008

Nie, B., & Sun, S. (2017). Using Text Mining Techniques to Identify Research Trends: A Case Study of Design Research. *Applied Sciences, 7*(4), 401. doi:10.3390/app7040401

Niles, M. T., Lesley A. Schimanski, Erin C. McKiernan, and Juan P. Alperin. (2019). Why we publish where we do: Faculty publishing values and their relationship to review, promotion and tenure expectations. *BioRxiv*, 706622.

Olmeda-Gómez, C., Ovalle-Perandones, M.-A., & Perianes-Rodríguez, A. (2017). Co-word analysis and thematic landscapes in Spanish information science literature, 1985–2014. *Scientometrics, 113*(1), 195–217. doi:10.1007/s11192-017-2486-8

Onan, A., Korukoğlu, S., & Bulut, H. (2016). Ensemble of keyword extraction methods and classifiers in text classification. *Expert Systems with Applications, 57*. doi:10.1016/j.eswa.2016.03.045

Pakgohar, N., Javad Eshaghi Rad, Gholamhossein Gholami, Ahmad Alijanpour, and David W. Roberts. . (2021). A comparative study of hard clustering algorithms for vegetation data.". *Journal of Vegetation Science, 32*(3), e13042.

Papatheodorou, C., Kapidakis, S., Sfakakis, M., & Vassiliou3, A. (2003). Mining User Communities in Digital Libraries. *Information Technology and Libraries, 22*(4), 152-157.

Peng, T., Liu, L., & Zuo, W. (2014). PU text classification enhanced by term frequency–inverse document frequency-improved weighting. *Concurrency and computation: practice and experience, 26*(3), 728–741 doi:10.1002/cpe.3040

Peters, I., & Bar-Ilan, J. (2014). Informetrics, bibliometrics, altmetrics What is it all about. *Proceedings of the American Society for Information Science and Technology, 51*(1), 1-4.

Petrus, J. (2019). Soft and Hard Clustering for Abstract Scientific Paper in Indonesian. *In 2019 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, 131-136.

Piotrowski, C. (2013). Citation analysis for the modern instructor: An integrated review of emerging research. *Journal of Educators Online, 10*(2). Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-84880336306&partnerID=40&md5=6f51226f19bdfb7cbdc99654afe0187f

Popper, K. (2005). The logic of scientific discovery. *Routledge*.

Pratikno, A. S. (2018). Comparative research on top five universities' research productivity in Indonesia and Malaysia. *Stanisław Juszczyk, 17*.

Pritchard, A. (1969). Statistical bibliography or bibliometrics. *Journal of Documentation, 25*, 348.

Rahman, N. A. A., & Hassan, M. S. H. (2015). Social media research trends in Malaysia: an analysis of three major communication journals in Malaysia. 1-18.

Ramírez, L. J. C., Sánchez-Cañizares, S. M., & Fuentes-García, F. J. (2019). Past Themes and Tracking Research Trends in Entrepreneurship: A Co-Word, Cites and Usage Count Analysis *Luis Javier Cabeza Ramírez, Sandra M. Sánchez-*

*Cañizares * [OrcID] and Fernando J. Fuentes-García, 11*, 3121. doi:10.3390/su11113121

Ramírez, L. J. C., Sánchez-Cañizares, S. M., & Fuentes-García., F. J. (2019). Past Themes and Tracking Research Trends in Entrepreneurship: A Co-Word, Cites and Usage Count Analysis. *Sustainability, 11*(11), 3121.

Ravikumar, S., Agrahari, A., & Singh, S. N. (2015). Mapping the intellectual structure of scientometrics: a co-word analysis of the journal Scientometrics (2005–2010). *Scientometrics, 102*(929). doi:10.1007/s11192-014-1402-8

Reis, T. L., Mathias, M. A. S., & Oliveira, O. J. d. (2017). Maturity models: identifying the state-of-the-art and the scientific gaps from a bibliometric study. *Scientometrics 2*(110), 643-672.

Rogers, E. M. (1976). New product adoption and diffusion. *Journal of consumer Research, 2*(4), 290-230.

Rokach, L., & Maimon, O. (2005). Clustering methods. *Data mining and knowledge discovery handbook*, 321-352.

Rons, N. (2018). Bibliometric approximation of a scientific specialty by combining key sources, title words, authors and references. *Journal of Informetrics, 12*(1), 113-132. doi:10.1016/j.joi.2017.12.003

Ros, F., & Guillaume, S. (2019). A hierarchical clustering algorithm and an improvement of the single linkage criterion to deal with noise. *Expert Systems with Applications, 128*, 96-108. doi:10.1016/j.eswa.2019.03.031

Rossetto, D. E., Bernardes, R. C., Borini, F. M., & Gattaz, C. C. (2018). Structure and evolution of innovation research in the last 60 years: Review and future trends in the field of business through the citations and co-citations analysis. *Scientometrics, 115*(3), 1329-1363. doi:10.1007/s11192-018-2709-7

Rousseau, F., & Vazirgiannis, M. (2015). Main Core Retention on Graph-of-Words for Single-Document Keyword Extraction. *European Conference on Information Retrieval*, 382-393.

Rudd, M. A. (2017). What a Decade (2006–15) Of Journal Abstracts Can Tell Us about Trends in Ocean and Coastal Sustainability Challenges and Solutions. *Frontiers in Marine Science, 4*. doi:10.3389/fmars.2017.00170

Ruiz-Real, Luis, J., Uribe-Toril, J., Torres, J. A., & Pablo, J. D. (2021). Artificial intelligence in business and economics research: trends and future. *Journal of Business Economics and Management, 22*(1), 98-117.

S.Bhanuse, S., D.Kamble, S., & Kakde, S. M. (2016). Text Mining using Metadata for Generation of Side information. *Procedia Computer Science, 78*, 807 – 814.

Sagar, A., Kademani, B. S., Bhanumurthy, K., & Ramamoorthy, N. (2014). Research Trends in Radioisotopes: A Scientometric Analysis (1993-2012). *DESIDOC Journal of Library & Information Technology, 34*(4), 349-358. doi:10.14429/djlit.34.7833

Saha, B. (2018). Green computing: current research trends. *International Journal of Computer Sciences and Engineering, 6*(3), 467-469.

Saisubramanian, S., Galhotra, S., & Zilberstein., S. (2020). Balancing the tradeoff between clustering value and interpretability. *In Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 351-357.

Sajana, T., Rani, C. S., & Narayana, K. V. (2016). A Survey on Clustering Techniques for Big Data Mining. *Indian Journal of Science and Technology 9*(3). doi:10.17485/ijst/2016/v9i3/75971

Saket, J. S., & Pandya, D. S. (2016). An overview of partitioning algorithms in clustering techniques. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), 5*(6), 1943-1946.

Salleh, N., and Azlin Nordin. (2014). Trends and perceptions of evidence-based software engineering research in Malaysia. *The 5th International Conference on Information and Communication Technology for The Muslim World (ICT4M)*, 1-6.

Sánchez-Prieto, J. C., Adriana Gamazo, Juan Cruz-Benito, Roberto Therón, and Francisco J. García-Peñalvo. (2020). AI-driven assessment of students: Current uses and research trends. *In International Conference on Human-Computer Interaction, Springer, Cham*, 292-302.

Santonen, T., & Conn, S. (2015). Research Topics at ISPIM: Popularity-based Scientometrics keyword analysis. *ISPIM Conference Proceedings. The International Society for Professional Innovation Management (ISPIM), 1*(15).

Saxena, A., Prasad, M., Gupta, A., Bharill, N., Patel, O. P., Tiwari, A., . . . Lin, C.-T. (2017). A review of clustering techniques and developments. *Neurocomputing, 267*, 664-681.

Sedighi, M., & Jalalimanesh, A. (2017). Mapping research trends in the field of knowledge management. *Malaysian Journal of Library & Information Science 19, no. 1 (2017). 19*(1).

Shen, W., Wang, Y., Bai, X., Wang, H., & Latecki, L. J. (2013). Shape clustering: Common structure discovery. *Pattern Recognition, 46*(2), 539-550.

Shier, M. L., & Handy, F. (2014). Research trends in nonprofit graduate studies: A growing interdisciplinary field. *Nonprofit and Voluntary Sector Quarterly, 43*(5), 812-831.

Small, H., Boyack, K. W., & Klavans, R. (2014). Identifying emerging topics in science and technology. *Research Policy, 43*(8), 1450-1467. doi:10.1016/j.respol.2014.02.005

Soares, V. H. A., Campello, R. J., Nourashrafeddin, S., Milios, E., & Naldi., M. C. (2019). Combining semantic and term frequency similarities for text clustering. *Knowledge and Information Systems, 61*(3), 1485-1516.

Sohrabi, B., Vanani, I. R., Jalali, S. M. J., & Abedin., E. (2019). Evaluation of research trends in knowledge management: a hybrid analysis through burst detection and text clustering. *Journal of Information & Knowledge Management, 18*(4), 1950043.

Sohrabi, B., Vanani, I. R., & Namavar, M. (2019). Investigation of Trends and Analysis of Hidden New Patterns in Prominent News Agencies of Iran Using Data Mining and Text Mining Algorithms. *Webology, 16*(1).

Soleymani, M., Garcia, D., Jou, B., Schuller, B., Chang, S.-F., & Pantic, M. (2017). A survey of multimodal sentiment analysis. *Image and Vision Computing, 65*(3-14).

Steinbach, M., Ertöz, L., & Kumar, V. (2004). The challenges of clustering high dimensional data. *In New directions in statistical physics,, Springer, Berlin, Heidelberg*, 273-309.

Su, H.-N., & M.Moaniba, I. (2017). Investigating the dynamics of interdisciplinary evolution in technology developments. *Technological Forecasting and Social Change, 122*, 12-23. doi:10.1016/j.techfore.2017.04.024

Subramanian, A. M., and Pek-Hooi Soh. (2017). Linking alliance portfolios to recombinant innovation: The combined effects of diversity and alliance experience. *Long Range Planning, 50*(5), 636-652.

Sun, L., & Yin, Y. (2018). Discovering themes and trends in transportation research using topic modeling. *Transportation Research Part C: Emerging Technologies, 77*(2017), 49-66.

Synnestvedt, M. B., Chen, C., & Holmes, J. H. (2005). CiteSpace II: Visualization and Knowledge Discovery in Bibliographic Databases. *AMIA Annual Symposium Proceedings Archive*, 724–728.

Taheriyan, M. (2011). Subject classification of research papers based on interrelationships analysis. *Proceeding KDMS '11 Proceedings of the 2011 workshop on Knowledge discovery, modeling and simulation* 39-44. doi:10.1145/2023568.2023579

Tahira, M., Alias, R. A., & Bakri, A. (2013). Scientometric assessment of engineering in Malaysians universities. *Scientometrics, 96*(3), 865-879. doi:10.1007/s11192-013-0961-4

Tang, K.-Y., Chang, C.-Y., & Hwang, G.-J. (2021). Trends in artificial intelligence-supported e-learning: a systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*, 1-19.

Toit, A. S. D. (2015). Competitive intelligence research: An investigation of trends in the literature.

Trotta, D., & Garengo, P. (2018). Industry 4.0 key research topics: A bibliometric review. *2018 7th International Conference on Industrial Technology and Management (ICITM)*, 113-117. doi:10.1109/ICITM.2018.8333930

Trstenjak, B., Mika, S., & Donko, D. (2014). KNN with TF-IDF based Framework for Text Categorization. *Procedia Engineering, 69*, 1356-1364. doi:10.1016/j.proeng.2014.03.129

Trujillo, C. M., & Long, T. M. (2018). Document co-citation analysis to enhance transdisciplinary research. *Science Advances, 4*(1), e1701130. doi:10.1126/sciadv.1701130

Tyler, T. (2018). Citation metrics and impact factors fail as measures of scientific quality, in particular in taxonomy, and are biased by biological discipline and by geographic and taxonomic factors. *Finnish Zoological and Botanical Publishing Board, 55*(1-3), 185-191. doi:10.5735/085.055.0123

Úbeda-Sánchez, Manuel, Á., Fernández-Cano, A., & Callejas., Z. (2019). Inferring hot topics and emerging educational research fronts. *On the Horizon.*

V.Mäntylä, M., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review, 27*, 16-32. doi:10.1016/j.cosrev.2017.10.002

V.Smirnova, N., TheresaLillis, & KristinaHultgren, A. (2021). English and/or Russian medium publications? A case study exploring academic research writing in contemporary Russian academia. *Journal of English for Academic Purposes, 53*, 101015. doi:10.1016/j.jeap.2021.101015

Vanani, I. R., & Jalali, S. M. J. (2018). A comparative analysis of emerging scientific themes in business analytics. *Int. J. Business Information Systems, 29*(2), 183-206.

Varga, A. (2018). Novelty and Foreseeing Research Trends: The Case of Astrophysics and Astronomy. *The Astrophysical Journal Supplement Series, 236*(1), 21. doi:10.3847/1538-4365/aab765

Walton, C., & Morris, A. (2013). A bibliometric study of taxonomic botany. *Journal of Documentation, 69*(3), 435-451. doi:10.1108/JD-09-2012-0121

Wang, J.-J., Chen, H., Rogers, D. S., Ellram, L. M., & Grawe., S. J. (2017). A bibliometric analysis of reverse logistics research (1992-2015) and opportunities for future research. *International Journal of Physical Distribution & Logistics Management*. doi:10.1108/IJPDLM-10-2016-0299

Wang, M., & Allen, G. I. (2021). Integrative generalized convex clustering optimization and feature selection for mixed multi-view data. *Journal of Machine Learning Research, 22*, 1-73.

Wang, P., & Yao, Y. (2018). CE3: A three-way clustering method based on mathematical morphology. *Knowledge-Based Systems, 155*, 54-65.

Wang., Q., & Waltman., L. (2016). Large Scale Analysis of the Accuracy of the Journal Classification Systems of Web of Science and Scopus. *Journal of Informetrics, 10*(2), 347–364. doi:10.1016/j.joi.2016.02.003

Wanga, D., Zhang, H., Liu, R., Lv, W., & Wang, D. (2014). t-Test feature selection approach based on term frequency for text categorization. *Pattern Recognition Letters, 45*, 1-10. doi:10.1016/j.patrec.2014.02.013

Wartena, C., & Brussee., R. (2008). Topic detection by clustering keywords. *19th international workshop on database and expert systems applications*, 54-58.

Westergaard, D., Stærfeldt, H.-H., Tønsberg, C., Jensen, L. J., & Brunak, S. (2018). A comprehensive and quantitative comparison of text-mining in 15 million full-text articles versus their corresponding abstracts. *PLoS computational biology, 14*(2), e1005962.

Westgate, M. J., Barton, P. S., Pierson, J. C., & Lindenmayer, D. B. (2015). Text analysis tools for identification of emerging topics and research gaps in conservation science. *Conservation Biology, 29*(6), 1606-1614. doi:10.1111/cobi.12605)

Westgate, M. J., Barton, P. S., Pierson, J. C., & Lindenmayer, D. B. (2019). Text analysis tools for identification of emerging topics and research gaps in conservation science. *Conservation Biology, 29*(6), 1606-1614. doi:10.1111/cobi.12605

Wu, X., Chen, X., Zhan, F. B., & Hong, S. (2015). Global research trends in landslides during 1991–2014: a bibliometric analysis. *Landslides, 12*(6), 1215-1226. doi:10.1007/s10346-015-0624-z

Xue-Li Liu , Gai, S.-S., Shi-Le Zhang, & Wang, P. (2015). An Analysis of Peer-Reviewed Scores and Impact Factors with Different Citation Time Windows: A Case Study of 28 Ophthalmologic Journals *PLOS, 10*(8), e0135583. doi:https://doi.org/10.1371/journal.pone.0135583

Yamamoto, M., & Church, K. W. (2001). Using suffix arrays to compute term frequency and document frequency for all substrings in a corpus. *Computational Linguistics, 27*(1), 1-30. doi:10.1162/089120101300346787

Yan Fang, J. Y., & Wu, B. (2017). Climate change and tourism: a scientometric analysis using CiteSpace. *Journal of Sustainable Tourism*. doi:10.1080/09669582.2017.1329310

Yang, D., Kleissl, J., Gueymard, C. A., Pedro, H. T., & Coimbra, C. F. (2018). History and trends in solar irradiance and PV power forecasting: A preliminary assessment and review using text mining. *Solar Energy, 168*. doi:10.1016/j.solener.2017.11.023

Yıldız, T. (2019). Examining the concept of industry 4.0 studies using text mining and scientific mapping method. *Procedia Computer Science, 158*, 498-507.

Yin, J., Chao, D., Liu, Z., Zhang, W., Yu, X., & Wang, J. (2018). Model-based clustering of short text streams. *In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2634-2642.

Yu, C., Davis, C., & Dijkema, G. P. (2014). Understanding the evolution of industrial symbiosis research: A bibliometric and network analysis (1997–2012). *Journal of Industrial Ecology, 18*(2), 280-293. doi:10.1111/jiec.12073

Yu, D., & Xu, C. (2017). Mapping research on carbon emissions trading: a co-citation analysis. *Renewable and Sustainable Energy Reviews, 74*. doi:10.1016/j.rser.2016.11.144

Yu, Z. (2020). Visualizing Co-citations of Technology Acceptance Models in Education. *Journal of Information Technology Research (JITR), 13*(1), 77-95. doi:10.4018/JITR.2020010106

Yua, D., Xub, Z., Pedrycz, W., & Wang, W. (2017). Information sciences 1968–2016: A retrospective analysis with text mining and bibliometric. *Information Sciences*, 619-634. doi:10.1016/j.ins.2017.08.031

Zainab, A. N., Abrizah, A., & Raj, R. G. (2013). Adding value to scholarly journals through a citation indexing system. *Program: electronic library and information systems, 47*(3), 239-262. doi:10.1108/prog-05-2012-0022

Zhang, L., & Liu, B. (2014). Aspect and entity extraction for opinion mining. *Data mining and knowledge discovery for big data*, 1-40.

Zhang, S., Yang, Z., Xing, X., Gao, Y., Xie, D., & Wong, H.-S. (2017). Generalized pair-counting similarity measures for clustering and cluster ensembles. *IEEE Access, 5*, 16904-16918.

Zhang, X., Wei, F., & Zhou, M. (2019). HIBERT: Document level pre-training of hierarchical bidirectional transformers for document summarization. *Computation and Language*.

Zhang, X., Zhong, Q., Zhang, R., & Zhang, M. (2020). People-Centered Early Warning Systems in China A Bibliometric Analysis of Policy Documents. *International Journal of Disaster Risk Reduction, 51*(101877). doi:10.1016/j.ijdrr.2020.101877

Zhao, F., Shi, B., Liu, R., Zhou, W., Shi, D., & Zhang, J. (2018). Theme trends and knowledge structure on choroidal neovascularization: a quantitative and co-word analysis. *BMC Ophthalmology, 18*(1), 86. doi:10.1186/s12886-018-0752-z

Zhou, L., Amadi, U., & Zhang, D. (2020). Is self-citation biased? An investigation via the lens of citation polarity, density, and location. *Information Systems Frontiers, 22*(1), 77-90. doi:10.1007/s10796-018-9889-9

LIST OF PUBLICATION

Indexed Journal

1. Nadzar, N. M. A. M., Bakri, A., & Ibrahim, R. (2017). A bibliometric mapping of malaysian publication using co-word analysis. *Int. J. Adv. Soft Comput. Appl*, *9*(Papatheodorou, Kapidakis, Sfakakis, & Vassiliou3), 90-113.
2. Nadzar N.M.A.M., Bakri A., Ibrahim R. (2019) The Study of Co-occurrences Index's Keywords for Malaysian Publications. In: Saeed F., Gazem N., Mohammed F., Busalim A. (eds) Recent Trends in Data Science and Soft Computing. IRICT 2018. Advances in Intelligent Systems and Computing, vol 843. Springer, Cham

Non-indexed Journal

1. Bakri, A., Azura, N. M., Nadzar, M., Ibrahim, R., & Tahira, M. (2017). Publication Productivity Pattern of Malaysian Researchers in Scopus from 1995 to 2015. *Journal of Scientometric Research*, *6*(2), 86-101.

Indexed Conference Proceeding

1. N. M. Azura Md Nadzar, D. Aryati Bakri and R. Ibrahim, "Mapping research theme for Malaysia's publication," 2018 IEEE Conference on Big Data and Analytics (ICBDA), Langkawi Island, Malaysia, 2018, pp. 38-43. doi:10.1109/ICBDAA.2018.8629698
   URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8629698&isnumber=8629578
2. N. M. Azura Md Nadzar, A. Bakri and R. Ibrahim, "Potential implementation of subject areas in Malaysia's research assessment Co-word analysis study," 2017 IEEE Conference on e-Learning, e-Management and e-Services (IC3e), Miri, 2017,pp.91-96.doi:10.1109/IC3e.2017.8409244
   URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8409244&isnumber=8409225