Research Trends with Cross Tabulation Search Engine

Chengjiu Yin, Research Institute of Information Technology, Kyushu University, Fukuoka, Japan

Sachio Hirokawa, Research Institute of Information Technology, Kyushu University, Fukuoka, Japan

Jane Yin-Kim Yau, Department of Computer Science, Malmö University, Fukuoka, Sweden

Kiyota Hashimoto, Osaka Prefecture University, Sakai, Japan

Yoshiyuki Tabata, Research Institute of Information Technology, Kyushu University, Fukuoka, Japan

Tetsuya Nakatoh, Research Institute of Information Technology, Kyushu University, Fukuoka, Japan

ABSTRACT

To help researchers in building a knowledge foundation of their research fields which could be a time-consuming process, the authors have developed a Cross Tabulation Search Engine (CTSE). Its purpose is to assist researchers in 1) conducting research surveys, 2) efficiently and effectively retrieving information (such as important researchers, research groups, keywords), and also 3) providing analytical information relating to past and current research trends in a particular field. Their CTSE system employs data-processing technologies and emphasizes the use of a "Learn by Searching" learning strategy to support students to analyze such research trends. To show the effectiveness of CTSE, a pilot experiment has been conducted, where participants were assigned to do research survey tasks and then answer a questionnaire regarding the effectiveness and usability of the system. The results showed that the system has been helpful to students in conducting research surveys, and the research trend transitions that our system presented were effective for producing research trend surveys. Moreover, the results showed that most students had favorable attitudes toward the usage and usability of the system, and those students were satisfied in gaining more know ledge in a particular research field in a short period.

Keywords: Analysis, Learning by Searching, Research Trends, Search Engine, Text mining

1. INTRODUCTION

When initiating a new research project, researchers are often required to conduct a research survey (also known as a literature review), collect papers of relevance and analyze past and emerging research trends. These activities are integral to the research process for researchers of all types including those situated in governments or professional organizations as well as academic ones. As indicated by Hwang and Tsai (2011), "results analysis could help

DOI: 10.4018/jdet.2013010103

policymakers in governments and researchers in professional organisations to allocate the necessary resources and make plans for supporting future research and applications; Doing research survey could be good references for educators and researchers who plan to contribute to the relevant studies." Being able to predict and detect emerging research topics is often a desired (or essential, even) element of conducting high-quality research. A number of methods have been proposed in the literature review within the area of information-based systems to help researchers identify emerging topics within a research area. Five of these different methods/systems are described below:

1. Bun (2005) proposed an Emerging Topic Tracking System (ETTS) which is an information agent for detecting and tracking emerging topics from a particular information area on the Web. It uses a new TF*PDF (Term Frequency * Proportional Document Frequency) algorithm to detect the changes in the information area of the user's interest and generate a summary of these changes for the user at set intervals of time. This summary of changes consists of the latest most discussed research issues and may, as a result, reveal an emerging topic or topics.
2. Decker et al. (2007) used a semantic approach that proposes a method for extracting the names of those researchers in their early stages of a research area, indicated by the amount of high-quality publications. It can be effective in retrieving many exact matches of researchers that have major contributions within the research area being explored.
3. The Hierarchical Distributed Dynamic Indexing (HDDI) system mentioned in Bouskila and Pottenger (2000) aims to identify features and methods for improving the automatic detection of emerging trends by generating clusters based on semantic similarity of textual data. The rate of change in the size of clusters and in the frequency and association of features is used as input for applying machine learning techniques to classify topics as emerging or non-emerging.
4. Collaborative Inquiry-based Multimedia E-Learning (CIMEL) is a multi-media framework for constructive and collaborative inquiry-based learning (Blank et al., 2001). The semi-automatic trend detection methodology described in (Roy et al., 2002) has been integrated into the CIMEL system in order to enhance computer science education. Citations traces are used with pruning metrics to generate a document set for an emerging trend. Following this, threshold values are tested to determine the year that the trend was emerged. A multimedia tutorial has been developed to guide students through the process of emerging trend detection.
5. Moreover, some researchers used bibliometric methodologies to analyze the trends and forecasts in different domains, such as e-commerce, supply chain management and knowledge management (Tsai, 2011; Tsai & Chi, 2011; Tsai & Chiang, 2011). Using a bibliometric approach, (Tsai & Yang, 2010) analyzed data mining and CRM research trends from 1989 to 2009 by locating headings "data mining" and "customer relationship management" or "CRM" in topics in the SSCI database. Especially, the approach utilized search categories such as publication year, citation, country/territory, document types to explore the differences in the two fields.

As mentioned above, research trend survey is an essential preliminary step for conducting any academic or non-academic research. In particular, many junior researchers experience difficulties in locating the appropriate keywords and subsequently experience difficulties for conducting a literature review in their research field(s). With the development of ICT technologies such as data-processing, it is possible to design search engines to address such learning difficulties or research needs. Data-processing is a broad category which includes functions such as search engines, data mining, recommendations, and image recognition; and such

technologies have been utilized to build our Cross Tabulation Search Engine (CTSE) system.

In this paper, we describe this system - Cross Tabulation Search Engine (CTSE) which we constructed in order to support in particular junior researchers for conducting and analyzing research trend surveys in a more straight-forward manner. Our CTSE system can analyze large amounts of information in a very short time in order to provide insight into the distinct changes occurring in different research fields and subsequently present appropriate and detailed emerging research trends for students/users.

Our CTSE search engine has been populated with 13326 articles and papers from SciVerse Scopus (http://www.info.sciverse.com/scopus/) that were publish from the year 1992 to 2012, and contain the keyword "e-learning", these articles are all related to the use of computer technologies in education, which include "mobile learning". To use our system, the user inputs keywords of topics that they intend to search and find information on. The system then presents to them a list of publications including the years, authors, countries, etc. The user can then choose the x and y axes on the system to compare these publications using different categories (such as years, authors). The relevance of inputted keywords is calculated by our system using co-occurrence frequency. This means that if a word has a higher relevance to the keywords given in the abstracts of papers (i.e. co-occurrence frequency), these words are listed up as feature words. Subsequently, the system gives the user effective visualizations to understand the research trend transitions.

The previous studies were about proposing new methods in the literature review within the area of information-based systems. Our CTSE system is distinct from the methods/systems in the five categories described above for analyzing research areas, in that our system not only presents the emerging topics of the completed research, but also presents current occurring emerging topic trends and gives the user effective visualizations to understand the research trend transitions.

In terms of pedagogical significance and effectiveness, our CTSE system has two features to facilitate these:

1. Users can use the system to analyze and compare knowledge/information and research methods in the literature review. The system can also retrieve and present related literature and papers to users and therefore helping the students to learn how to do literature retrieval and analysis naturally.
2. Students are able to retrieve large volume of information regarding the research trend progresses and keyword search transitions of their research field(s) using the system. Therefore, the pace in which students conduct scientific research and research trend surveys can be speeded up using the system and to subsequently obtain scientific research results and achievements at a quicker pace. Using this information, students become more knowledgeable about emerging research trends in their research area and decide their research topics accordingly, and make important decisions about their topics of interest, as well as allowing students to more accurately predict how they should position their own research in the future.

The organization of the paper is as follows. In section 2, a literature review including technologies, methods and learning strategies relating to the foundation of CTSE is presented. In section 3, the CTSE system is described together with the implementation of it. In section 4, we present a pilot study using the CTSE system and a discussion and implication of the results obtained from this study. Finally, in section 5, conclusion and future works are discussed.

2. LITERATURE REVIEW

In this section, a literature review concerning some of the technologies, methods and learning strategies motivating the design of the CTSE

system is presented. Specifically, these are *Mindtools*, and the "*Learning by Searching*" learning strategy.

2.1. Mindtools

Mindtools are computer-based learning applications which serve as extensions of the mind (Jonassen, 1996). Jonassen (1999, p9) described Mindtools as "*a way of using a computer application program to engage learners in constructive, higher-order, critical thinking about the subjects they are studying*". Mindtools include databases, spreadsheets, concept maps, computer conferencing, simulation programs, and communication facilities such as online discussion forums and search engines (Wu et al., 2010). These tools are used to demonstrate how learners can model what they know, and build model, externalize learners' conceptions (Bratt & McCracken, 2007). For example, Hwang et al. (2011) developed a Mindtool to help students organize and share knowledge for differentiating a set of learning targets based on what they have observed in the field.

Our CTSE system can be regarded as a Mindtool in the sense that it allows information relating to research articles or papers to be organized and the relationships between these can be displayed visually to users. Users can use the CTSE to search for and analyze articles or papers in order to answer specific questions relating to their research. The system also teaches students to explore and examine the data in order to reveal important information relating to their research. The system also allows relational databases to be displayed. These databases enable users to interrelate information in several tables, which could support higher level thinking for advanced learners.

2.2. "Learning by Searching" Strategy

The act of searching for information is a cognitive process that acquires knowledge actively and we have defined this act as 'learning by searching', for the purpose of this study. Searching information online using search engines such as Google is a common everyday practice carried out by many people in order to obtain information, for solving problems or completing study tasks. Searching information online has becoming a part of many people's everyday learning processes or activities. Given the advancement of internet and search engine capabilities, 'learning by searching' has become an important style of learning or obtaining of information utilized by academic/non-academic junior as well as senior researchers. Liu (2008) argued that the skill to 'learn by searching' is necessary for students and researchers. The cognitive processes underlying information searching have been investigated by a number of researchers. For example, State (2009) examined the search habits of 72 participants while conducting a total of 426 searching tasks. It was found that the searching of information imitated a learning process rather than simply a way of obtaining information. Therefore implying that 'learning by searching' is a learning strategy, as opposed to a searching process. Similarly, Bruner (1967) suggested that students were more likely to remember concepts if they discovered them on their own. In the 21st century, 'learning by searching' provides newly developed pedagogy to meet the knowledge needs of learners and it is a strategy that advocates students to take the initiative to acquire knowledge. However, search engines currently do not provide returned results which are categorized into different groups or areas (of research, for example).

As part of our CTSE system, we have developed a search engine which is able to support the 'learning by searching' learning strategy or process. Our search engine enables students to master some of the basic concepts and methods of searching and locating scientific literature survey during the process of document retrieval. Students can recognize and learn the research trends in depth through accessing and searching the data and viewing the analysis results provided by the system. Our search engine can support students' individual learning, it can be used to broaden their sources of knowledge and improves their self-learning ability. The role of

the instructors is switched from the providers of information to the facilitators of students' active and initiated individual learning.

3. THE IMPLEMENTATION OF THE SYSTEM

We downloaded 13326 articles and papers which range from 1992 to 2012, as of June 24, 2011. Note that the number of documents of 2011 is small, since most of them are not open to public yet. Each document of the project contains, *title, project id, *period,*represe ntative,*members, genre, keywords, outline and amount. We indexed the marked(*) items to construct a search engine. We ignored the representative and members in our implementation. We used special marks such as "y:2009" and "n:SachioHirokawa" to distinguish the year information and the name of the researcher from usual keywords, when they are indexed.

3.1. General Approach

In order to help students to grasp the outline, problem, method and solution of the literature efficiently, our system provides a method of extracting sentences which describe existing research problems automatically; these problems are retrieved from paper abstracts in the literature review using clue words. Johan et al. (2005) determined that "when two terms frequently co-occur in, for example, the same sentence, they are related. The more frequently they co-occur compared to how often they each occur individually the stronger we will assume their relationship to be". Therefore, the feature words of search results should have the highest co-occurrence frequency with inputted keywords. Our system calculates relevance of inputted keywords using co-occurrence frequency. If the words have higher co-occurrence frequency, then the words have higher relevance of inputted keywords, and these words are list up as feature words.

3.2. Interface of the CTSE

As show in Figure 1, a user can control his/her process by specifying (a) the queries for search and (b) the features to be displayed. The search results are shown in (d) matrix form together with (c) the ordinary listing of articles (see Appendix). The characteristics of our search engine is in the listing of articles obtained as a search result, but in the way that feature words of search result are displayed. User specifies 2 features from "word", "year", "country", "author", "source" or "organization". The number of articles that matched the pair of feature words are shown in a matrix map. Imagine that a list of articles outlines are obtained for a query "q", that R1,R2,R3,R4 and R5 are the top 5 researchers and that W1,W2,W3,W4 and W5 are top 5 feature words in the search result. The system conducts 5*5 search with "q and Ri and Wj" to calculate the number projects that matches the condition. The number is displayed in the (i,j)-th cell of the matrix.

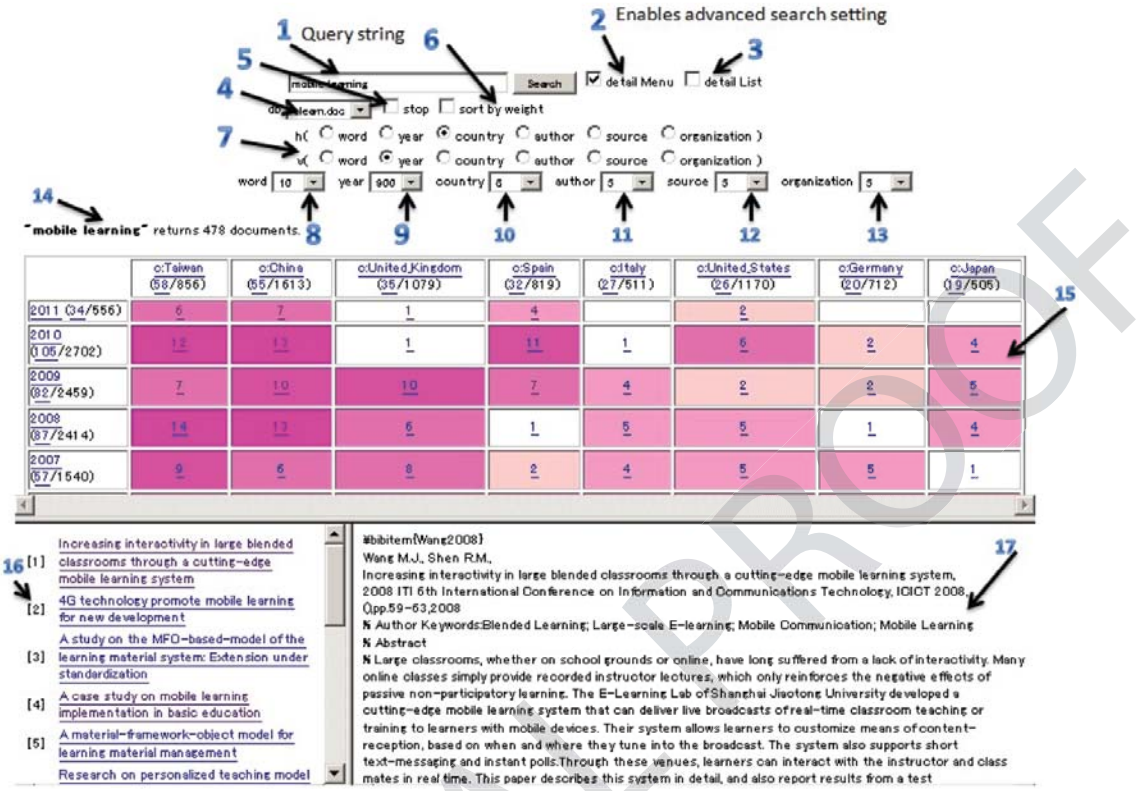
3.2.1. Input Query and Parameters

The "input query and parameters" includes "keywords ", "Detail menu", "Detail list", "DB", "Exclusion of stop words", "Sort by weight","Axes of matrix", "Word", "Year","Country", "Author", "Source" or "Organization"..

3.2.2. Output of Features

The output have features such as "to show the number of search results", "to show the number of articles per year", "to show the top-N of authors" and "to show top-M related words". The fraction shows the number of articles that match the query and the keyword. The denominator displays the number of articles that match the query. A click on the faction yields a narrowing search using the query and the keyword. A click on the denominator yields a new search with the keyword without the query.

Figure 1. Interface



3.2.3. Matrix display

The search result is shown in a matrix, where the x-axis and y-axis are determined in 5 and the number in each cell means the number of documents that match both of x-axis and y-axis. The keywords, years, names of researchers in the first column and in the first row can be used for a new search. A click on a cell displays the list of titles of the articles in the lower frame.

3.2.4. List of Article Titles

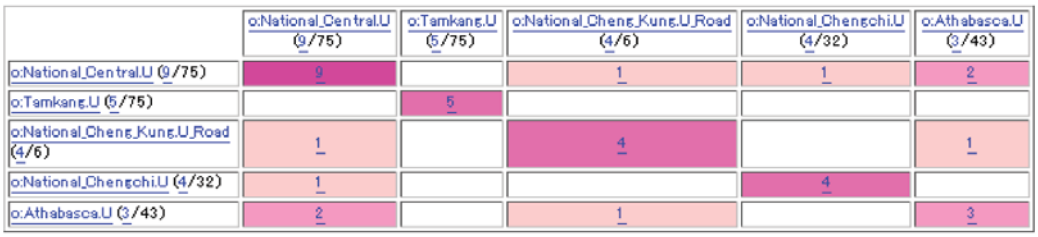
When users click on the cell. The top 10 titles of the articles will be shown from the list.

3.3. Analysis Samples

In the remainder of this section, we show a case study with respect to the query "mobile learning". Figure 1 displays the 478 articles retrieved by the query where the publication year and the country of the authors were chosen as the x and y axes, where the top 8 countries were chosen to be compared. We can see that the number of articles are increasing in Taiwan and China. On the other hand, UK, which had a large publication until 2008, has few articles in 2010. Spain is increasing the number of articles after 2009. The list of titles in lower-left of the Figure 1 displays the articles publish in 2008 by Chinese authors. The list is generated by a simple click on the cell where the line of "2008" and the column of "c:China" crosses. The detailed meta-data of the first article is shown in the lower-right of the Figure 1. A click on a title in the left lists gives the detailed information of the article. Analysis begins by a query from a user. However, once a user gives a query, he can choose the axes for his viewpoints, select the cell or the article to see in detail by a click with his mouse. All he needs to view the tabulation and evaluate the importance of particular target to be analyzed further.

Figure 2 displays the top 5 organizations with respect to a query "mobile learning c:Taiwan". Figure 2 is obtained by choosing the organization as both x and y axes. The cross

Figure 2. Analysis with organization by Organization



	o:National_Central.U (9/75)	o:Tamkang.U (5/75)	o:National_Cheng_Kung.U_Road (4/6)	o:National_Chengchi.U (4/32)	o:Athabasca.U (3/43)
o:National_Central.U (9/75)	9		1	1	2
o:Tamkang.U (5/75)		5			
o:National_Cheng_Kung.U_Road (4/6)	1		4		1
o:National_Chengchi.U (4/32)	1			4	
o:Athabasca.U (3/43)	2		1		3

tabulation shows the relationship of organizations. We can found that the four universities (except Tamkang U.) have collaborative articles with joint authorship.

To find the key researchers of each university, we chose the authors as y axis and obtain the Figure 3. The five articles by Tamkang U. are made by the group of "a:Shih_T.K. (4/33)" and "a:Chang_C.-Y. (3/5)". The first line of Figure 3, we see the names of researchers followed by two numbers. The first number shows the number of the articles that matches the query by the researcher. The second number shows the number of all articles by the researcher. Thus, we can interpret "a:Shih_T.K. (4/33)" as that Shin T.K. is a senior researcher who have been working not only in mobile learning but in other related area and that Chang C.-Y. might be a young researcher who started recently and is focused in mobile learning.

Figure 4 displays the author*author cross tabulation of the search result with respect to a query "mobile learning c:Taiwan o:Tamkang_University". This query is determined by clicking the number "5" at the display of "o:Tamkang.U (5/75)" in Figure 3. We can see two groups. The first group consists of Kao T.-C, Chang C.-Y, Shih T.K. and Shih K.-P. The other consists of Chang C.S. and Chen C.H.

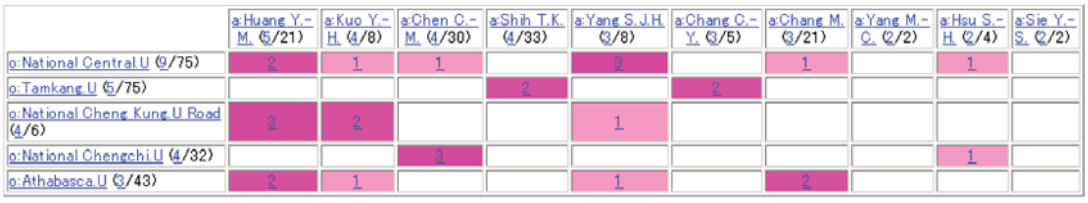
Figure 5 shows the relationship of top 10 researchers with respect to "learning c:Taiwan o:Tamkang_University". Figure 5 is an expanded analysis of Figure 4 by removing the keyword "mobile". Figure 4 shows the organizational approach of Tamkang U. in the field of learning.

4. PILOT EXPERIMENT

In order to evaluate the effectiveness of our CTSE system, an experiment was conducted. Nine students from the graduate school of Information Science and Electrical Engineering of four university participated in this experiment. These included four fourth-year undergraduate, four master and one PhD students. We required that participants must have some previous experience of conducting literature review or research surveys manually using Internet search engines in order to evaluate the system. All nine participants met this requirement.

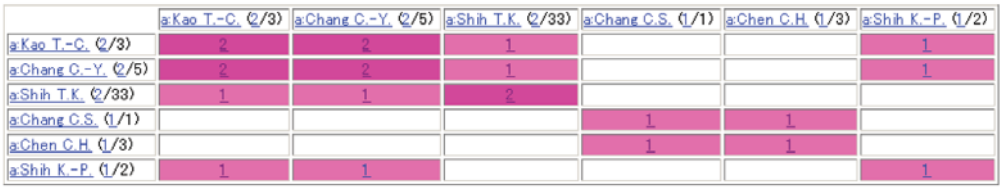
In this section, first the experiment methodology is described, followed by the study results from this experiment, followed by the results from a questionnaire study relating to the CTSE system, finally some qualitative feedback regarding the usefulness of the system.

Figure 3. Analysis with organization by Author



	a:Huang Y.-M. (5/21)	a:Kuo Y.-H. (4/8)	a:Chen C.-M. (4/30)	a:Shih T.K. (4/33)	a:Yang S.J.H. (3/8)	a:Chang C.-Y. (3/5)	a:Chang M. (3/21)	a:Yang M.-C. (2/2)	a:Hsu S.-H. (2/4)	a:Sie Y.-S. (2/2)
o:National Central.U (9/75)	2	1	1		3		1		1	
o:Tamkang.U (5/75)				2		2				
o:National Cheng Kung.U Road (4/6)	3	2			1					
o:National Chengchi.U (4/32)			3						1	
o:Athabasca.U (3/43)	2	1			1		2			

Figure 4. Analysis with Author by Author



	a:Kao T.-C. (2/3)	a:Chang C.-Y. (2/5)	a:Shih T.K. (2/33)	a:Chang C.S. (1/1)	a:Chen C.H. (1/3)	a:Shih K.-P. (1/2)
a:Kao T.-C. (2/3)	2	2	1			1
a:Chang C.-Y. (2/5)	2	2	1			1
a:Shih T.K. (2/33)	1	1	2			
a:Chang C.S. (1/1)				1	1	
a:Chen C.H. (1/3)				1	1	
a:Shih K.-P. (1/2)	1	1				1

4.1. Methodology

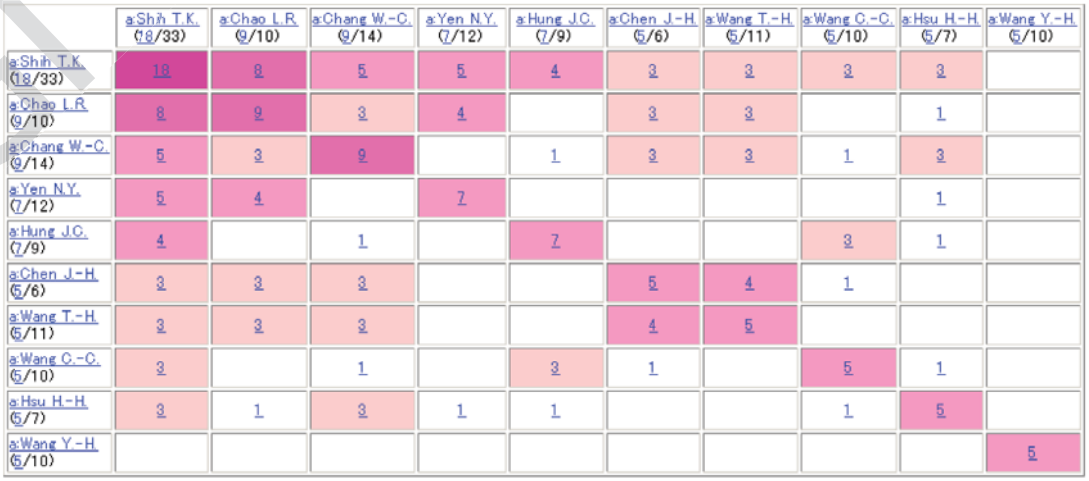
Participants were explained the functions of our CTSE system, how the system works and how they can use the system to perform various individual tasks such as how to find the researchers, investigate and analyze the research trends. The experiment required participants to perform four tasks (shown in Table 1) which consisted of conducting research surveys in the field of mobile learning using the CTSE system. They were asked to write down their research survey results accordingly with pen and paper. They spent about two hours in total on the tasks. They completed these surveys independently, without any instructions or any help from others. Their written papers were collected afterwards for data analysis.

4.2. Study Results

Table 2 shows the participants' results of conducting the four above-listed tasks in the form of research surveys using our CTSE system. The participants' answers to the tasks showed that they retrieved the relevant information or knowledge about the mobile learning field. Tasks 1 to 3 required participants to retrieve factual information and participants have retrieved these successfully; the answers by the nine participants were almost identical. Task 4 is a research-based question requiring more analytical and in-depth skills from participants, in order to obtain the answers to this task. From these results, we can conclude that using the CTSE system, it is possible for students to conduct research surveys in a shorter time and it can support their acquiring of information and knowledge.

After the experiment had taken place, an evaluation questionnaire regarding the CTSE system was given to all of the participants to complete. Table 3 shows the questionnaire and the results. The questionnaire consists of 7 questions - Q1 to Q7 were measured using four-point Likert scales varied from '1 - strong-

Figure 5. Expanded Analysis



	a:Shih T.K. (18/33)	a:Chao L.R. (9/10)	a:Chang W.-C. (9/14)	a:Yen N.Y. (7/12)	a:Hung J.C. (7/9)	a:Chen J.-H. (5/6)	a:Wang T.-H. (5/11)	a:Wang C.-C. (5/10)	a:Hsu H.-H. (5/7)	a:Wang Y.-H. (5/10)
a:Shih T.K. (18/33)	18	8	5	5	4	3	3	3	3	
a:Chao L.R. (9/10)	8	9	3	4		3	3		1	
a:Chang W.-C. (9/14)	5	3	9		1	3	3	1	3	
a:Yen N.Y. (7/12)	5	4		7					1	
a:Hung J.C. (7/9)	4		1		7			3	1	
a:Chen J.-H. (5/6)	3	3	3			5	4	1		
a:Wang T.-H. (5/11)	3	3	3			4	5			
a:Wang C.-C. (5/10)	3		1		3	1		5	1	
a:Hsu H.-H. (5/7)	3	1	3	1	1			1	5	
a:Wang Y.-H. (5/10)										5

Table 1. Tasks

Tasks	Content
(a)	Find some of the main researchers in mobile learning area.
(b)	Find some influential research groups in the mobile learning area.
(c)	Find some main research keywords in the mobile learning area.
(d)	Investigate the research trends of mobile learning in the last 10 years and the appearance of mobile learning.

ly disagree' to '4 - strongly agree'. The participants were also asked to comment on their answers and to provide some suggestions for improving the system.

From Q1, the mean score of the question was close to 4 (strongly agree), which means that the participants strong agreed that the system was easy to find the required information such as researchers; Q2 and Q3 presented that 66.7% of the participants agreed that it was easy to find researcher groups and main research keywords, 33.3% of participants did not agree; From Q4, 83.3% of the participants confirmed the changes of mobile learning research trend for each year, 16.7% of them cannot; whereas, Q5 indicated that the opinions of the participants were divided, which means that it was not easy to investigate when the appearance of mobile learning occurred. Q6 presented that all of them agreed that it was easy to compare the differences of mobile learning with courtiers. From Q7, 83.3% of the participants agreed that it was easy to compare the feature keywords of mobile learning, 16.7% of them did not. These results indicate that our CTSE can help researchers to perform research surveys easier than if done manually without the system. In some way, the 'learning by searching' strategy is also supported as shown by the open-ended research task of finding the mobile learning research trend (task 4), participants must have acquired the information/knowledge in order to answer this question. It can be seen that some 'learning by searching' had taken place in this scenario.

Table 2. Results Summary From All Nine Participants

Tasks	Content
(a)	Martin S.S, Castro M.S., Barron- Estrada M.L., Huang Y.M.
(b)	National Central U, Athabasca U., Tamkang U.
(c)	Anywhere, Anytime, Wireless, Context, Phone, Mobile, Digital, M-learning, Devices, PDA, smart phone and tabulate PC
(d)	•Some paper were published in 1997, until 2003 the number of paper is still limited. But the number of the paper surged in 2004, until 2010 the number continued growth. •After 2007, the research on mobile learning increased very fast in China and Taiwan •From 2004, the number of the papers increased every year. •National Central U is very activity on the research of mobile learning •There are many papers from China and Taiwan, but there are not so many paper in Japan. •There are many papers from China and Taiwan, the papers increased recent two years in Spain •There are many studies on mobile learning from 2008 to 2010. •Asia is the most active areas in mobile learning research. •Spain has higher growth rate on mobile learning research.

Table 3. Results Summary From All Nine Participants

Questions		SA/A	D/SD	M	SD
Q1	It is easy to find researchers in mobile learning	100%	0.0%	3.67	0.52
Q2	It is easy to find researcher groups in mobile learning	66.7%	33.3%	3	0.89
Q3	It is easy to find main research keywords of mobile learning.	66.7%	33.3%	3	0.89
Q4	It is easy to confirm the changes of mobile learning research trends.	83.3%	16.7%	3.5	0.84
Q5	It is easy to investigate a appearance of mobile learning	50.0%	50.0%	2.5	0.55
Q6	It is easy to compare the differences of mobile learning with courtiers.	100%	0.00%	3.5	0.55
Q7	It is easy to compare the feature keywords of mobile learning.	83.3%	16.7%	3	0.63

Note. SA/A-Strongly Agree and Agree; D/SD- Disagree and strongly disagree; M-Means; SD-Standard Deviation

4.3. Qualitative Feedback Regarding The Usefulness Of The System

Positive feedback regarding the CTSE system was given by the majority of the participants. As shown by the study results, participants retrieved the main concepts of mobile learning such as "anytime, anywhere" and they found that many of these research papers were written by researchers in Taiwan, which was indeed the case. They can also describe the changes of the mobile learning research trends for each year accurately. Some of the positive feedback given by participants includes the following:

- The system is useful for conducting research trends survey in that research transitions, researchers and research groups are highlighted by the system. One participant would recommend the system to others for conducting research surveys.
- The system is helpful in providing visualizations of the search results using time series, as well as locating accurately popular and new emerging research topics.
- The system is relatively easy to use and participants were able to conduct the research survey more efficiently and effectively.
- The system is helpful for searching for articles or papers under several different categories such as year, feature, word, country, organization. It is useful as results can be displayed from different perspectives. Useful comparisons from these categories can also be made between one another and can be very useful.

5. CONCLUSION AND FUTURE WORKS

From SciVerse Scopus, we collected 13326 "e-learning" articles and papers that were published from the year 1992 to 2012. Using data-processing technologies, we built a Cross Tabulation Search Engine (CTSE) system, which is aimed for assisting junior researchers to conduct more efficient and effective research including development research surveys and analyzing research trends. The system emphasizes the "learning by searching" learning strategy. As shown in our study experiment, participants were able to use the system to retrieve information about a research field and to obtain the necessary and accurate data regarding researchers, research groups, and research trends. Using this information, students can decide upon their research topic, narrow down their focus within a particular area and be able to keep up-to-date with current existing research, as well as knowing about research trends that have previously been taken place and which are of importance to the research work of current researchers. The CTSE system can provide researchers such information is a much shorter time than if they had done the search manually using Internet search engines such as Google.

As demonstrated in our pilot study, two hours was necessary for participants to find out the answers to the required questions/tasks. Normally speaking, this is a much shorter time than one would require to find on the Internet. This is because any Internet search engine would generate a much larger quantity of information on any particular keywords such as "mobile learning", "anytime", "anywhere", etc, and young researchers may not be able to identify immediately which articles are of more importance than others. Therefore, it would be more difficult to answer those questions in a short time by junior researchers using a standalone Internet search engine.

The current limitations in the present pilot study are that the findings were based on an experiment with nine students; therefore, generalizations about students' attitudes could be difficult to make based on such a small sample size. However, this study has been sufficient to make the preliminary conclusions that our system can be effective for junior researchers to conduct research and retrieve relevant information. The research students at our university are currently using this system to conduct their research surveys and have commented the usefulness and effectiveness of the system. In the future, we plan to conduct an larger-scale experiment to evaluate the learning effectiveness of the system, as well as to make more in-depth comparisons between the participants' results obtained using the CTSE system and using normal Internet search engines. To summarize, the main feature of our system is to analyze effectively and efficiently large amounts of information in the shortest time, providing appropriate research information and trends for students/users.

REFERENCES

Baker, R. S., Corbett, A. T., & Koedinger, K. R. (2004). Detecting student misuse of intelligent tutoring systems. In *Proceedings of the 7th International Conference on Intelligent Tutoring Systems*, Maceio, Brazil (pp. 531-540).

Baker, R. S. J. D., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, *1*(1), 3–17.

Blank, G. D., Pottenger, W. M., Kessler, G. D., Herr, M., Jaffe, H., & Roy, S. … Wang, Q. (2001). CIMEL: Constructive, collaborative inquiry-based multimedia e-learning.In *Proceedings of the 6th Annual Conference on Innovation and Technology in Computer Science Education (ITiCSE).*

Bollen, J., Nelson, M. L., Manepalli, G., Nandigam, G., & Manepalli, S. (2005). Trend analysis of the digital library community. *D-Lib Magazine, 11*(1). <doi:10.1045/january2005-bollen>.

Bouskila, F., & Pottenger, W. M. (2000). The role of semantic locality in hierarchical distributed dynamic indexing. In *Proceedings of the 2000 International Conference on Artificial Intelligence (IC-AI 2000)*, Las Vegas, NV.

Bratt, S. E., & McCracken, J. (2007). [Reveiew of the book *Modeling with technology: Mindtools for conceptual change,* by David H. Jonassen]. *Journal of Educational Technology & Society*, *10*(2), 225–227.

Bruner, J. S. (1967). *On knowing: Essays for the left hand.* Cambridge, MA: Harvard University Press.

Bun, K. K. (2005). Topic trend detection and mining in world wide web. Japanese Society for Artificial Intelligence, 27(6).

Decker, S. L., Aleman-Meza, B., Cameron, D., & Arpinar, I. B. (2007). *Detection of bursty and emerging trends towards identification of researchers at the early stage of trends*. Athens, GA: University of Georgia, Computer Science Department.

Desmarais, M. C., & Pu, X. (2005). A Bayesian student model without hidden nodes and its comparison with item response theory. *International Journal of Artificial Intelligence in Education*, *15*, 291–323.

Hwang, G. J., Chu, H. C., Lin, Y. S., & Tsai, C. C. (2011). A knowledge acquisition approach to developing Mindtools for organizing and sharing differentiating knowledge in a ubiquitous learning environment. *Computers & Education*, *57*, 1368–1377. doi:10.1016/j.compedu.2010.12.013.

Hwang, G. J., & Tsai, C. C. (2011). Research trends in mobile and ubiquitous learning: A review of publications in selected journals from 2001 to 2010. *British Journal of Educational Technology*, *42*(4), 65–70. doi:10.1111/j.1467-8535.2011.01183.x.

Jonassen, D. H., Carr, C., & Yueh, H. P. (1998). Computers as Mindtools for engaging learners in critical thinking. *TechTrends*, *43*(2), 24–32. doi:10.1007/BF02818172.

Kelly, G. A. (1955). *The psychology of personal constructs*. New York, NY: Norton.

Kindberg, T., & Fox, A. (2002). System software for ubiquitous computing. *IEEE Pervasive Computing / IEEE Computer Society [and] IEEE Communications Society*, *1*(1), 70–81. doi:10.1109/MPRV.2002.993146.

Liu, H. (2008) Learning by searching. In K. McFerrin et al. (Eds.), In *Proceedings of Society for Information Technology & Teacher Education International Conference 2008* (pp. 3843-3844). Chesapeake, VA: AACE.

Pavlik, P., Cen, H., Wu, L., & Koedinger, K. (2008, June 20-21). Using item-type performance covariance to improve the skill model of an existing tutor. In *Proceedings of the 1st International Conference on Educational Data Mining,* Montreal, Canada (pp. 77-86).

Roy, S., Gevry, D., & Pottenger, W. M. (2002, April 11-13). Methodologies for trend detection in textual data mining. In *Proceedings of the Textmine '02 Workshop, Second SIAM International Conference on Data Mining*, Chicago, IL.

State, P. (2009). Search engines are source of learning. *ScienceDaily*. Retrieved May 2, 2011, from http://www.sciencedaily.com /releases/2009/11/091119111417.htm.

Tsai, H. H. (2011). Research trends analysis by comparing data mining and customer relationship management through bibliometric methodology. *Scientometrics*, *83*(2), 425–450. doi:10.1007/s11192-011-0353-6.

Tsai, H. H., & Chi, Y. P. (2011). Trend analysis of supply chain management by bibliometric methodology. *International Journal of Digital Content Technology and its Applications, 5*(1), 285–295.

Tsai, H. H., & Chiang, J. K. (2011). E-commerce research trend forecasting: A study of bibliometric methodology. *International Journal of Digital Content Technology and its Applications, 5*(1.12), 101–111.

Tsai, H. H., & Yang, J. M. (2010). Analysis of knowledge management trend by bibliometric approach. In *Proceedings of the WASET International Conference on Knowledge Management and Knowledge Economy*, Paris, France (vol. 62, pp. 174–178).

Wu, P.-H., Hwang, G.-J., Tsai, C.-C., Chen, Y.-C., & Huang, Y.-M. (2010). A pilot study on conducting mobile learning activities for clinical nursing courses based on the repertory grid approach. *Nurse Education Today*. doi: doi:10.1016/j.nedt.2010.12.001 PMID:21196068.

Chengjiu Yin received his PhD degrees from the Department of Information Science and Intelligent Systems, Tokushima University, Japan, in 2008. He is an Assistant Professor in the Research Institute for Information Technology, Kyushu University. Currently he is committing himself in mobile learning, ubiquitous computing, language learning, text mining and SNS. He received the best paper award from ICIE in 2009. Dr. Yin is a member of JSiSE, JSET, and APSCE.

Sachio Hirokawa is professor of Research Institute for Information Technology, Kyushu University, Japan. He studied mathematics and computer science at Kyushu University. He was appointed to a research assistant at Shizuoka University in 1979, moved to Kyushu University in 1988 as Associate Professor and Professor in 1996. He received PhD in 1992. He has been involved in research and teaching in the area of mathematical logic and computer science. Since late 90s, His research focuses on search engine and text mining, where frequency analysis and visualization are the key features. He conducted 3 years project on search engine and became founder of start-up company Lafla (http://www.lafla.co.jp) to realize his technologies for commercial services.

Jane Yin-Kim Yau was a postdoctoral fellow at the Center for Learning and Knowledge Technologies at Linnaeus University, Sweden from November 2010 - May 2012. In June 2012, she joined the Dept of Computer Science at Malmö University as a postdoc in Computer Science. She obtained her doctorate from the Dept. of Computer Science, University of Warwick, UK in 2011, which was focused on a mobile context-aware learning schedule framework with Java learning objects. She has published around thirty articles in journals and conferences in mobile learning.

Kiyota Hashimoto has a BA and MA in Linguistics from Kyoto University, and Doctor of Engineering from Nara Institute of Science and Technology. He worked at Seiwa College and Osaka Women's University, and currently is an associate professor at College of Knowledge & Information Systems, Osaka Prefecture University, Japan. Some of his interests are natural language processing, educational engineering for language learning/teaching, tourism informatics, and soft computing in general.

Yoshiyuki Tabata is a Professor in the Research Institute for Information Technology, Kyushu University. He received his BA and MA degrees from Tokyo University of Foreign Studies, Japan, in 1983 and 1986. He is a member of JGG, JDV, GDDJ, JAECS and JEI. His current interests are in language learning environments supported by ICT.

Tetsuya Nakatoh is the assistant professor of Research Institute for Information Technology, Kyushu University, Japan. He studied computer science at Kyushu University. He was appointed to a research assistant at Kyushu University in 1993. He was appointed to an assistant professor at Kyushu University in 2007. He received PhD in 2010. Since the start of the 2000s, his research area is search engine, feature extraction, pattern discovery and string matching algorithm.

APPENDIX

System Function

Table 4.

	Title	Functions
(a)	Input query and parameters	1. Keywords: search query string, the special keyword "z" returns all the documents. 2. Detail menu: toggle advanced search settings display. 3. Detail list: Toggle verbose results output. 4. DB: choice of databases such as e-learning or foreign. 5. Exclusion of stop words: The frequent words are excluded as stop words. 6. Results sorting method: Choose sort function from document frequency or weight(default). 7. Axes of matrix: x-axis and y-axis in matrix display can be chosen from "word", "year", "country", "author", "source", and "organization". 8. Word: Number of top ranked keywords to be analyzed for each year. 9. Year: year range. 10.Country: The number of countries to be analyzed. 11. Author: Number of top ranked authors to be analyzed for each year. 12. Source (e.g. Publisher): Number of top ranked sources be analyzed for each year. 13. Organization (e.g. University): Number of top ranked organizations for each year.
(b)	Output of Features	14. The number of search results.
(c)	Matrix display	15.The search results are shown in a matrix and the number of articles are also shown.
(d)	List of articles titles and abstract	16. The titles of the articles are shown by clicking each cell. 17. The abstract of the article is shown by clicking the titles.