APPLICATION FOR FACULTY RESEARCH GRANT

**All questions must be completed to be considered for grant award.

Choose one:	
[ ] Creative [x] Research	**Date of Last FRG Award** (Semester and Year awarded): __Fall 2004__
	Date of ATU Faculty Appointment (Semester and Year): __Fall 2005__

1. Project Title: ___Visual Data Representation and Information Efficiency in Knowledge Discovery___ .

2. Name of Principal Investigator/Project Director:_____Dr. Roger Fang_____

3. School (abbrev): ____Sys Sci_____   **4. Department:** __Computer and Information Science____

5. Campus Mail Address: _Corley 261__ **6. PI/PD Campus Phone:** __498-6082_____

7. Amount Requested: $__2,000_____   **8. Total Cost of Project:** $__2,000_____

9. Does this project involve: **10. Duration of Project:** __4/01/06 – 1/01/07_____

Yes No
[ ] [X] human subjects?
[ ] [X] animals/animal care facility?
[ ] [X] radioactive materials?
[ ] [X] hazardous materials?
[ ] [X] biological agents or toxins restricted by the USA Patriot Act?
[ ] [X] copyright or patent potential?
[ ] [X] utilization of space **not** currently available to the PI/PD?
[ ] [X] the purchase of equipment/instrumentation/software currently **available** to the PI/PD?

NOTE: If the answer is "yes" to any of the above questions, the investigator must attach appropriate documentation of approval or justification for use/purchase.

SIGNATURES

Department Contribution (if applicable): $_____

Account Number: _____

School Contribution (if applicable): $_____

Account Number: _____

Jany Moull _3/14/06_
Chairperson   Date

John West _3-14-06_
Dean   Date

This Section to be completed by the Office of Academic Affairs

FSBA Committee Award Recommendation: Yes____ No____
FSBA Committee Proposal Rank: _____ of _____ Total Proposals.
Recommendation of VPAA: Yes____ No _____
Recommendation of President: Yes____ No _____
Award Date: _____

ABSTRACT

The booming and widespread IT applications that provide users data warehousing and data mining functions have made statistical and analytical data that are crucial to business operations more available than any time in the history. Commercial software tools currently in the marketplace have dramatically made this type of data process even easier to implement and control. However, in many of the tools and packages, we also found that an important issue regarding the effectiveness of user access of a system is mostly left either unsolved or overlooked. Lacking effective data access, an information system, including all data warehouse systems, is like a computer keyboard with all keypads being shuffled randomly that no one can easily and quickly make a correct input, which results in poor *information efficiency* to all applications in such a computer.

During our previous project we found that possible reasons of why user's data access is not well addressed and supported by these commercial products are twofold: (1) The complexity of the underlying multi-dimensional data structures of a data warehouse has hindered the developers' attempts for a user-friendly interface; and (2) the stiff learning curve of applying the data warehousing query protocols has prevented many users from direct access to the data warehouse.

As a new extension to our ongoing research "**Knowledge Discovery in a University Data Warehouse,**" which is a FRG-funded project approved in fall 2004, this project proposes to investigate and resolve issues related to the information efficiency of a data warehousing application.

PURPOSE/OBJECTIVES

The purpose of this project is to extend our previous work funded by FRG in fall 2004 to study and develop an effective multi-dimensional data access interface with certain

1

visual representation technologies [1] and, therefore, to improve the information efficiency of a data warehouse system. To achieve this, we will extend our research to

- Conduct a survey of data access and representation technologies that are currently available in software tools and packages of the marketplace. New methods will possibly be developed to enhance the existing technologies.

- Integrate a data access and representation component with our in-house data warehouse system. This data warehouse has been loaded with the student transcripts and course scheduling data of Arkansas Tech University for the past two decades (1985 ~ 2005). *For privacy and security reasons, no social security numbers and student names are included in this data warehouse.*

- Implement a user interface prototype with visual representation capability.

SIGNIFICANCE/NEED

Effective data access is considered one of the key factors that determine the *efficiency of information* of using a data warehouse. A data warehouse with high information efficiency means users can (1) enter the data or specify the queries easily and precisely without wasting time for trial-and-errors; and (2) understand the returned data or information quickly and correctly without cumbersome interpretation. Both types of data access are especially important to the OLAP (online analytic processing) users who are typically the executive decision makers of an organization. Misleading of these people due to poor data access of an OLAP system can cause catastrophic business problems or loses and, therefore, has been long time recognized to be a high operational risk to an organization. This has also motivated more and more academic studies [2] and industrial implementation on incorporating business intelligence in an organization. In light of this emerging information technology trend, this

research is aimed at devising an effective data access as a measure to improve the overall information efficiency of a data warehouse system.

PROCESS FOR ATTAINMENT OF OBJECTIVES/GOALS

By following the results of our previous knowledge discovering project, this research will be conducted in two extended stages: (1) data preparation and query design, (2) visual representation prototyping. During the preparation and design stage, possible data access methods will be studied and identified based on the input and output characteristics of the underlying data and queries. Efficiency of different methods and formats will be compared. Selected optimal methods and formats for final data representation of the data warehouse will be implemented in the visual representation prototyping stage. A tentative schedule for completing the proposed research activities is sketched below.
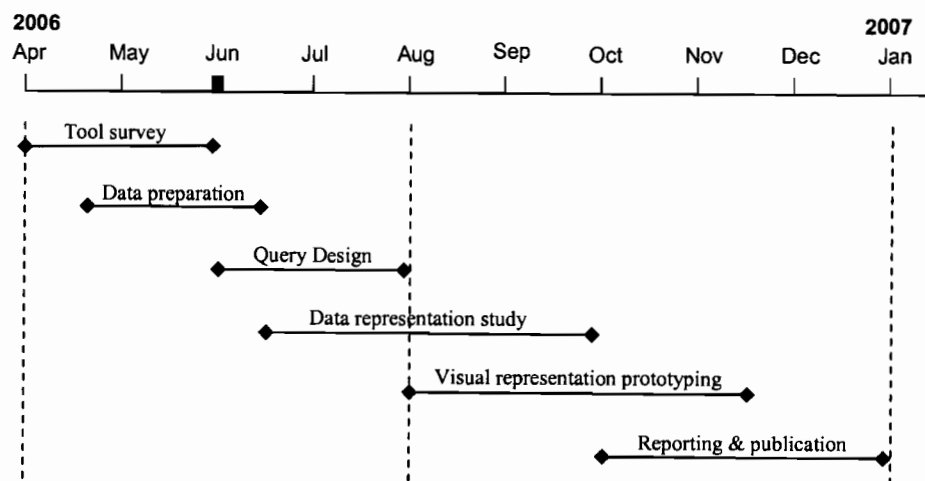
DISSEMINATION OF RESULTS

The result and experience of the proposed research work will be fully implemented on an in-house data warehouse system to provide OLAP functions to interested users of various offices and departments. We also plan to use the completed system as a case study and demo project in several courses for our BSIS, BSIT, and MSIT majors. This will

3

definitely enrich and strengthen our IS/IT curriculum of the department. Finally, we believe this research will provide a good opportunity for not only encouraging a larger scale of project to allow more faculty members of the department to work together, but also exposing positively the department and the university to other schools and industry through publications and conference presentations.

BUDGET   (See the attached sheet on next page.)

BIBLIOGRAPHY

[1] Marakas, G. M. Modern Data Warehousing, Mining, and Visualization. Prentice Hall, 2002.
[2] Keim, D., Müller, W., Heidrun, S., Visual Data Mining, 2002
     http://wwwicg.informatik.uni-rostock.de/~schumann/papers/2002+/VisDM_Star.pdf

APPLICATION VITA

The PI/PD of this project has taught courses in databases, data warehousing, and information systems fields for several years since 1999 in British Columbia Institute of Technology (BCIT, Canada), Kingston College (Canada), and Arkansas Tech University (starting 2001). Before teaching, the author of this proposal worked for a governmental research lab in Taiwan, senior technical consultant for Rational Software, and director of technology of Infocomm International. The author has been granted the Faculty Research Grants with an amount of $1,800 in November 2004 for a research in this area at Arkansas Tech University. A paper that describes the project results is currently in progress and will be submitted in summer 2006 for a conference presentation.

PROPOSED BUDGET
FACULTY RESEARCH GRANT

1. Graduate assistant stipend $ 400
   Fringe benefits @ .4% (4/10 percent) of graduate assistant stipend

2. Non-work study stipend $ 0
   Fringe benefits @ .4% (4/10 percent) of non-work study stipend

3. *Supplies (please list items to be purchased and estimated price
   per item including taxes and shipping, if appropriate):

Item No. 1 publication purchase	Estimated Price	$ 400
Item No. 2 computer software	Estimated Price	$ 1,200

            Total estimated supplies $ 1,600

4. Travel (please list travel expenditures by date
   and estimated costs):

            Total estimated travel $ 0

5. *Capital Outlay (please list items to be purchased and estimated
   price per item including taxes and shipping, if appropriate):

            Total estimated capital outlay $ 0

            TOTAL PROPOSED BUDGET $ 2,000

*Items purchased under $2,500 (including taxes and shipping) are considered supply items. Capital Outlay items are those which cost $2,500 or more (including taxes and shipping). Please contact the Purchasing Office for questionable items.