# 推薦導向之專題知識檢索系統

# A Recommender-based Knowledge Support System for Senior Project Design

作者：涂嘉安、王昱森、黃伊萍
系級：資訊四甲
學號：D9120458、D9155925、D9156180
開課老師：竇其仁 老師
課程名稱：專題研究（三）
開課系所：資訊工程學系
開課學年： 94 學年度 第 一 學期

# 中文摘要

現今的工程教育中，終端課程扮演了相當重要的角色，藉著終端課程的完成，可以讓學生破除只懂理論不會實作的弊病，也可在求學階段培養專案管理、團隊合作的能力。而終端課程也可以顯示一個系所的研究方向，甚至可藉由終端課程評鑑系所。學生在終端課程的學習中，不約而同地會遭遇同樣的問題，往往每年都重複花上一段時間去解決同樣的問題。本篇專題著眼於「專題知識檢索系統」的建立、知識擷取的功能以提供對使用者有用的資料、利用關聯式索引將獨立的專題串聯起來變成知識網，藉由推薦功能讓使用者一層一層深入知識核心，進行資料挖掘，以及歷年趨勢分析看出各個系所的方向及特色，也可以從中看出老師的教學歷程。在這個系統建立完成之後，開放使用，紀錄使用者的網站尋訪軌跡，蒐集意見回饋並進一步改善之，如此一來，可觀察分析使用者的知識吸收慣性，也可以加強「專題知識檢索系統」的實用性，由實驗結果數據也表示出：「專題知識檢索系統」的確是可以在終端課程的學習中提供實質的貢獻。

**關鍵詞:** 推薦者系統、知識檢索、關聯式索引、專題製作。

# Abstract

Senior project design plays an essential role in an engineering educational curriculum. How to properly judge a student's abilities, and how to convert senior project profiles into information useful for reuse, are our key concerns here. Moreover, while designing, students often encounter the same difficulties and spend lots of time solving repeated and previously solved problems. According to issues thus said, this work designs and implements a knowledge support system that employs knowledge retrieval and relational index technology to provide knowledge and decision support, making it possible for students to tap easily on the knowledge established previously. Moreover, senior projects from 1998~2004 in the Department of Information Engineering and Computer Science of the Feng Chia University are digitalized and analyzed, and trend information mined from them is also presented. The experimental results show that our system improves the overall quality of senior project design course in high education institutions.

**Keywords:** Recommendation system, Senior Project Design, Relational Index, Knowledge Retrieval

# Table of Contents

# List of Figures



# List of Tables

# Chapter 1 Introduction

Senior project design plays an essential role in an engineering educational curriculum. As required by ABET EC2000, students must be prepared for engineering practice through the curriculum to learn a major design experience based on the knowledge and skills required in earlier course work, while incorporating engineering standards and realistic constraints (ABET, 2000). In general, such a major design experience can only be provided in the form of senior project design courses. For computer science programs, besides providing opportunities for senior students to apply what they have learnt by demonstrating the practical applications of system analysis and software engineering tools, senior project design is, to most undergraduate students, built on the experience of previously done projects([1] Chang, C. C. ). From the perspective of program and curriculum, senior project design could also be an indicator for institutional self-assessment ([7] Enderle, J; [4] George, D.; [11] Norton-Meier, L. A.). With properly statistical and analytical work, senior project design trends, thus the education outcome of a program could be mined for evaluation or other purposes ([2] Dow, C. R.; [9] Hales, J. L.).

However, much effort needs to be done in advance before we can get the most out of senior project design. The most essential task is assessment. When large scale designing works and longer learning duration are concerned, traditional assessment tools such as examinations or performance appraisal may not be as effective as they evaluate students' learning outcomes only at certain time intervals ([10] Prus, J.). While starting the course, students tend to spend much time on choosing advisors, exploring project possibilities, and forming topics. They will also be facing technical problems that have been previously solved by former students. Thus comes in the need of knowledge supporting system. If done efforts are preserved and modeled in someway so that they are made "reusable", students could use their time more effectively on creative thinking, thus improving the overall quality of senior project design. According to issues thus said, this work designs and implements a knowledge supporting portal, which enables an easy access to previously established project documents and provides decision support. Students' responses are acquired through usage questionnaire survey. Moreover, projects from 1999 to 2004 in the Department of Information Engineering and Computer Science of the Feng Chia University are

analyzed, and trend information mined is presented. The experimental results show that our system improves the overall quality of senior project design course in high education institutions.

# Chapter 2 Related Work

Computer science programs in higher education trains or drill students' ability through the conducting of a senior-level, semester-long capstone course in which teams of students apply all their knowledge in an integrative fashion, working on a "real" project. In the department of Information Engineering and Computer Science in Feng Chia University there is a "special topic" courses which stands for senior project design. Students are to freely choose an IT related research/application area, such as bioinformatics, web service, network management, micro system, etc, and seek a faculty member with such or similar research interest for mentorship. During the process of the course students perform system analysis, setting goals, master required tools, and finally set out to implementation and evaluation stages.

During the process of senior project courses, the most essential for students is the assistance of information and knowledge retrieval. The retrieving of data and that of knowledge are quite different, but also share some common concerns. While retrieving information, unstructured data such as documents are concerned. IR (Information Retrieval) has drawn much attention recently due to the Internet's popularity. The most basic idea of IR starts with the representation of document. As the Internet grows however, researchers have been trying to discover new methods for information retrieval such as intelligent Web agents, as well as extend DB and data mining techniques, thus creating the term 'Web Mining'. Among the many approaches in Web Mining ([5] Larose, D. T.), web usage mining ([8] Roddick, J. F.) involves the automatic discovery of user access pattern from one or more Web servers. Pieces of information such as the page that are visited most, most users' click history, or even their purchasing habits, are the production of Web Mining, pieces of information that Web managers are extremely interested in knowing so that could better improve their site, or devise commercial strategies.

While it seems that information seems can be found everywhere, facing a great deal of rough and unfamiliar domain-based content, how do student get the useful knowledge for themselves? Now, there are many searching engines, which are like Google, Yahoo, etc. And there are also many on-line bookstores, which are like Amazon, Caves Books and so on. Of course, in the commonwealth of learning, there are also many pages, which are like NCBI (National Center for Biotechnology

Information). Although they are different character, they have the similarity service, but may in differ ways. When the user searches some words in these pages, they will not only give the user information which the user inputted, but also give the user some information which is related to the words that the user inputted. This is called a recommender system, which itself has been a research area in knowledge discovery.

# Chapter 3 System Architecture

System architecture is described in this chapter. The learning model is described in Section 3.1. The overview of system described in Section 3.2. The mechanism of relational index and recommendation are explained in Section 3.3.

## 3.1 Learning Model

As shown in Figure 3.1, the learning model of senior project design is composed of four stages: exploration, design, implementation and presentation. Now the program portfolio is a knowledge supporting system representing past knowledge of senior projects. During the exploration stage, students may have problem deciding which research area to go, or which advisor to find mentorship, etc. Onto design stage where system analyzing was concerned, proper developing tools such as programming language, platform, or hardware has to be decided. In above two stages, trend information is especially nectar for quick solution, since students usually do not have a clue of where to start. In implementation stage, numerous technical problems will be encountered, and most of them are usually solved already by previous students, and which can be solved again if past experience were offered from the program portfolio.

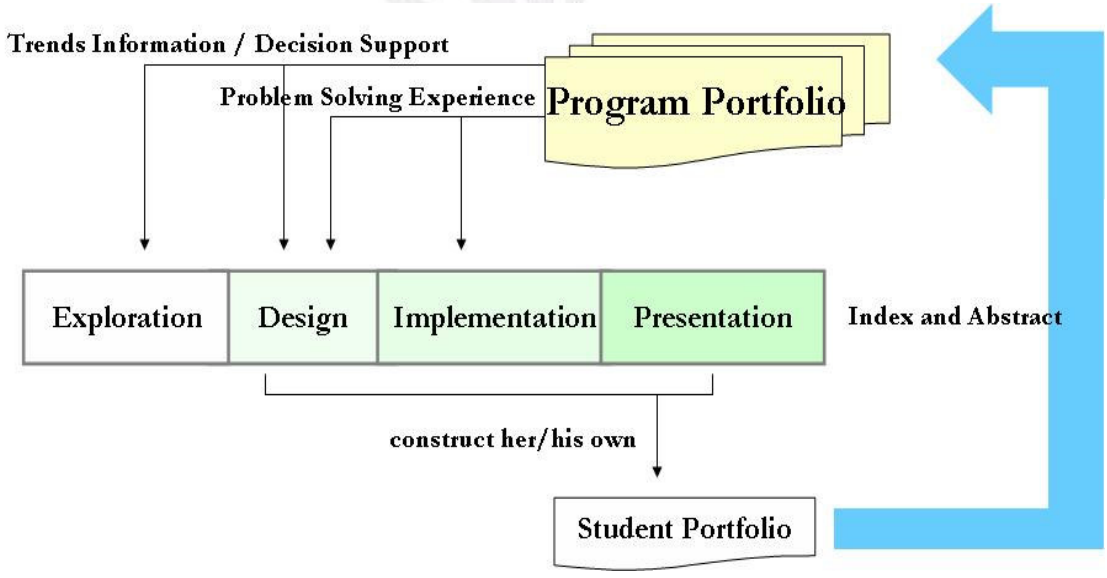


Figure 3.1 Learning Model

During the process, students accumulate and collect their own learning materials through this system to make their own personal senior project, which will be presented in presentation stage, and regarded as final product. Then these learning materials, thus ‘student portfolio’, will be under some kind of abstraction and

indexing, then goes into the program portfolio pool to provide new knowledge for future students.

## 3.2 Overview of System

The proposed system architecture of a knowledge support system is shown in Figure 3.2, which is composed of several main parts including retrieving system, recommender system, and trend analysis, with relation mining as supporting database preprocessor. Users interact with the knowledge supporting system either through a Web interface, or self-developed client application.

The most basic function of the system is document retrieval. Such information is prepared in the database, and users could directly access them through a keyword search or browse through classified category.

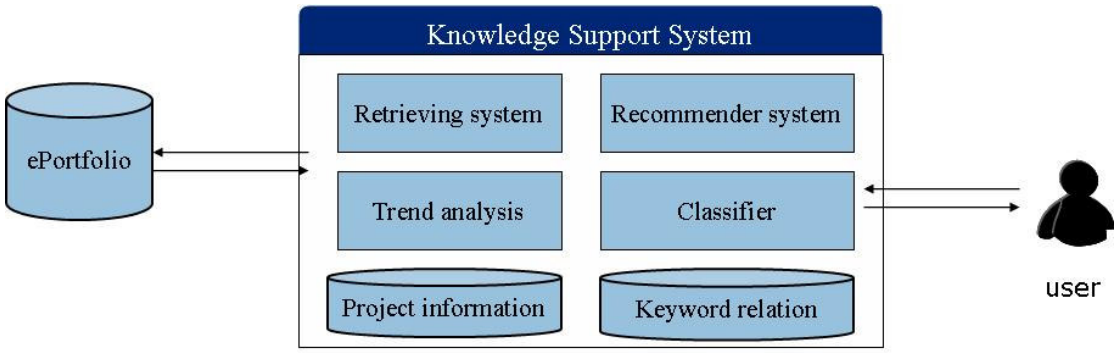


Figure 3.2 System Architecture

Project database is built with the indexes, abstracts and keywords are drawn out from project reports, which could be used later as features for query. When system receives input submitted from user, the system will analyze user's input and will transfer the input into database query to search the materials in the database. We use the way of indistinct compare to query, so if projects which have some keywords or abstracts is about the input will be regarded as the relational projects. The result will be a lot of disorderly project data, it will be sorted by the level of relation, and the more relational projects can be seen easily. The ordered data will be packed into html and be paged, then presented to user.

Recommender system will help user do the further inquiry by recommending keywords related to the projects of search result. So how to find the relational keywords with input is the problem. In order to let keywords have relation with each other, the relational index is designed. The relational index is an important part in our database, which is built by the arrangement of keywords from each project. So if we

use a keyword to query the index, the result will be the words related to the input word. Every times when user searches the projects, the system not only retrieves relational projects, but also will use user's input to transfer into database query to search the relational index, and then the result will be the recommendable keywords.

Trend analysis can help students to definitely understand the research trend with analytic information. Projects information can be analyzed to provide an OLAP, like decision support, which have effective help for exploration stage and design stage in leaning model. In our system, user has many statistical ways to select. System can recognize which way user selected is, and then use that to transfer into different database query to search the database. The result materials will be graphed into a statistical chart, and then presented to the user, that user can easily analyze research trend by the statistical chart.

## 3.3 Relational Index / Recommendation

We use the arrangement of keywords to make up the relational index. As shown in Figure 3.3, suppose Dn is a project with 3 keywords, A, B and D, then use the arranged way: (A,B), (B,D), and(A,D) to record into the database, and give each set initial value 1. Then use this way to add each keyword set of all projects into relational index. If we find the same set, then its value will be incremented.

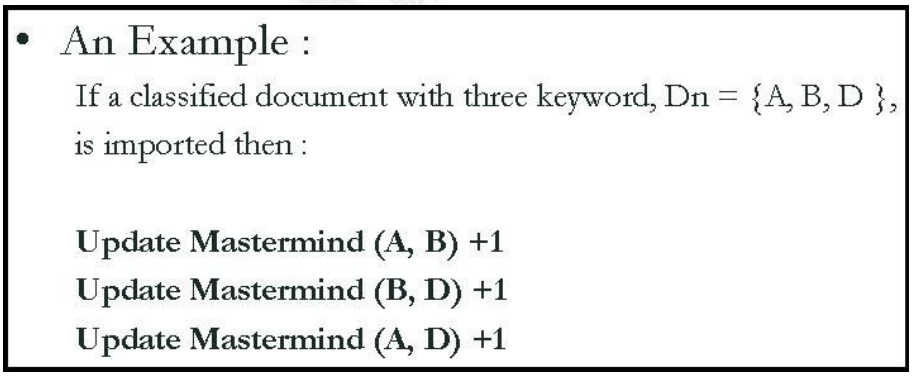


Figure 3.3 Example of Find Relation

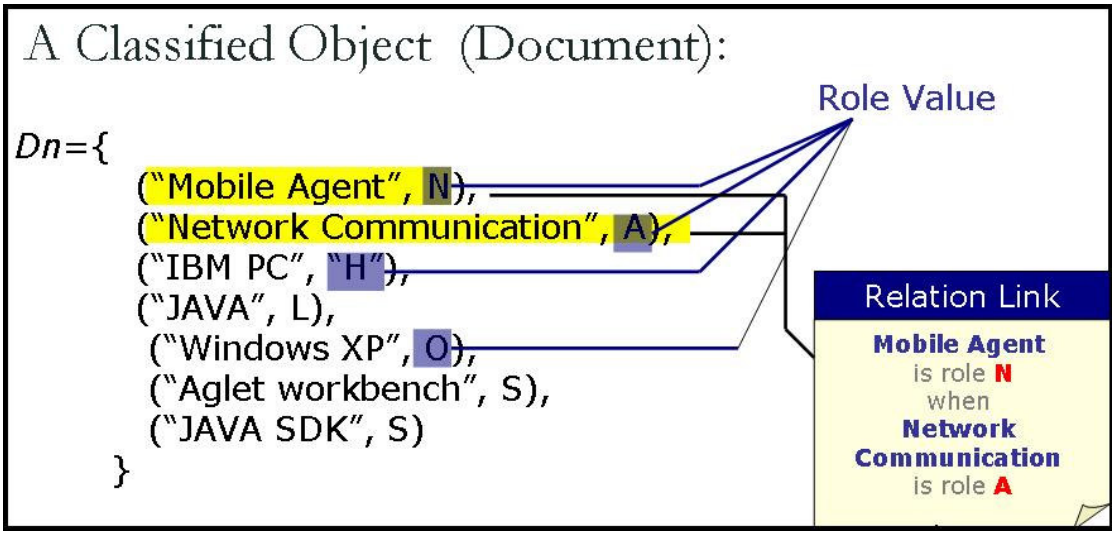


Figure 3.4 Example of Classification

On the other hand, we will classify keywords with its functional property before recording arrangements. By using the way, we can see the relation between fields. Suppose a keyword set as shown in

Figure 3.4, we will use different code to represent the property of each keyword. For example, Mobile Agent is N(Other), Network Communication is A(Area), Windows XP is O(OS). And the appropriate code will be recorded together with keywords when making up the relational index.

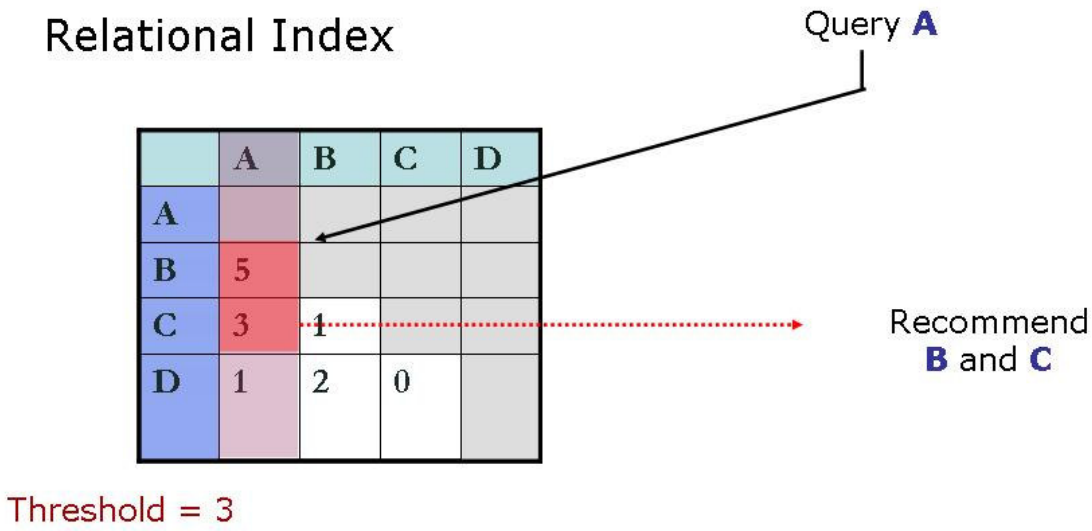


Figure 3.5 Example of Recommendation

As shown in

Figure 3.5, in order to find out the relation keywords of A, we only need use A to query, and set up a threshold value (in the figure is 3), so that we can find relation keywords of A to recommend (in the figure is B & C).

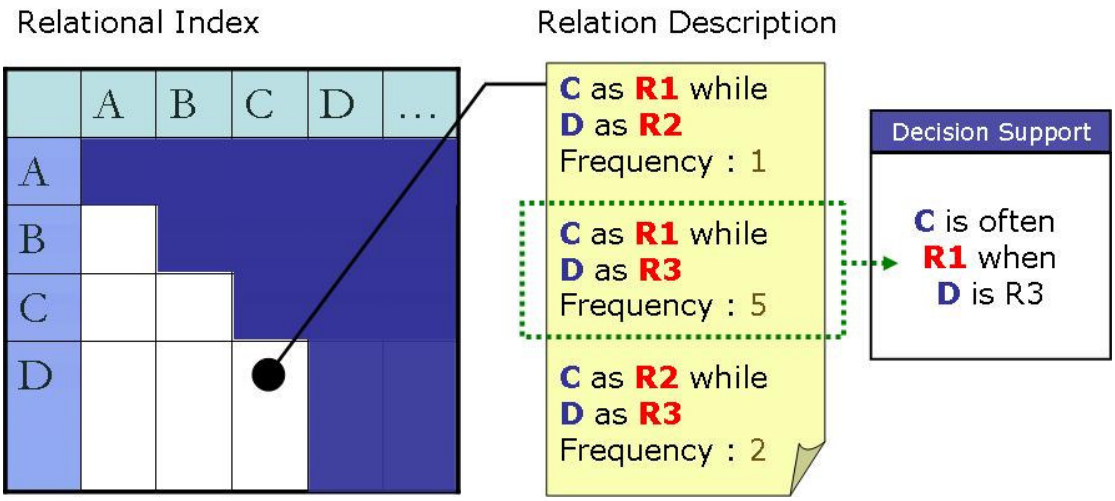


Figure 3.6 Relation Query

As shown in Figure 3.6, the same keyword set may have more than one relation in different fields. But we can find the most relational filed by relational index classified. In the figure, we use C to query, and we can find the relation of C & D in different fields. Then we can use frequency value to see the set of C & D has the strongest relation in other field, therefore we will recommend D to user in other field.

# Chapter 4 System Implementation and Prototype

According to the architecture, knowledge support system for senior project design was implementation and described in this chapter.

## 4.1 System Implementation

The subject of our implement is Dept of IECS (Department of Information Engineering and Computer Science) of Feng Chia University. Senior project design in the department has been going for more than thirty years. Reports were stored in PDF files since 1998, which would be the raw data for our knowledge system. 350 projects were classified into 13 research areas including Web applications, information security, multimedia, image processing, embedded systems, communication network, database, digital signal, distance education, algorithms, electronic commerce, distributed processing and artificial intelligence, and patched with proper keywords according to technologies involved, hardware/software, development platform, etc. Finally, 32 teachers were associated with these projects, thus forming our database. And according these source data, we built our relational index links.

The scope of our implementation is exactly as proposed earlier, with a knowledge support system for knowledge retrieval, which is named the Corpus. Both were realized in ASP.NET applications in a Web Server with Microsoft Internet Information Service and .NET Framework, in combination with MySQL database system. All sub-applications are presented in active web pages; enabling users to access the system from any place without the need to install any client software. The whole system is constructed under MyIECS, the departmental Web portal, thus taking advantage of its authentication system.

As shown in Table 4.1, our system was composed of these tools.

| Tool Options | Tool Item |
|---|---|
| Operating System | Microsoft Windows XP |
| Develop Platform | Microsoft .NET Framework |
| Develop Environment | Microsoft Visual Studio .NET 2003 |
| Web Server | Microsoft Internet Information Service (IIS) |
| Database | Microsoft Office Access |
| Program Language | ASP.NET / C# |

Table 4.1 System Tools

And according to system architecture, we described a system flow chart as shown Figure 4.1. This flow chart was described how project was inserted into our database, how system was done recommended, and how user got information.

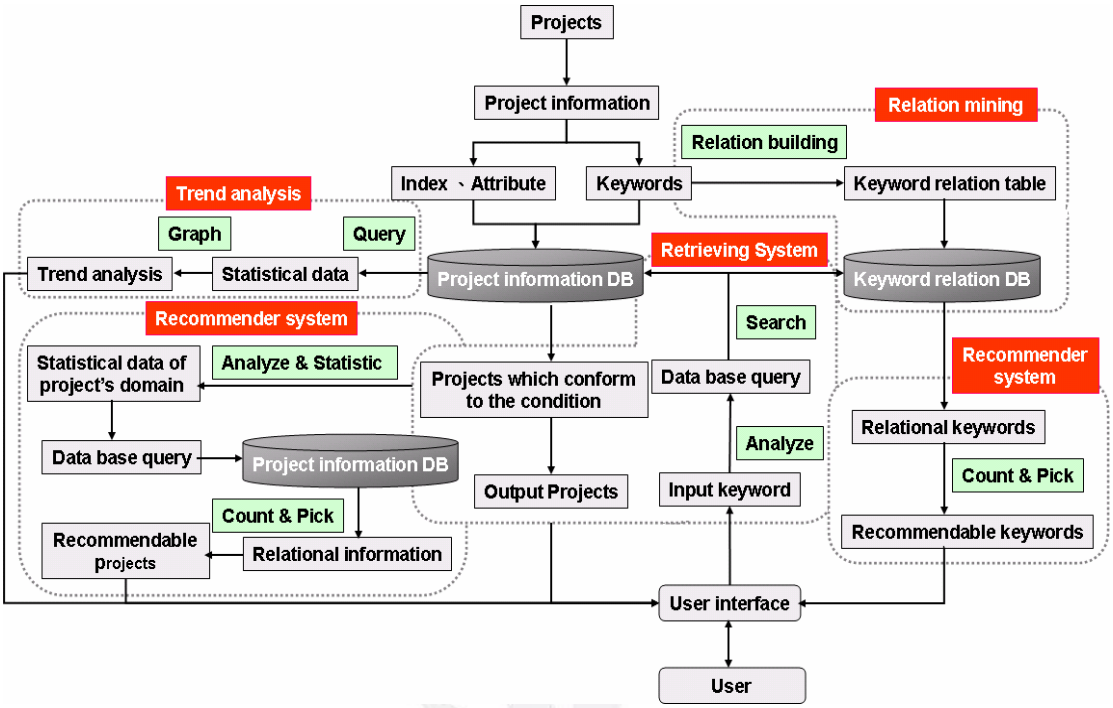


Figure 4.1 System Flow Chart

## 4.2 System Prototype


Figure 4.2 Corpus Home Page

Because Corpus is a support system, searching was played an important part. There are two search systems in Corpus. As shown in Figure 4.2, one of them is

simple search system which is keyed in words by the user in the homepage. The other is advanced search system which is chose research area, school year or teacher, or keyed in keywords by the user, and shown in Figure 4.3.


Figure 4.3 Advanced Search Page

No matter simple search system or advanced search system, Corpus will give result as shown in Figure 4.4.

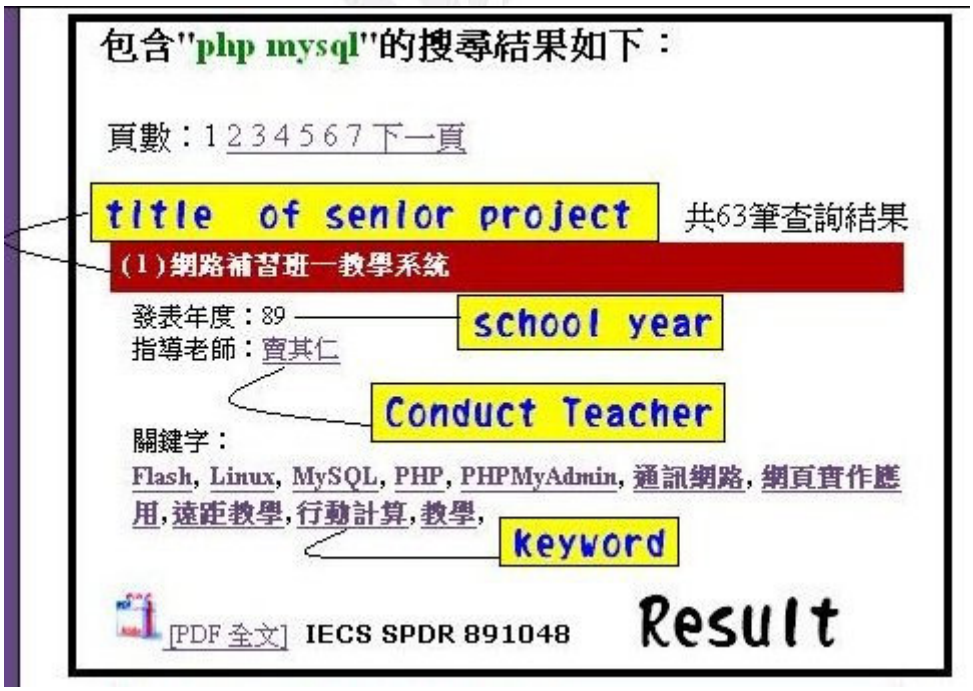


Figure 4.4 Result

And according to keyword of the user input, result page is shown title of senior project, school year, teacher, keywords, number of senior project and PDF connected.

More detailed functions are provided in advanced search. In simple search system, as long as any keywords from database accord with the words that the user inputs, result page will show up. For example, when the user key in “mobile”, and then result page is shown up result is like “mobile” or “mobile agent.” But in advanced search system, all keywords from database must match with inputted. For example, when the user chose “communication network” and “Dow, C. R.,” and then result page is shown up result including “communication network” and “Dow, C. R.”

As shown in Figure 4.5, Corpus collects times that keyword is queried by the user, and then shows top ten keywords in the homepage, which was named “Hot Search.”


Figure 4.5 Hot Search

As shown in Figure 4.6, according to the user inputs keywords, Corpus will show up the related senior project papers, teachers, technologies, software, languages and operating systems.

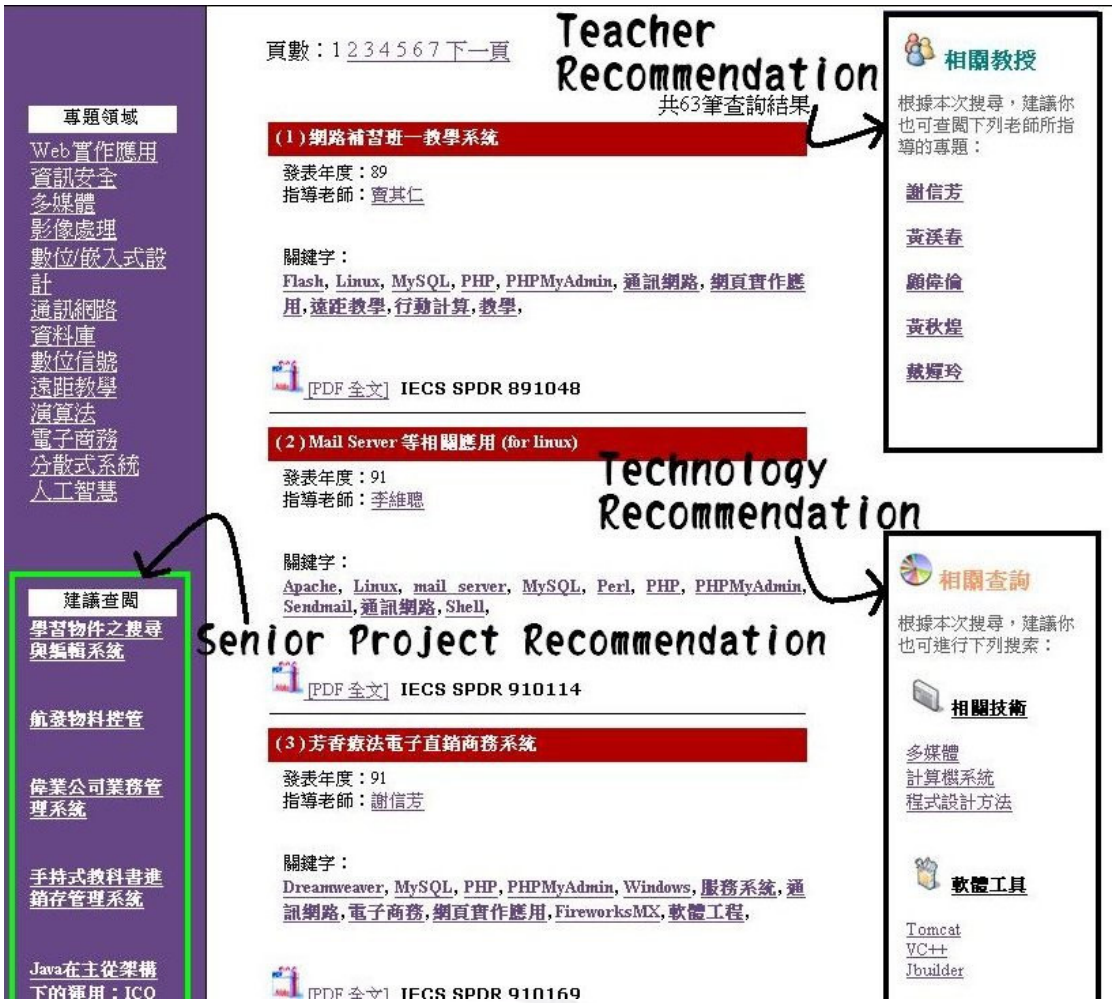


Figure 4.6 Recommendation

In the statistics page, there are “Research Area Statistics Data” and “Teacher Research Area Statistics Data” now. As shown in, school year has been selected to read by the user. Corpus will show the data that all students do senior project paper from 1999 to 2003. As shown in, school year and teacher have been selected to read by the user. Corpus will show the data that teacher conducts students who do research area of senior project paper.


Figure 4.7 Research Area Statistics


Figure 4.8 Teacher Research Area Statistics

# Chapter 5 Experiment Results

The experimental results are presented in this section, including trend analysis and usage analysis.

## 5.1 Trend Analysis

In this section, the data in Corpus are summarized into graphs to show learning trends concerning senior projects, including research area, dynamic web page languages and database system. Of course, if needed, other analysis could be performed.

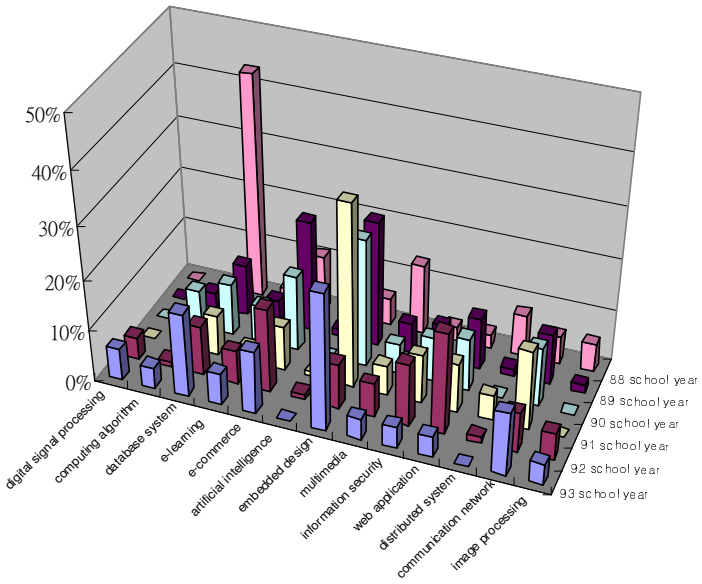


Figure 5.1 Research Area Analysis

As shown in Figure 5.1, we categories the senior projects in the past five years into 13 Research areas as we mentioned above, including artificial intelligence, multimedia, distributed system etc. The analysis shows there are three research areas which are getting popular. The most popular is embedded design, which has been increasing since 89th school year, and even reached 37 papers in 91st school year. And it's just the time the experiment target, the Dept of IECS of Feng Chia University, set more programs for embedded design. It shows that students are interested in what they learn and eager to put their efforts on modeling embedded design. Other popular fields are communication network and web application. These two fields are also the trend of the industrial needs. It means when students choose the direction of senior project,

they do think about the future development. Here leaves some issue we can go further to explore, like the relation between conducting teachers' professional specialty and the trend of Research areas.

According to the trend analysis of research areas, we observed that web application is a popular trend. Web application often focuses on the interaction between web pages and database systems. Therefore we analyze the usage of dynamic web page language and the database system from the projects of web application. See the Figure 5.2; we observe that the most popular language students choose to program web pages is ASP. Although PHP is free and also popular, it is not as convenient as ASP. When it comes to database systems, as shown in Figure 5.3, the most popular is still Access because of its simplicity. Thus, we know that student would prefer to choose the interface they familiar with and feel easy to use.

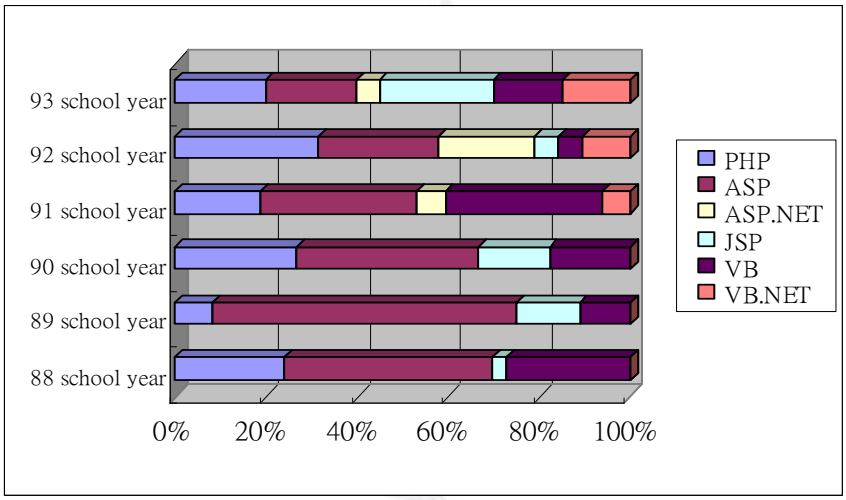


Figure 5.2 Dynamic Web Page Language

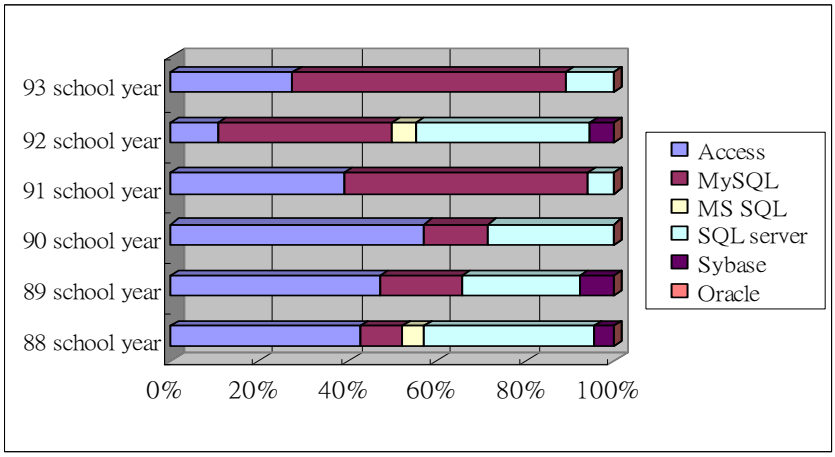


Figure 5.3 Database System

## 5.2 Usage Analysis

By collecting user feedback about access log and satisfaction questionnaire survey, we analyze the frequency user use the knowledge system to do their senior project and how they think about whether the system can help them.

The system opens for students of the department since May 23rd in 2005. By collecting the user access log during May 23rd and June 8th and analyzing the log, as shown in Figure 5.4, we observed that students use this knowledge support system eagerly during weekdays (drawn in light color). It shows students work harder during weekdays so that they can take a break on weekend. Prior to semester due, when students have to hand in their gradation paper, the access times increased sharply. Thus we can be sure that they are willing to seek for help from the knowledge support system.

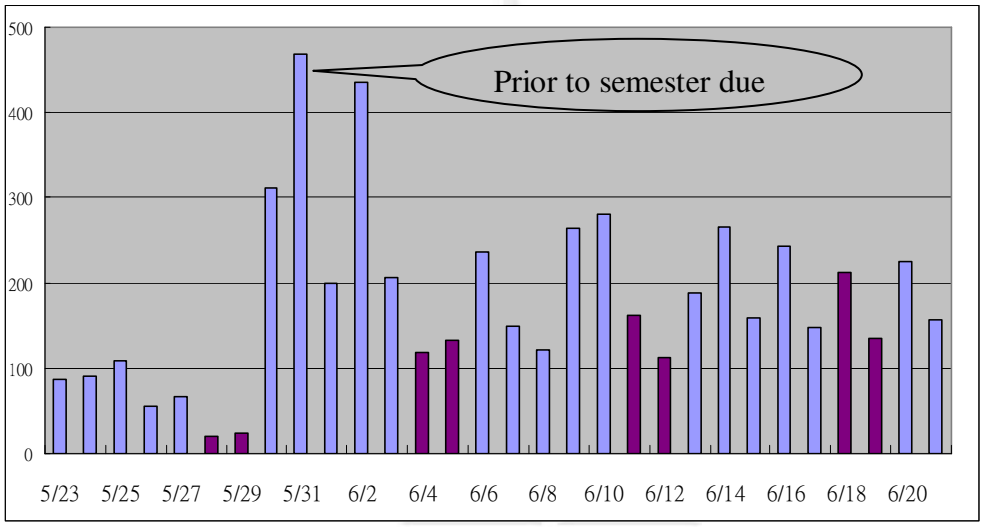


Figure 5.4 User Access Log Analysis

We conducted a web questionnaire survey. As mentioned above, the questions in the survey include whether the system helps with decision problem, like technique choosing or language choosing, when produce the senior projects, and whether users satisfied with the recommend function supported etc. Analyzing the result of satisfaction questionnaire survey, as shown in Figure 5.5, the average is between four and five degree. It tells, generally, students satisfied with this knowledge support system on the user interface and the information they can get. We can see that hot search function helps students a lot when they have any clues to find information. And according to the answer to questions about whether the system is encouraging for users to work harder and whether recommend support them inspiration, we know that,

the most important, students can get inspiration from this system when studying senior project.

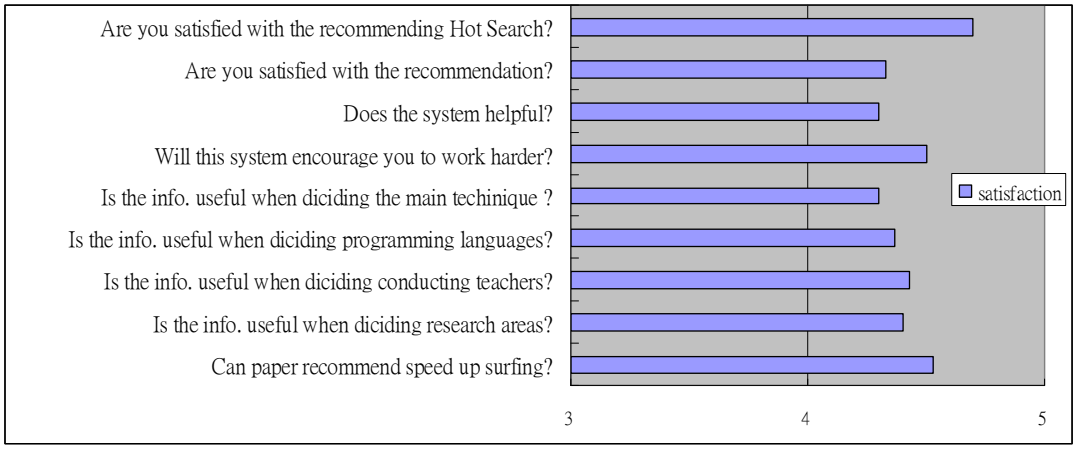


Figure 5.5 Questionnaire Survey Result

# Chapter 6 Conclusions

## 6.1 Difficulty and Solution

Difficulty 1: How to build up the prototype and set up type of webpage? And how to make the result of searching?

Solution 1: At first, we spent much time on discussion which is like what data did we must insert our database. And then, we began to design prototype of webpage. We did the simple search in the homepage, and let it show result in the result page. Because we did not have experience of using ASP.net, we did result page for a span. But as long as the result was written, other page could follow this way. And when showing results, we also discussed what data we must show up to the user. When we did above, our prototype of webpage was finished.

Difficulty 2: How did we design our recommend system at the beginning?

Solution 2: At the beginning, we have no idea to design our recommend system. So we started to collect materials from some searching websites, like Google, Yahoo, and Amazon. Then we found these websites have the similar characteristics, they can use users' input to recommend relational materials are related to the input. So we started to discuss how to let our implemental projects have relation with each other in database, then we finally designed the relational index to help us to find the relation between each keywords.

Difficulty 3: What is our policy to search the query string that user inputs? How does the query string be split into meaningful tokens?

Solution 3: We cut the query string when meeting a space character and combined those tokens into different meaningful string. And after getting the combined string to search, we search them in the database successively. Every searching result will response a value which stands for how much the result fits in with user's expectation. And the system will show the result in order of its value. Thus, users can get the effective result intuitively.

## 6.2 Summary

This work designs and implements a knowledge supporting system, which enables an easy access to previously established project documents and provides decision support. Students' responses are acquired through usage questionnaire survey. Senior projects from 1998~2004 in the Department of Information Engineering and Computer Science of the Feng Chia University are digitalized and analyzed, and trend information mined from them is also presented. The experimental results show that our system improves the overall quality of senior project design course in high education institutions.

# Chapter 7 Feedbacks

## Y. P. Huang

Tree semester past, we finally accomplish this work. By this course, I realized that communication is very important in teamwork. How to allot the work, how to make work in progress, and how to discuss with partners all need the communication. Without communication, we cannot work well but delay and quarrel. I also learned that good time-management help me get better results. Students are busy all the time, but also free all the time. It just depends on how we use time. After the senior project course, I got some very valuable experience.

## Y. S. Wang

By doing research projects in the three terms, I think what I learn most is the importance of teamwork. Under the leading of teaching assistant, we will have regular meeting each week. Except check the progress report, we also use the way of brainstorming to discuss the new function or something needed to improve. When decide the next target, we will assign those work to every member by their ability, then we can effectively accomplish our work. By doing research projects, we experience a very good teamwork model. We learn how to effectively work with our partner, and how to communicate with each other. I think the experience is very precious wherever I go in the further.

## J. A. Tu

At first, I felt that senior project design was very hard and difficult. Because I had no idea what I could do. When I found out classmates who worked together in doing senior project design, we discussed what we could do. Finally, we decided that we did senior project design with C. R. Dow. Our team had much difficulty during doing senior project, but we discussed with each other, solved problems together and K. H. Chen who is a candidate for doctor's degree in Dept of IECS of FCU would help us. Let our senior project could finish without blemish. In this senior project design, I not only get ability but also understand how to team work. This is my best results.

# Acknowledgement

There are too many people to thanks for their help through the senior project design. This senior project would never be accomplished without their help. First of all, we would like to thank our advisor, Prof. C. R. Dow, who gives us good advice and support in many ways. Studying with him has been a pleasure and a great learning experience. We would also like to thank K. H. Chen, who is Prof. C. R. Dow's a candidate for doctor's degree. He gives us much help and advice.


# References


[1] Chang, C. C. (2002), Building A Web-based Learning Portfolio for Authentic Assessment, Proceedings of the International Conference on Computers in Education, Vol. 1, pp. 129-136.

[2] Dow, C. R., Chen, K. H., Pong, C. Y., Lu, C. P. and Lin, W. C. (2004), Trends Analysis of Senior Project Design Using an ePortfolio System, Proceedings of the Conference on Information Management and Implementation.

[3] Dow, C. R., Lee, W. B. and Chen, K. H. (2004) SPDP: A Versatile Senior Project Design Portfolio System, Proceedings of the First International Conference on Teaching and learning in Higher Education, pp.449-454.

[4] George, D. et al. (2002) How Electronic Portfolios Add Coherence to Educational Programs, Proceedings of the International Conference on Computers in Education (ICCE'02), Vol. 2, pp.1406-1407.

[5] Larose, D. T. (2005), Discovering Knowledge in Data, Connecticut: Wiley-Interscience.

[6] Engineering Criteria 2000: Criteria for Accrediting Engineering Programs, Accreditation Board of Engineering and Technology, available at http://www.abet.org.

[7] Enderle, J. et al. (2003), "The ABCs of Preparing for ABET," IEEE Engineering in Medicine and Biology Magazine, 22(4). pp. 122-132.

[8] Roddick, J. F. and Spillopoulou, M. (2002), A Survey of Temporal Knowledge Discovery Paradigms and Methods, IEEE Transactions on Knowledge and Data Engineering, 14(4), pp.750-767.

[9] Hales, J. L. and Pisarski, S. J. (1998), Evolution of a Capstone Course: a Twenty-year Perspective, Proceedings of the Frontiers in Education Conference, Vol. 1, pp. 456-459.

[10] Prus, J. and Johnson, R. (1994) Assessment & Testing Myths and Realities, New Directions for Community Colleges, 88(94), pp. 5-94.

[11] Norton-Meier, L. A. (2003) To Efoliate or Not To Efoliate? The Rise of the Electronic Portfolio in Teacher Education, Journal of Adolescent & Adult Literacy, 46(6), pp. 516-518.

[12] Upchurch, R. L. and Sims-Knight, J. E. (2002), Portfolio Use in Software Engineering Education: An Experience Report, Proceedings of the 32nd Frontiers in Education Conference, Vol. 3, pp. S2G-1 – S2G6, 2002.