Jurnal Teknologi

Full Paper

TREND ANALYSIS IN AGEING AND ICT RESEARCH

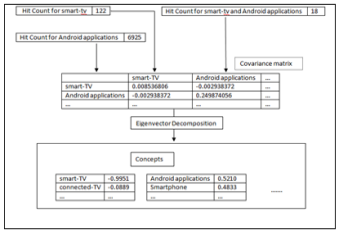
Loo Yow Cherng, Nurul Hashimah Ahamed Hassain Malim*, Manmeet Mahinderjit Singh

School of Computer Sciences, Universiti Sains Malaysia, Penang, Malaysia

Article history
Received
2 February 2015
Received in revised form
24 March 2015
Accepted
1 August 2015

*Corresponding author
nurulhashimah@cs.usm.my

Graphical abstract



Abstract

The population of the elderly is growing rapidly and become a major concern in twentieth-century. The impact of this global ageing phenomenon is significant to the human life in all aspect including social, political and economic of all social class. The implementation of ICT application can help to cope with this phenomenon and improve the quality of life of the elderly. This study aims to perform a bibliometric analysis on the field of ageing and ICT to reveal the trend of research and technologies related in the recent years. Bibliometric information such as keywords and number of publication is extracted from the search result of online publication databases and then further analyzed using computational technique such as growth rate analysis and Latent Semantic Analysis (LSA) to identify the relationship and the growth of the information extracted. The outcome is a set of terms that are grouped and ranked according to their relevance and growth in the recent year. This result can then provide a brief understanding on the trend of ageing and ICT for the researchers that wish to research on this field.

Keywords: ICT, LSA, Bibliometric information, growth rate analysis

Abstrak

Populasi warga tua sedang berkembang dengan pesat dan menjadi perhatian utama pada abad kedua puluh. Kesan fenomena penuaan global ini adalah penting kepada kehidupan manusia dalam semua aspek termasuk sosial, politik dan ekonomi. Aplikasi ICT boleh membantu untuk mengatasi fenomena ini dan meningkatkan kualiti hidup warga tua. Kajian ini bertujuan untuk membuat analisis bibliometrik di bidang penuaan dan ICT untuk mendedahkan trend penyelidikan dan teknologi yang berkaitan dalam tahun-tahun kebelakangan ini. Maklumat Bibliometrik seperti kata kunci dan beberapa penerbitan diekstrak dari hasil carian pangkalan data penerbitan dalam talian dan kemudian dianalisis menggunakan teknik pengkomputeran seperti analisis kadar pertumbuhan dan Analisa Semantik Pendam (LSA) untuk mengenal pasti hubungan dan pertumbuhan maklumat yang diekstrak. Kajian ini berjaya menghasilkan satu set terma yang dikumpulkan dan disusun mengikut kesesuaian dan pertumbuhan mereka dalam tahun baru-baru ini. Keputusan ini boleh memberikan pemahaman yang ringkas mengenai trend penuaan dan ICT bagi penyelidik yang ingin memulakan penyelidikan di bidang ini.

Kata kunci: ICT, LSA, maklumat Bibliometrik, analisis kadar pertumbuhan

1.0 INTRODUCTION

It is important for researchers to understand the state of their field of interest to identify the research area that may play an important role in the future and hence alert to the growth and research of these areas. However, one may face difficulty in identifying the trends and significant, promising areas of research from the ever-increasing size of research. Researchers often faced problems of identifying the value of research that is published. Hence, often the first step of a research is to perform a bibliometric analysis on the interest field. By analyzing bibliometric information such as citation and year of publication, the application of bibliometric analysis can help the researchers to have a brief overview of the trend of their interest field.

This research aim to provide a brief review to the current trend and technologies related to the global ageing phenomena in the field of Information, Communication and Technology (ICT). The global population ageing phenomena has a profound effect to the world and hence has a valuable research value. The use of bibliometric analysis has been performed on various domains to identify the state of their research; however, none of it is focusing on the ageing in ICT context. In this research project, we will perform bibliometric analysis on the bibliographic database and analyze the result to identify terms that are significant and related to the field of ageing in ICT context. Bibliometric information such as "keywords" and "the number of publication" will be extracted from the online publication databases such as ACM Digital Library. The result will then be further analyze to narrow down the related and important terms and hence help researchers to identify the current trend of research in the area of ageing and ICT.

In terms of individual, ageing means the natural process of an individual growing old. On the other hand, population ageing define the process by which older individuals become a proportionally larger share of the total population [1]. In a report published by the United Nation in the year 2012, the ageing population has increased from 205 million in 1950 to 810 million in years 2012 in the world population. In the past 60 years, the population of the elderly has grown more than double in the world population. The growth rate of the population aged 60 or over is faster than the total population for most countries. The elder population is expected to be double into 2 billion in the year 2050 and make up more than 30 percent of the population in 64 countries [2]. This global ageing phenomenon is caused by the decline in fertility and mortality. The effect of population ageing is unprecedented and has profound implication to the human life and hence has a valuable research value. As the population of the aged people becomes the larger portion of the population, the world will need to face the problems and challenges of adapting the elderly to the world of ICT.

The global population ageing phenomena raises challenges in social, health, and economic aspect at the individual, family, and society level. Financial security and health are the major concern in supporting the elderly life in the ageing world [2]. Furthermore, this phenomena is happening faster in developing countries compare to developed countries (as indicated in Figure 1), it is projected that nearly four in five of people aged 60 or over will live in a developing in the year 2050 [2]. Therefore, it is crucial for Malaysia and other developing countries to be prepared in managing the ageing phenomena, as well as supporting the elderly life in different aspect, ensuring that the elderly population is protected and able to participate and contribute to the society meaningfully.

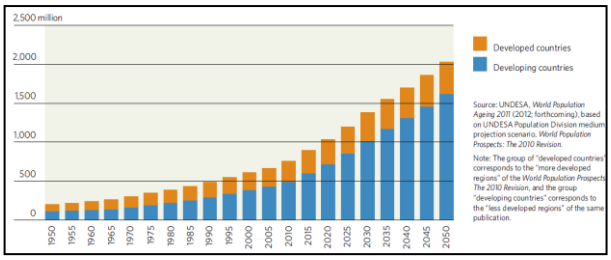


Figure 1 Number of people aged 60 or over; World, developed, and developing country, 1950 – 2050 [2]

However, based on our observation via random searches on articles related to ageing and ICT published by researchers from developing countries, we found that these researches are very focused on handling the ageing population in the healthcare sector [4-5-6]. Whilst, a few other work look into helping the elderly to adapt to the use of ICT such as mobile phone and user interface [7-8]. In developed countries, various researches have been conducted to support the independence life of the elderly people in all fields including health care, security, transport, and ICT, this is important to ensure that aged people can also participate and contribute to the society in a meaningful way as well as providing them essential protection and care that they need [9-10-11-12]. For instance, Smart House system which enhanced and automated houses using ICT has been proven to be useful in assisting elderly independence life and hence increased the quality of life. This have proof that developing countries particularly Malaysia is still far from capable to handle ageing population.

As a stepping stone towards catching up in the preparation of becoming Ageing Nation, researchers from developing countries should start looking into the trends in ICT that are being cultivated by other developed countries. Hence, identifying fast growing technology in ICT is crucial for researchers to understand what are the important areas related to this field so that they are aware to the growth of these areas which may play an important role in the

future and hence worth researching. Unfortunately, researchers often faced difficulties in identifying these related fields that are important to their field of interest. With the large amount of information that are available and keep updating, researchers may find themselves spending days just to analyzing this information to find an area worth to research, this is time consuming and not efficient for the researchers. Besides, one may not possess all the knowledge related to their field of interest and hence may missed certain important areas related to their field of interest as they doesn't understand the value of the area as seen by other researchers. Although the application of bibliometric analysis can help researchers to understand the state of a research field, none of the bibliometric analysis has been performed on the research area of ageing phenomena with ICT.

One of the methods that can be used to identify the trend of a field is through the use of bibliometric analysis. Bibliometrics, the application of mathematics and statistical methods to books and other media of communication [3] can be used to analyze research publications to evaluate the impact of the article to a particular field. Such quantitative approach included analyzing the bibliographic references, publication, title, author, and keywords. Information such as keywords can represent the field that is relevant to the field of interest; while information such as the number of publication can be analyze to identify the growth rate of the field.

The focus of this study is to perform a bibliometric analysis on the field of Ageing and ICT by using only the bibliometric information that can be easily obtain from the online publication databases, and hence identify the trend of research area in the recent years which in turn provide a brief introduction of the current research state for the researcher that are interested on the research of Ageing and ICT.

2.0 BACKGROUND

Ageing is the natural process of an individual growing older; while population ageing is the process where older individuals become a proportionally larger share of the total population. Population ageing is one of the most distinctive demographic events of the twentieth century [1] and a report published by the United Nation shows that the population of the older people will become double in the year 2050, reaching 2 billion; besides, there will be 64 countries that has more than 30 percent of population aged 60 or above [2]. This global population ageing phenomena is unprecedented, pervasive, long-lasting and has a deep affection to human life. This global population ageing process is caused by the decline in fertility and mortality and has a significant impact on economic, political, and social of both the individual and society level [1-2]. Table 1 indicates the causes of poulation ageing which are the decreased of fertility rate and the increased of life expectancy in two time frame i.e. 1950 to 1955 and 2010 to 2015.

Table 1 Cause of population ageing

Year	1950 - 1955	2010 - 2015
Fertility rate	4 children per women	2.5 children per women
Life Expectancy (average)	45 – 50 years	68 – 78 years

2.1 The Use of ICT in Facing the Ageing Society

Concern are growing about providing elderly a better quality of life in different aspect including financial, security, health, as well as participation of elderly in the society. The use of Information and Communication Technologies (ICT) is believed to be able to improve the quality of life for the aged people. To help the elderly to adapt to the use of ICT, research on motivation theory is preformed to encourage the elderly to learnt to use ICT [15]. In Japan, interactive robot is also used to support the elder people's independent life by providing mental support to the lonely elderly through communication with the robot, and hence improving the smart home environment [16]. Similarly, to improve the independent life of elderly in a smart home environment, ICT and networking can be used to communicate with friends and family through social networking, and hence improving the life of individual elderly which are suffering from social and physical isolation life style at home that might lead to cognitive impairment [17]. The design and the accessibility of the web site is also researched to identify the what are the suitable web design that can encourage the use of ICT among the elderly [7-18]. Spain that also facing the global population ageing phenomena research on how to bridge the digital divide for elderly people [19]. Devices such as HealthPa and teleWEAR improve the health of the elderly by monitoring the health condition of the elderly and alert the family member or the friends of the elderly [9-20].

2.2 Bibliometric Analysis

Bibliometric analysis is defined as the application of mathematics and statistical methods to books and other media of communication [3]. The bibliometric information that can be used from a publication includes authors, references, citation from other publication, co-citation with other publication, keywords, number of publication, and the publication year. Bibliometric analysis can be used as an evaluation tools to evaluate the impact of a publication to the research area as well as a tracking tools to monitor the level of activity in a research field [21]. By using bibliometric analysis to analyze the

large amount of metadata of past publications, researchers can then understand the patterns of the researches in the field of interest and even forecast the important technology in the future, it is believe that an emerging technology will showed a positive growth for the number of publications in the recent years while a mature technology will have reduced number of publications [22].

Citation analysis is performed on a paper publish by [23] to identify the current structure of their research. Citation data of publications related to their field is collected to form a citation network. Clustering is then performed on the citation network to organize the publications into different cluster, and hence reveal the structure and relationship between them. Ding [24] has performed co-word analysis by using paper's keywords that occur together within the paper. Terms are extracted not only from the keywords defined in the databases but also from the publication's title and abstract. The number of times a pair of term appears in the same document is recorded to form correlation matrix using Pearson's correlation coefficient. Keyword clustering is then performed for the data mapping of the keyword's relationship. In a paper published by [25], publication count is used investigate the growth rate of the interest field with respect to time, country, organization, and subject area. The keywords of the papers are extracted and then the concept of keyword clustering is applied to the frequently occurred keywords. By considering the co-occurrence of the keywords in a document, the keywords are grouped into different cluster to determine the relationship between keywords.

Woon [26] claim that the use of bibliometric analysis techniques such as co-citation analysis could be time consuming and expensive for massive data set because the bibliometric information such as abstracts, author names, or even the full text of the articles will need to be downloaded and processed locally, such as what is done by Janssens [27] and Kajikawa [23]. In the case of Kajikawa, bibliometric records of 79,705 articles were collected and analyzed, this amount of data could be very hard to be collected and analyze. Besides, methods such as co-citation and inter-citation analysis sometimes may result in a map that contains a huge number of nodes, which could be difficult to process and interpret. The author claims that bibliometric analysis is hard to be perform if one doesn't provided will the full access to the bibliometric documents, and hence proposed a method that can be used without the need to explore the full documents, which is by just simply analyzing via s simple search interface. The authors choose to explore the relationship of research areas by investigating co-occurrences of keywords in the publication by utilizing the hit count from the online search engine, using methods modified from Normalized Google Distance. The concept is based on the intuition that the appearances of a particular pair of keywords in a large number of publications indicate that these pair of keywords is closely related to each other. The result is then further organized using visualization or clustering process such as k-means algorithm to group keywords into closely related clusters.

2.3 Latent Semantic Analysis (LSA)

Ziegler [21] claims that the technique of keyword clustering is not effective enough to determine relationship between keywords. This is because in these technique keywords is grouped into disjoint clusters, meaning that every keyword can only appear in exactly one cluster. This is inaccurate for homonyms terms where a single terms can have multiple meaning. Hence, Ziegler proposed a technique modified from Latent Semantic Analysis (LSA), a technique that can identify relationship among keywords without restricting the keywords to appear in only one cluster. The application of LSA also explores the co-occurrence of keywords to assign those keywords that is highly related to each other into the same group, however, LSA produces groups which is called "concepts", where a keyword is possible to appear in multiple concepts.

Given a set of terms, LSA analyze the co-occurrences of the terms and then organize them into groups called "concepts". Inside each of these concepts, terms are given a value that represents the weight of the term in this particular concept. Terms that are related to each other will be given a higher weighting while terms that are irrelevant to the concept will be given negligible weighting. Hence, by analyzing the weight of terms inside these concepts, we can then identify their relationship. The advantage of LSA is that it can also solve the problem of homonym terms, which is unsolvable in other method that identity term's relationships such as the clustering method. Unlike clustering which restrict each of the terms to be groups into one and only one cluster, LSA allowed a term to be present in more than one concept. In LSA, a homonym term maybe given a high weighting in multiple concepts, indicating that the term have multiple meanings.

The application of LSA is based on a general statistical technique known as Principle Component Analysis (PCA), which explores the relationship of the data by analyzing the variance of the input data, while LSA is a straight forward application of PCA that explore the term frequency in the document that formed term-document matrix by using terms as the rows and documents as the columns of the matrix, while the values Ai,j in the matrix represent the frequency of term i's occurrence in the document .j Singular Value Decomposition (SVD) is then performed on the term-document matrix to produce groups that is called "concept". Besides, SVD also produces a singular values for each of these concepts, which can be used to determine which concepts are most representative of data. By performing LSA, the information initially collected can then be reduced into a smaller set of important and uncorrelated concepts, For a more detailed

explanation of the concept of PCA and LSA, reader can refer to the published of [21-28-29].

Ziegler [21] proposed a novel approaches modified from LSA, which allowed the process of determining relationship between terms collected to be simplified. In the original LSA, term-document matrix is formed by searching the terms occurrences in the documents, this will required exploring the whole document to find the terms frequency. Ziegler proposed an idea that can simplify the process of forming the term-document matrix. By assuming that a term will occurs 0 or 1 times in the document, the algorithm of LSA is simplifies to using only the hit count returned from the search result to form a term covariance matrix. Instead of exploring the whole document to find the term occurrence, we only need to extract the hit count for the terms when using the terms as a search term in the online publication databases.

Many researchers has used methods adapt from bibliometric analysis to forecast emerging technologies in their field of interest [21-23-24-25-26]. Unfortunately, none of the research is focusing on forecasting technologies related to ageing in ICT context.

2.3 Growth Rate Analysis

To determine the rate at which a technology is growing, trends of the number of publications over time can be analyze as a more popular technology will have more number of publications in the year. As mention by Ziegler [21], there is several ways in how to define the term's growth rate by using the hit count over time. The first metric that can be used to determine growth rate is to simply take the numerical difference between the term's final year hit count and initial year hit count. This straightforward metric can represent how the term grown over the entire period of time however it has two downsides. One is that the hit count in between the first and final year is ignored. Another more important downsides is that this metric will favored term that have high hit count in initial years, even though there might be terms that have low initial year hit count but a higher growth than other terms that have high initial year. The second downsides of the difference metric mentioned can be solved by using another metric, which is the ratio metric that take the ratio of the final years hit count over initial year hit count. This ratio is equivalent to the logarithm equation

$$\log\frac{X}{Y} = \log X - \log Y \quad (1)$$

Where X is the final year hit count and Y is the initial year hit count. From this logarithms equation, we know that log X decreases with increasing X, and hence this metric favored terms that have low initial hit count. The downside of this ratio metric is that the growth rate will be undefined for terms that have zero hit count in the initial year. The third metric that can be used consider all hit count for the range of years, this metric fit the exponential curve to the hit count time series and take the exponential parameter as the growth rate. For example, if is the best fit exponential curve, then r is the growth rate. The fourth metric also consider all hit count over the period of time, which it calculate the average year of publication. This metric takes the formula of , where it sum the product of all years yi and the hit count from that year hi to divide by the sum of hit count for all years. This metric has produce time frame that tells which years the terms was the most prevalence.

3.0 METHODOLOGY & IMPLEMENTATION

This research involved quantitative research where data will be extracted and collected from the online publication databases and then further analyze using computational and algorithms techniques to produce a list of terms that researchers may interested on the issues of ageing in the ICT context. The research will be performed using the methodology that is modified from the bibliometric analysis method proposed by Ziegler [21]. The methodology of this research is indicated in Figure 2.

The first step of this research is to identify what are the fields or technologies that are related to our interest field, which is ageing and ICT. This can be done by using information that is available from the online publication databases. By using ageing and ICT as the search term to query the online publication databases, the databases will return search results that are related to our interest field. These search results will contain keywords that are defined by the authors. Assuming that these keywords defined by the authors are terms that is potentially related and can represents fields or technologies that are related to our interest field, these keywords will be extracted automatically using available web data extraction software. These keywords will then be further analyzed to narrow down the results into list of fast growing and highly relevant terms in the field of ageing and ICT.

Since we are only interested with the technologies or fields that are currently emerging and popular, growth rate analysis will need to be performed on the keywords that are extracted previously. This can be done by analyzing the trend of the keyword's hit count over a time span. Assuming that a popular technologies or fields will have more research publications, the hit count of a keyword (number of publication) will be used to calculate its growth rate. For a given range of years, the keywords will be used as the search term to the online publication databases and the hit count will be extracted for every year. This extraction is performed on each of the keywords collected previously and hence will provide us a time series for the keyword's hit count for every keyword collected. Growth rate of the keywords can then be calculated by using the

information collected. As a result, we then obtain a list of keywords with their respective growth rate and hence we can identify keywords that have the highest growth rate, which can represent the growing and important technologies or fields that is related to ageing and ICT, in the given time span.

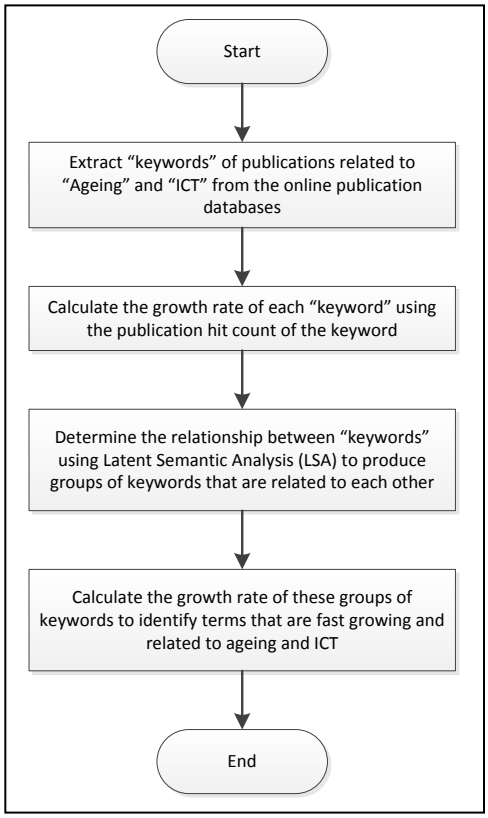


Figure 2 Research methodology

The next step in this study is to group the keywords into related groups by using the co-occurrences of the keywords in the publication. If two keywords frequently appear together in different documents, we then can conclude that these two keywords are related to each other and hence should be group together. We use a technique proposed by Ziegler [21] that is modified from the mathematical technique known as Latent Semantic Analysis (LSA) to identify the relationship between keywords by analyzing the co-occurrences of each of the keywords and hence organize them into related groups which is called "concept".

The result from the LSA is a list of concepts, which inside each concept are terms that are relevant to each other. Hence, each of these concepts can then represent a field or technologies that are related to our initial search term, which is ageing and ICT. This result can be used by the researcher to identify what are the current research area that is related to the domain ageing and ICT. Furthermore, growth rate analysis will again be performed onto these concepts, arranging these concepts according to their growth rate and hence help researchers to identify concepts that are fast growing and popular in the recent years.

The result of this research after performing all the steps above is a list of terms in groups that represents different concepts or fields, which is fast growing and related to the domain field ageing and ICT. This result can then be used by researchers that are interested to the ageing phenomena in related to ICT to understand what is the current state of research and hence ease the effort in determining the trends and challenges in facing the global ageing phenomena in the fields of ICT.

3.1 Keyword Extraction

In choosing the online publication databases to be used for the data extraction, two important criteria of the databases need to be considered. First, the online publication databases must contain a large number of publications that is related to our field of interest. Second, the search result of the databases must be displayed in a way that allows the data to be extracted automatically using the web data extraction tools. Hence, the online publication database ACM Digital Library has been chosen as it fulfills the two criteria mention above.

We are interested to the fields that are related to ageing and ICT, and hence a search has been performed on a few databases by using ageing and ICT as the search term. As a result, IEEE Xplore Digital Library and Scopus returns 25 and 243 search result respectively, while ACM Digital Library returns a search result of 475 publications. This shows that the ACM Digital Library contain more publication that is related to our field of interest. Besides, the search results of ACM Digital Library also displayed in a way that allows the data such as keywords and hit count to be extracted automatically. Large scholar web search engine such as Google Scholar are not suitable to be used because the search result is displayed as links to other websites, making the extraction of the keywords extremely difficult to be perform automatically using the web extraction tools because of the result that are not standardize.

There are plenty of automated web scrapping tools available, which vary in cost and features [31], such as Visual Web Ripper, Web Content Extractor, Mozenda, and many more. One of the most simple web data extraction tools for inexperience users is the Visual Web Ripper [30], which provide a very user friendly visual project editor, and easy to learnt features for first time users of web extraction tools. Visual Web Ripper also provide extra functionalities for users which need to perform tricky extraction on more difficult cases, by providing the users the free to use techniques such as XPath, Regex, and programming scripts. By using Visual Web Ripper, users can repeatedly submit forms automatically for all different possible input values, this feature is suitable to be used to extract keywords and hit count repeatedly for this research. Besides, the official website of Visual Web Ripper also provide step by

step tutorial on how to use the feature of the tool. The Figure 3 shows the screenshot of the Visual Web Ripper.

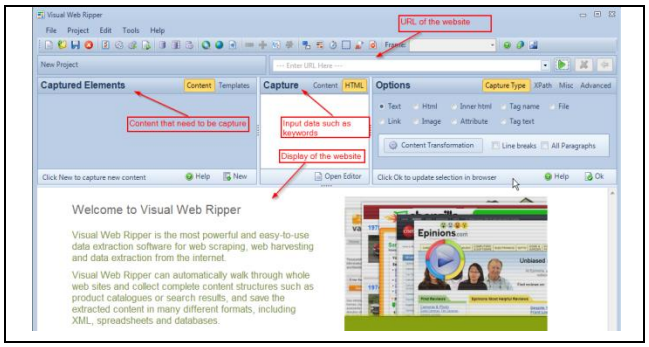


Figure 3 Visual Web Ripper tool screenshot [30]

By using the web data extraction tool Visual Web Ripper, keywords of an article can be extracted automatically from the ACM Digital Library. The screenshots below shows the search result by using ageing and ICT as the search term in the ACM Digital Library. The result is displayed in a way that the keywords of each publication is showed in the abstract and hence can be extracted easily using the Visual Web Ripper.

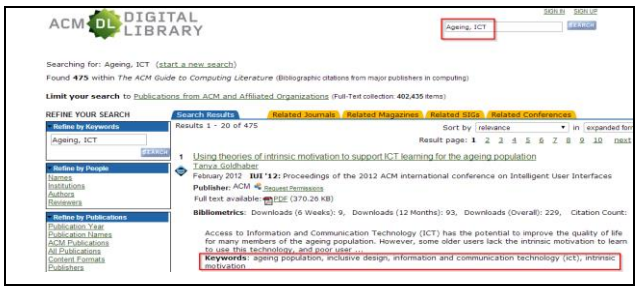


Figure 4 Search result page of ACM Digital Library by using the ageing, ICT as the search term

3.1 Growth Rate Analysis of Keywords

For each of the keywords extracted as mentioned in section 3.1, growth rate analysis will need to be performed on these keywords to identify which keywords is fast growing and prevalent in the recent years. This is done by analyzing the number of publications over time for each of the keywords. Researchers often use the research publications as channel to publish their research findings, and hence a vast number of publications related to a specific topic can indicate that this topic is currently popular and is a focus in the research field. Therefore, we assume that the publication count can be a useful data to be used in the growth rate analysis. Given a range of years, each of the keywords extracted previously will be used as the search term in the ACM Digital Library to extract its publications counts for every year. The hit count (publication counts) can be easily extracted as it is displayed in the search result page.

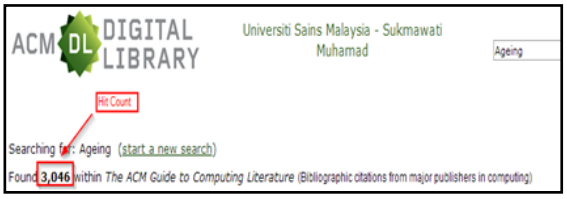


Figure 5 Hit Count for Ageing return by the ACM Digital Library

There are many methods that can be used to represent the growth rate of the terms, such as using the difference between the final year's hit count and the initial year's hit count, ratio of final year's hit count over the first year's hit count, and the average year's hit count, each of this methods has their own pros and cons.[21] In this research, the ratio of the final year hit count over the first year hit count will be used to represent the growth rate.

$$\text{Growth Rate} = \frac{\text{Final Year's Hit Count}}{\text{Initial Year's Hit Count}} \quad (2)$$

This method will generate high growth rate for terms that has low hit count in the beginning but high hit count in the recent year, allowing us to identify terms that are fast growing.

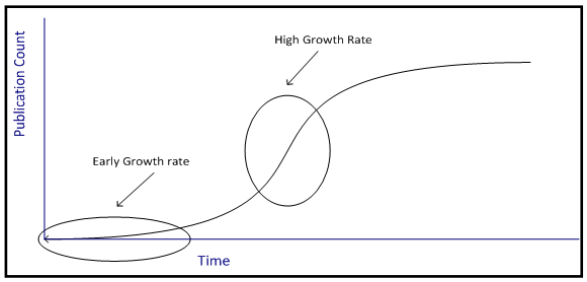


Figure 6 Growth rate analysis identify terms that have high hit count in recent years but low hit count in the initial year

3.1 Implementation of LSA

To identify the important research areas that are fast growing in related to ageing and ICT, it is inaccurate by simply taking the keywords that are extracted from the previous steps. This is because terms extracted from the part 4.1 might be synonym or homonym.

In the case of synonym, different authors might define different terms for the same technology, and hence giving us lower hit count if we only consider hit count for a single term. For example, the terms "smart home" and "smart house" are representing the same or similar technology, a search of "smart home" will return a hit count of 17,369 while a search of "smart

house" will return a hit count of 7,881. On the other hand, a search of "smart home" or "smart house" will return a hit count of 21,041 instead. This showed that these terms should be group together to identify hit count that are more accurate then by just considering a single term's hit count. In the case of homonym, the hit count of a single term might be higher than the actual number that is representing what we are interested to the field of ageing and ICT. A term that is homonym can have multiple meaning, giving us search result for publications that are irrelevant to ageing and ICT. For instance, according to the Oxford Dictionaries, the keyword "cloud" can represent the technology "cloud computing" or the "condensed water vapor floating in the atmosphere" . Since we are only interested to the fields that are related to ageing and ICT, taking the hit count for the term "cloud" will give us inaccurate growth rate for the articles that are unrelated to our interest.

To overcome the problems caused by synonym and homonym, we need to identify the relationship between terms extracted and analyze the growth rate of these terms in groups. In this research, we are using the technique proposed by [21], for identifying the relationships of terms by modifying the mathematical technique known as Latent Semantic Analysis (LSA). This technique explores the relationship of the terms by analyzing the co-occurrences of the terms extracted.

Given a set of terms as the input to the LSA, each of the terms will present in every concepts that is produced, with their weight in the respective concepts, which represent the relevance of that term in that particular concept. Hence, n terms of input to the LSA will always produce n concepts. Furthermore, the concepts as a whole is also weighted according to their prevalence, therefore, concepts with low weights can be ignore as well as the individual terms that have low weight in the concept.

In this research, we are using a modified LSA technique that is proposed by Ziegler [21], which simplifies the works in forming the term-document matrix that is used as the input to the LSA. In the original LSA algorithm, the term-document matrix is a matrix that is formed by using the term occurrences in the document. To determine the value for each of the entries in the matrix, we will need to explore the whole document text to calculate the term occurrence. This will required a full access to the documents text, which is not feasible in this research since we are only accessing the bibliometric information available in the online publication databases. In this modified technique, only the hit count (number of publications) of the terms will be used to from the matrix. These hit counts are then used to form a covariance matrix by using the covariance expression below.

$$\mathrm{cov}(i,j) = \frac{1}{n}\left(h_{i,j} - \frac{1}{n}\left(h_i h_j\right)\right) \quad (3)$$

Where cov(i,j) is the entries of the matrix at row i and column j, hi and hj is the hit count of the term i and term j respectively, hi,j is the hit count for searching the both terms together, and n is the number of documents being searched.

Eigenvector decomposition is then performed on this covariance matrix. Given an input of n terms, an n x n covariance matrix is formed using the covariance expression, and hence n eigenvector will be formed from the eigenvector decomposition, each with their own eigenvalues. These eigenvectors are the groups of terms that we called "concept" in the application of LSA, while eigenvalues is the value that represent the prevalence of that particular concept, which we mentioned above.

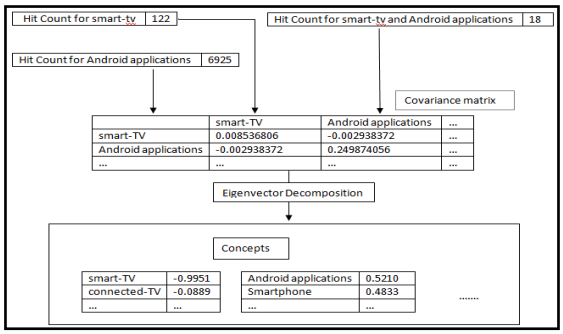


Figure 7 Term's Hit Count is used to produce covariance matrix which form concepts through eigenvector decomposition

Each of the terms in the concepts will be given a value that represent its weight in that particular concept, we then rearrange the terms by sorting in a descending order of the absolute weight. Terms with the highest value indicating that it is representing that particular concept and those terms on the top of the concepts are highly related to each other. For example, in Figure 6, smart-tv and connected-tv are the terms with the highest absolute weight in the concept and hence they are representing that concept, and these two terms are highly related to each other. Note that absolute weight is used because a term with a high negative weight is as important as a term with a high positive weight in representing the term's relationship with each other.

3.1 Growth Rate Analysis for Concepts

The result of the application of LSA is that the terms that we extracted previously is now organized into groups that are called concepts, and this result can be further analyze by calculating the growth rate to determine the groups there are fast growing. Note that we already performed growth rate analysis on the keywords that is extracted; this is actually a first approximation of the fast growing technologies or fields, which we use to reduce the set of input terms into the LSA. The implementation and result of the LSA become hard to interpret if the number of input data is great, and hence we used growth rate analysis to reduce the set of data to be use as the input of LSA. Growth rate analysis is then again performed on the result of the LSA to identify fields or technologies that are prevalence in related to ageing and ICT.

In this section, we perform the growth rate analysis by first calculating the hit count for the whole concept. The hit count for a concept is calculated by summing the weight of the terms inside the concept with their respective term's hit count. For example, if a concept contains the term smart-tv with weight 0.9951 and the term connected-tv with weight 0.0889, and their respective hit count is 122 and 45, then the hit count for this concept is given by $0.9951 * 122 + 0.0889 * 45$. After we calculated the hit count of each concept for a specific time span, we than can calculate the growth rate by using the ratio of the final year' hit count over initial year's hit count. As a result, the concepts are now ranked based on their growth rate and we can hence identify fields or technologies that is fast growing in related to ageing and ICT. The summary of the steps taken and outcome from the initial search term to related field being identified is shown in Figure 8.

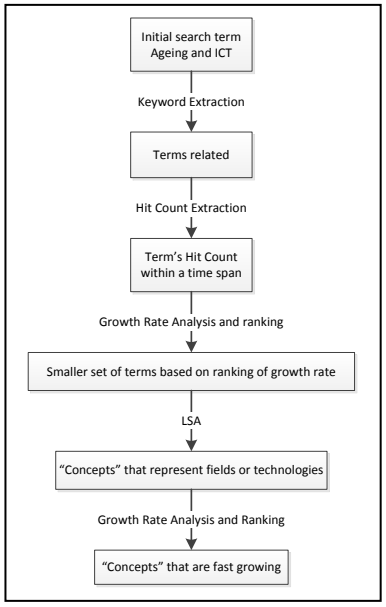


Figure 8 The process and result in each step

4.0 RESULTS AND DISCUSSION

In this section, the result of our research by using ageing and ICT as the initial search term will be presented in accordance to the phases mentioned in section 3.

4.1 Keywords Extraction

For the search terms that is used for keyword extraction, synonyms of ageing and ICT is also used in order to collect result that are more accurate. The search terms ageing, aging, and elderly is used interchanged together with the search terms ICT and computer, which is also used interchangeably. For example, "ageing, ICT" is first used as the search term to extract the keywords from the results, and then the search term is replaced by "ageing, computer" in the second search. Therefore, 6 searches are performed to extract the keywords from the search result. Each of these searches returns approximately 500 to 10,000 search result and each of the search result contain at least 4 keywords in the abstract. Table 2 shows the number of publications returned from each pair of the search terms and also the unique keywords extracted from all of the publications returned.

Table 2 Search result and keywords count by using different search terms

Search Terms	Search Result (Num. of article)	Number of Keywords extracted from publications
Ageing, ICT	488	1,441
Aging, ICT	1,133	2,777
Elderly, ICT	1,212	2,886
Ageing, Computer	2,605	6,791
Aging, Computer	9,296	19,129
Elderly, Computer	8,984	16,454
Total num. of unique Keywords	33,554	

If we extract every keyword from every search result by using the 6 search terms mentioned above, approximately 33,000 unique keywords will be collected. However, this number of keywords is too huge to be analyzed in order to identify the growth rate. In the step of identifying the relationship between keywords, the difficulty of using LSA concepts increased as the number of terms increased. Furthermore, assuming that the search result is displayed by the online publication databases in a manner that the most relevant article is displayed first, then for search term that return a large search result might return keywords that have low relevance to our initial search term for articles that are ordered at behind. Hence, instead of using all of keywords that can be extracted from the set of search terms mention above, only a portion of

articles returned will be consider for the keywords extraction. We aim to determine 1,000 to 2,000 keywords to be the input for the growth rate analysis as this number is appropriate for the hit count extraction in the next step. A number greater than 2,000 will be time consuming for the hit count extraction. Since the keywords extraction of 100 articles will approximately returns 300 keywords for each of the search term mention above, we decided to consider only keywords extracted from the first 100 search result from each search term. As a result, a total of 1,483 unique keywords are extracted for further analyzing as shown in Table 3.

Table 3 Number of keywords extracted for the first 100 search result returned using different search term

Search Term	Number of unique keywords extracted from the 100 publications
Ageing, ICT	324
Aging, ICT	324
Elderly, ICT	405
Ageing, Computer	356
Aging, Computer	280
Elderly, Computer	394
Number of unique keywords	1,483

A small set of 50 keywords extracted is displayed in the Table 4.

Table 4 Partial keywords extracted from the ACM Digital Library

smart-TV	87.19.xr
Android applications	87.57.R
Android	accessibleTV
adaptive smartphone	commvault galaxy
smartphone	cross-age FER
Aging-related bugs	gamification
Web accessibility	Genetic algorithm (GA) and Ziegler-Nichols (Z-N) method
einclusion	GEROM
connected-TV	gerontopsychology
tablet computers	Guardian ad Litem
Touchscreen tablet	homeML
Calcium-dependent modulation	ICT in language teaching
mosaic-type work	Later midlife
apps	L-type Ca2+ channels
Ageism	M-GTA
AAL platform evaluation	persona re-usage
senior cloud	PRIMER-ICT
web accessibility guideline	stereopicture
cloud computing	swallowing apnoea
Acessibility	Symbiotic dyadic e-health networks
wrist-worn	technology induced movement(TIM)
gaze pattern lock	user centred desing
Human-product interaction	Utilisation-importance matrix
mental wellbeing	Venn-network
comprehensive roadmapping methodology	Vestibular autorotation test

From the Table 4 we can see some of the promising terms such as smart-TV and adaptive smart phone, however, it also contain noise terms which is too general or too specialized such as 87.19.xr, Utilisation-importance matrix, and Vestibular autorotation test. Hence, the result from the first extraction need to be further analyzes using growth rate analysis and Latent Semantic Analysis.

4.2 Growth Rate Analysis of Keywords

Since we are using the ratio of the final year's hit count over initial year's hit count as the growth rate of terms, any terms with the value 0 as the initial year's hit count will causes the growth rate to be undefined and hence excluded from the ranking. Therefore, the time span that we choose must be not too long so that technologies that are just started to become known is not excluded from the ranking. Hence, we have chosen a range of 5 years for the hit count extraction of keywords, with year 2009 as the initial year and year 2013 as the final year. The year 2014 is not chosen as the final year because the hit count for the current year will be low. The Table 5 shows the part of the term's hit count.

Table 5 Terms hit count for year 2009 and year 2013

Keywords	2009 Hit Count	2013 Hit Count
3d human activity recognition and monitoring	210	270
3D virtual stores	744	841
42.30.Sy	2	1
6LoWPAN	66	102
87.19.xr	0	1
87.57.R	0	1
AAL	120	187
AAL platform	37	82
AAL platform evaluation	17	57
AAM	128	114
accelerometer	932	1273
acceptance	4,509	4386
Access Control	17,807	14798
access to the web	11,268	10487
accessibility	2,305	2196
Accessibility barriers	367	457
accessibility for elderly people	163	186
accessibility guidelines	525	566
accessible contact manager	301	335
accessibleTV	0	0

Growth rate analysis is then performed by using the hit count extracted, with the algorithm in equation 2.

Table 6 shows the fifty highest ranked terms with their respective growth rate.

Table 6 The top 50 terms with their growth rate

Terms	Growth Rate	Terms	Growth Rate
smart-TV	33	cyber physical systems	2.487
Android applications	10.39	Co-creation	2.471
adaptive smartphone	8.22	restful web service	2.453
smartphone	8.21	postgraduate health care education	2.448
Aging-related bugs	8	tree-augmented naïve Bayes	2.4
Web 2.0	5.91	aging friendly ICT technology	2.357
einclusion	5.69	Craniofacial growth	2.33
connected-TV	5	ICT4D	2.32
tablet computers	4.04	smart social platform	2.319
Touchscreen tablet	4.03	active societal engagement	2.264
Calcium-dependent modulation	4	AAL platform	2.216
mosaic-type work	4	latent prints	2.202
apps	3.82	TV-WEB	2.2
Ageism	3.5	LBP	2.199
AAL platform evaluation	3.35	volunteer-based organizations	2.167
senior cloud	3.20	Green ICT	2.153
Web accessibility	3.18	NFC technology	2.149
cloud computing	3.01	non-digital and digital artefact	2.14
Acessibility	3	Parental supervision	2.09
wrist-worn	2.95	CWL sensor	2
gaze pattern lock	2.89	Loneliness	2
Human-product interaction	2.83	multimedia hieroglyph	2
mental wellbeing	2.77	senior-friendly design	2
comprehensive roadmapping methodology	2.67	sociophysical interactions	2
persuasive mobile interfaces	2.56	Multi-touch interaction	1.993

The growth rate analysis shows some of the focusing fields of research in recent years, which is related to ageing and ICT. These fields are mostly related to the application of smart devices, including terms such as smart-TV, smart phone, Android applications, and tablet computers.

4.3 LSA Application

The application of LSA relies on algorithm outlined by equation 3. This algorithm requires the extraction of the joint hit count for every pair of terms used as the input of LSA. Considering for the time needed to extract the hit count of terms, this algorithm is only suitable to be used for a small set of input terms. An input of 1,000 terms will require us to extract the joint hit count of 1000 *1000 =1,000,000 terms. Furthermore, the results of LSA become difficult to interpret as the number of terms increases. Therefore, the result from the growth rate analysis will be used to reduce the size of the input terms in LSA. Approximately 50 terms as the input will be appropriate for considering the time of hit count extraction and the interpretation of the LSA results. Hence, the 50 highest ranked terms showed in the Table 6 will be used as the input for the LSA. Hit count and joint hit count is then extracted for these 50 terms using the range of years from year 2008 to year 2013. The total documents return from the initial search terms is used as the value of n. As a result, a 50 * 50 covariance matrix can be formed and hence 50 concepts are generated from by performing eigenvector decomposition.

Recall that a weight value is given to every concept produced in the process, which represents the concept's prevalence in the fields. This value is called eigenvalues, and our result from this research shows that after the 29th concepts, the remaining concepts shows eigenvalues of 0.00, indicating that these concepts are not representative of the data. Hence, these concepts are not presented in this research. Furthermore, for the conservation of space, only the highest weighted terms in each concept is displayed. The first six concepts with the highest five weight terms is displayed here in the Table 7. Note that from the process of forming the covariance matrix, some matrix might be given negative values, however when considering the correlation between terms, only the magnitude of the weighting is considered in judging the prevalence of the concepts or the terms, this is because a strong negative correlation between terms is as important as a strong positive correlation, where the magnitude will represents the strength of the linear relation.

The result from this stage shows that the terms are now organize into groups based on their relevance to each other. For example, Android application, smart phone and tablet computers are the highest weighted terms in the Concept 4, revealing that Concept 4 is representing smart devices that are related to Android. This result is useful in a way that helping the researchers to identify which technologies is related to each other. However, since we are interested to the technologies that are fast growing in the recent years, growth rate analysis is

performed on the following section on these concepts to rank them according to their growth.

Table 7 The first six concepts with their five highest weighted terms

Concept 1		**Concept 2**	
EigenValues	-2509.07	EigenValues	3.292
wrist-worn	-0.997	Web 2.0	0.9397
Web 2.0	-0.0626	cloud computing	0.297
cloud computing	-0.0403	Web accessibility	0.0815
Web accessibility	-0.0112	wrist-worn	-0.0749
tablet computers	-0.0107	Smart social platform	0.0581

Concept 3		**Concept 4**	
EigenValues	1.9341	EigenValues	0.8861
cloud computing	0.9376	Android applications	0.5210
Web 2.0	-0.31694	smartphone	0.4833
senior cloud	0.0657	tablet computers	0.4530
Android application	0.05842	apps	0.3572
apps	0.0553	smart social platform	0.2757

Concept 5		**Concept6**	
EigenValues	0.54167	EigenValues	0.4491
Web accessibility	0.9605	tablet computers	-0.85531
Android application	-0.16179	apps	0.3353
smartphone	-0.1422	Android applications	0.25873
apps	-0.1009	smartphone	0.1786
smart social platform	0.0823	Multi-touch interaction	-0.1286

4.4 Growth Rate Analysis of Concepts

Growth rate analysis is performed on these 29 concepts to identify concepts that are prevalence in recent years. Hit count of the concept is first calculated by using the terms hit count and their weight in the concepts.The hit count for these concepts is displayed in the Table 8.

Table 8 Concept Hit Count for year 2009 and year 2013

	Hit Count 2009	Hit Count 2013
Concept1	11549.63	1000.542
Concept2	171985.9	10521.64
Concept3	60131.2	10116.5
Concept4	17207.74	6884.654
Concept5	13664.24	3127.6
Concept6	3481.443	4559.466
Concept7	3504.402	3436.523
Concept8	2007.21	4351.997
Concept9	2943.997	4301.179
Concept10	577.2783	3678.937
Concept11	1832.611	1077.646
Concept12	633.2721	1193.756
Concept13	772.9818	1335.533
Concept14	902.1366	409.7885
Concept15	1980.482	673.1969
Concept16	1323.579	454.9357
Concept17	304.801	1214.544
Concept18	411.742	725.8364
Concept19	639.2458	582.1213
Concept20	713.3365	423.5407
Concept21	305.7013	677.4165
Concept22	402.8865	388.1526
Concept23	203.8311	238.2891
Concept24	322.8759	290.3546
Concept25	116.4384	213.54
Concept26	110.9087	204.1246
Concept27	69.20176	156.2789
Concept28	239.8525	168.3515
Concept29	56.31373	110.4274

These hit counts are then used for the growth rate analysis that takes the ratio of final year hit count over initial year hit count. The concepts are then ranked and displayed in according to their rank in the following Table 9.

Table 9 Concepts ranking based on the growth rate

	Growth Rate		Growth Rate
Concept10	6.37290102	Concept22	0.96342921
Concept17	3.98471109	Concept19	0.91063769
Concept27	2.25830762	Concept24	0.89927634
Concept21	2.21594246	Concept28	0.70189595
Concept8	2.16818187	Concept20	0.59374597
Concept29	1.9609327	Concept11	0.58803853
Concept12	1.88506043	Concept14	0.45424217
Concept26	1.84047471	Concept4	0.40009054
Concept25	1.83393178	Concept16	0.34371632
Concept18	1.76284283	Concept15	0.33991574
Concept13	1.72776744	Concept5	0.22888951
Concept9	1.46099989	Concept3	0.16824042
Concept6	1.30964815	Concept1	0.08662982
Concept23	1.16905186	Concept2	0.06117732
Concept7	0.98063034		

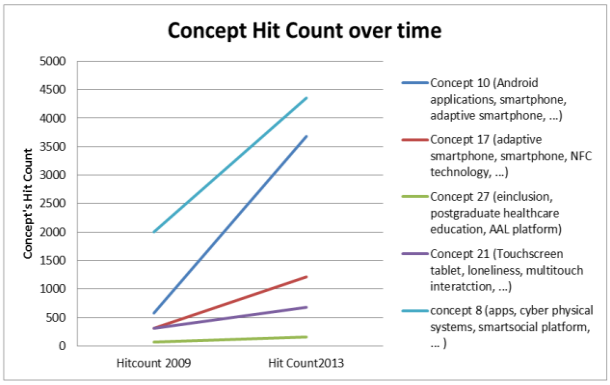


Figure 9 Hit Count for the 5 highest growth concepts

Based on Figure 9, although the concept 8 has the highest hit count, it actually has lower growth rate compare to the other concepts such as concept 17. The growth rate of the five highest ranked concepts is illustrated in the Figure 10 below.

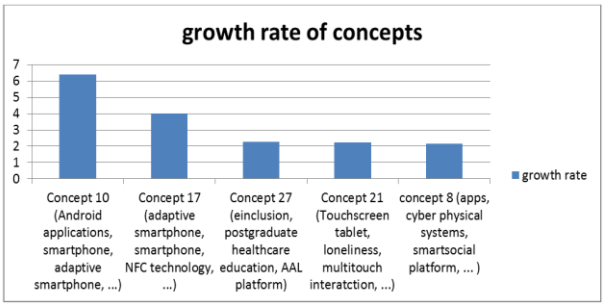


Figure 10 Growth rate of the five highest ranked concepts

The five highest ranking concepts is displayed here in Table 10. From the Table 10 we can see our final result for this research, which are the five concepts that have the highest growth rate in recent years. These concepts are produced by using terms that are highly related to ageing and ICT. The fastest growing concept (concept 10) consists of the terms that are related to Android application and smart phone as its highest weighted components, showing that smart phone are the prevalence technologies in the recent years, especially on the Android applications.

Table 10 The five highest ranking concepts

Concept 10		**Concept 17**	
EigenValues	0.283824	EigenValues	0.069955
Android applications	0.716827	adaptive smartphone	0.948633
smartphone	-0.66177	smartphone	-0.23734
adaptive smartphone	-0.18233	NFC technology	-0.20249
apps	-0.11455	latent prints	-0.03356
tablet computers	0.02265	Android applications	0.024168
smart social platform	-0.0173	apps	0.015671
restful web service	0.01624	smart social platform	-0.01247
Green ICT	-0.0162	restful web service	0.010744
cyber physical systems	0.014631	senior cloud	-0.01072
persuasive mobile interfaces	-0.00919	Touchscreen tablet	-0.00923

Concept 27		**Concept 21**	
EigenValues	0.017192	EigenValues	0.04777
einclusion	0.90953	Touchscreen tablet	0.979958
postgraduate health care education	-0.41285	Loneliness	0.125134
AAL platform	-0.02482	Multi-touch interaction	-0.11022
ICT4D	0.024305	tablet computers	-0.09417
AAL platform evaluation	-0.02369	active societal engagement	0.026661
Touchscreen tablet	-0.01102	AAL platform	0.026308
Web accessibility	-0.00986	mental wellbeing	0.019198
aging friendly ICT technology	0.009082	AAL platform evaluation	0.018698
mental wellbeing	-0.00873	Einclusion	0.013991
Co-creation	-0.00594	smartphone	-0.01093

Concept 8	
EigenValues	0.37305
apps	-0.7239
cyber physical systems	-0.42982
smart social platform	0.388308
smartphone	0.279626
Android applications	0.186678
tablet computers	-0.137
adaptive smartphone	0.084915
NFC technology	0.020627
cloud computing	0.019288
persuasive mobile interfaces	0.011008

4.0 CONCLUSION

In this study, we had performed a bibliometric analysis on the field Ageing and ICT by using only bibliometric information that can be easily extracted from the online publication databases. The objective of our research in determining research area and technologies that are fast growing and related to Ageing and ICT is fulfilled. More than thousands of related terms have been revealed in the first step. The growth rate of these terms is analyzed and a set of ranked terms is then produced, which can be used to identify the current trend in our research field. Furthermore, the relationship of the keywords is determined and organizes in their respective groups using LSA. We then identify the fast growing groups using the growth rate analysis. As a result, we had produced.

As the global ageing phenomena have encouraged various researchers in the field of ageing, Hence, this research has performed a bibliometric analysis to identify related work on the field of ageing and ICT. As a result, we have produced terms that are organized into groups that are called concepts, which are ranked according to their growth rate. These set of concepts has reveal the trend in the current fields of research related to ageing and ICT, and hence can be used as a guide for researchers that are interested to perform research related to ageing and ICT, so that they have a brief understanding of what are promising technologies and fields that are worth to be researched.

The application of our bibliometric analysis on the field of Ageing and ICT shows some reasonable and positive result in revealing the current state of researches. The keywords extraction performed on the online publication databases provided a view on the technologies and research area that is related to our interest field. However, it also contains some irrelevant terms such as 87.19.xr and overall too general terms such as survey. The result of growth rate analysis is reasonable as terms with the highest growth rate are those related to smart technologies, which is prevalence in the recent years. Keywords are then organized into groups according to their relevancy by using LSA and the number of co-hit count. The result is useful in a way that can help researcher to understand which fields are related to each other. Finally, the growth rate analysis is performed on the result of LSA and we had produced a set of concepts with high growth rate, in which the highest ranked concept is those technologies that are related to smart phone and Android application.

Acknowledgement

We are grateful to Universiti Sains Malaysia for providing us facilities to conduct this research.

References

[1] Department of Economic and Social Affairs, Population Division. 2002. *World Population Ageing: 1950-2050*. Available from: http://www.un.org/esa/population/publications/worldageing19502050/pdf/7introduction.pdf. [Accessed 18 Oct, 2013].

[2] UNFPA and HelpAge International. 2012. *Ageing in the Twenty-First Century: A Celebration and a Challenge*. Available from: http://www.unfpa.org/public/home/publications/pid/11584. [Accessed 18 Oct, 2013].

[3] Pritchard, A. 1969. *Statistical Bibliography or Bibliometrics*. Available from: http://www.academia.edu/598618/Statistical_bibliography_or_bibliometrics. [Accessed 19 Oct, 2013].

[4] Mafauzy, M. 2000. The Problems and Challenges of The Aging Population of Malaysia. *Malaysian Journal of Medical Sciences*. 7.

[5] Philip, J.-H. Poi, Duncan, R. Forsyth, and Daniel, K. Y. Chan. 2004. Services for Older People in Malaysia: Issues and Challenges. *Age and Ageing Journal*. 33: 444-446.

[6] Ambigga, K. S., Ramli, A. S., Suthahar, A., Tauhid, N., Clearihan, L. and Browning, C. 2011. Bridging the Gap in Ageing: Translating Policies into Practice in Malaysian Primary Care. *Asia Pacific Family Medicine*. 10.

[7] Hisham, S. and Alistair, D. N. Edwards. 2007. Culture and User-Interface Design for Older Users in Malaysia. *Gerontechnology*. 6: 217-223

[8] C. Y. Wong. 2011. Exploring the Relationship Between Mobile Phone and Senior Citizens: A Malaysian Perspective. *International Journal of Human-Computer Interaction (IJHCI)*. 2.

[9] Komninos, A. and Stamou, S. 2006. HealthPal: An Intelligent Personal Medical Assistant for Supporting the Self-Monitoring of Healthcare in the Ageing Society. In *Proceedings of UbiHealth 2006: The 4th International Workshop on Ubiquitous Computing for Pervasive Healthcare Applications.*

[10] Oguchi, K. 2012. ICT Technologies and Their Application to the Home Environment in Asian/ Oceania Regions: From the View Points of Aging and Home Healthcare. *Review of Asian and Pacific Studies*. 37:153-182.

[11] Johanna, E. M. H. van Bronswijk, William, D. Kearns and Lawrence, R. Normie. 2007. ICT Infrastructures in the Aging Society. *Gerontechnology*. 6: 129-134.

[12] Obi, T. 2012. *ICT* Accessibility Solutions for Japan's Ageing or Disability in Societies. UNDESA Expert Group Meeting April, 19. Tokyo.

[13] Murty, M. 2012. *Malaysia to be Ageing Nation by 2035*. New Straits Times. March, 20.

[14] Patrick Lee. 2013. *A Growing And Ageing Populace*. The Star Online. July, 29.

[15] Catharina R and Ulf K. 2008. *Bibliometric handbook for Karolinska Institutet*. Version 1.05, Karolinska Inst., Solna, Sweden. Available from: http://ki.se/content/1/c6/01/79/31/bibliometric_handbook_karolinska_institutet_v_1.05.pdf.

[16] Tatsuya, Y. 2012. *Communicative Robot Interface for the Ageing Society*. 12th International Conference on Control, Automation, Robotics & Vision, Guangzhou, China, , 5-7 December 2012. 668-671.

[17] Ballesteros, S., Peter, C., Waterworth, E. L., and Waterworth, J. A. 2012. *The role of ICT and networking in healthy ageing.* PETRA '12 Proceedings of the 5th International Conference on PErvasive Technologies Related to Assistive Environments, New York, USA, No. 44, 06 June 2012.

[18] Andrew A. 2009. Web Accessibility for Older Users - Successes and Opportunities. *Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibility (W4A), New York, USA*. 1-6, 20 April 2009.

[19] Norbart, M., Rukiye, O. and Marcelino, C. 2008. *Active Ageing and Independent Living Services: The Role of Information Communication. Technology*. JRC Scientific and Technical Report. European Commission Joint Research Centre Institute for Prospective Technological Studies.

[20] Andrea T, Lorna B, Hugh P, Craig W. 2012. *teleWEAR: Engaging Users and Suppliers of Telecare in Product Design.* Proceeding CHI EA '12 CHI '12 Extended Abstracts on Human Factors in Computing Systems, New York, USA, 05 May 2012. 2637-2642.

[21] Blaine, Z. 2009. *Methods for Bibliometric Analysis of Research: Renewable Energy Case Study*. M.Sc. thesis, Dept. Elect. Eng and Comput. Sci., Massachusetts Inst. of Technology, Cambridge, England.

[22] J. Bae, K. Seetharaman, P. Suntharasaj, Y. Ding. 2007. *Technology Forecasting of RFID by Using Bibliometric Analysis and Bass Diffusion Model* Proceeding PICMET 2007, Portland, USA, 5-9 August, 2007.

[23] Yuya, K., Yoshiyuki, T. 2008. Structure of Research on Biomass and Bio-Fuels: A Citation-based Approach. *Technological Forecasting & Social Change*. 75: 1349-1359.

[24] Ding, Y. and Chowdhury, G. and Foo, S. 2001. Bibliometric Cartography of Information Retrieval Research by Using Co-word Analysis. *Information Processing and Management.* 37(6): 817- 842.

[25] F. S. Wu, C. C. Shiu, P. C. Lee, H. N. Su. 2010. *Integrated Methodologies for Mapping and Forecasting Science and Technology Trends: A Case of Etching*. Technology Proceedings of PICMET 2010, Phuket, Thailand. 1-23.

[26] W. L. Woon, Stuart, M. 2010. *Semantic Distances for Technology Landscape Visualization*. MIT Sloan School Working Paper 4956-11, Cambridge, USA.

[27] F. Janssens, J. Leta, W. Glanzel, B.D. Moor. 2006. Towards Mapping Library and Information Science. *Information Processing & Management.* 42(6): 1614-1642, December 2006.

[28] M. W. Berry, S. T. Dumais, T. A. Letsche. 1995. Computational Methods for Intelligent Information Access. *Proceedings of ACM/IEEE on Supercomputing 1995, New York, USA*. 20.

[29] T. K. Landauer. 1998. An Introduction to Latent Semantic Analysis. *Discourse Processes*. 25: 259-284.

[30] Sequentum, Visual Web Ripper, [online] Available from: http://www.visualwebripper.com/ [Accessed 1 June 2014].

[31] S Michael, Software for Web Scraping, [online] Available from: http://scraping.pro/software-for-web-scraping/ [Accessed 1 June 2014].