HARDWARE,
SOFTWARE
AND PEOPLEWARE

SAICSIT 2001

Edited by
Karen Renaud
Paula Kotzé
Andries Barnard
HARDWARE, SOFTWARE AND PEOPLEWARE

South African Institute of Computer Scientists and Information Technologists
Annual Conference
25 – 28 September 2001
Pretoria, South Africa

SAICSIT 2001

Edited by Karen Renaud, Paula Kotzé & Andries Barnard
University of South Africa, Pretoria
Table of Contents

Message from the SAICSIT President ................................................ iv
Message from the Chairs ................................................................... vi
Conference Organisation .................................................................. vii
Referees .......................................................................................... viii

Keynote Speakers

Cyber-economies and the Real World ................................................ xi
Alan Dix
Computer-aided Instruction with Emphasis on Language Learning .................. xiv
Lut Baten
Internet and Security Trends............................................................... xv
Arthur Goldstuck
The Future of Data Compression in E-technology .................................. xvi
Nigel Horspool
Strategic Planning for E-Commerce Systems: Towards an Inspirational Focus .... xvii
Raymond Hackney

Research Papers

Human-Computer Interaction / Virtual Reality
The Development of a User Classification Model for a Multi-cultural Society .......... 1
M Streicher, J Wesson & A Calitz
Real-Time Facial Animation for Virtual Characters ........................................ 11
D Burford & E Blake
The Effects of Avatars on Co-presence in a Collaborative Virtual Environment .... 19
J Casanueva & E Blake

Education
Structured Mapping of Digital Learning Systems ..................................... 29
E Cloete & L Miller

Formal Methods
The specification of a multi-level marketing business .................................... 35
A van der Poll & P Kotzé
Finite state computational morphology - the case of the Zulu noun .................. 45
L Pretorius & S Bosch
Combining context provisions with graph grammar rewriting rules: the
three-dimensional case ....................................................................... 54
A Barnard & E Ehlers

Human-Computer Interaction / Web Usability
Web Site Readability and Navigation Techniques: An Empirical Study .............. 64
P Licker, R Anderson, C Macintosh & A van Kets
Jiminy: Helping Users to Remember Their Passwords .................................. 73
K Renaud & E Smith

Information Security
Computer Security: Hacking Tendencies, Criteria and Solutions ....................... 81
M Botha & R von Solms
An access control architecture for XML documents in workflow environments .... 88
R Botha & J Eloff
Graphics and Ethics
Model-based Segmentation of CT Images ................................................................. 96
O Marte & P Marais
Towards Teaching Computer Ethics........................................................................ 102
C de Ridder, L Pretorius & A Barnard

Human-Computer Interaction / Mobile Devices
Ubiquitous Computing and Cellular Handset Interfaces – are menus the best way
forward? ................................................................................................................... 111
G Marsden & M Jones
A Comparison of the Interface Effect on the Use of Mobile Devices...................... 120
J Franken, A Stander, Z Booley, Z Isaacs & R Rose
The Effect of Colour, Luminance, Contrast, Icons, Forgiveness and Closure on
ATM Interface Efficiency ...................................................................................... 129
A Stander, P van der Zee, & Y Wang

Object Orientation
JavaCloak - Considering the Limitations of Proxies for Facilitating Java
Runtime Specialisation .......................................................................................... 139
K Renaud

Hardware
Hierarchical Level of Detail Optimization for Constant Frame Rate Rendering........ 147
S Nirenstein, E Blake, S Windberg & A Mason
A Proposal for Dynamic Access Lists for TCP/IP Packet Filtering ......................... 156
S Hazelhurst

Information Systems
The Use of Technology to Support Group Decision-Making in South Africa ........ 165
J Nash, D Gwilt, A Ludwig & K Shaw
Creating high Performance I.S. Teams ................................................................... 172
D C Smith, M Becker, J Burns-Howell & J Kyriakides
Issues Affecting the Adoption of Data Mining in South Africa ............................... 182
M Hart, E Barker-Goldie, K Davies & A Theron

Information Systems / Management
Knowledge management: do we do what we preach? ...................................... 191
M Handzic, C Van Toorn, & P Parkin
Information Systems Strategic Planning and IS Function Performance:
An Empirical Study ............................................................................................. 197
J Cohen

Formal Methods
Implication in three-valued logics of partial information ..................................... 207
A Britz
Optimal Multi-splitting of Numeric value ranges for Decision Tree Induction .... 212
P Lutu
Abstracts of Electronic Papers

Lessons learnt from an action research project running groupwork activities on the Internet: Lecturers' experiences .............................................................. 221
   T Thomas & S Brown

A conceptual model for tracking a learners' progress in an outcomes-based environment
   R Harmse & T Thomas

Introductory IT at a Tertiary Level – Is ICDL the Answer? ........................................ 222
   C Dixie & J Wesson

Formal usability testing – Informing design .............................................................. 222
   D van Greunen & J Wesson

Effectively Exploiting Server Log Information for Large Scale Web Sites ................ 223
   B Wong & G Marsden

   E von Solms & J Eloff

A Pattern Architecture, Using patterns to define an overall systems architecture .............................................................. 224
   J van Zyl & A Walker

Real-time performance of OPC .............................................................. 224
   S Kew, & B Dwolatzky

The Case for a Multiprocessor on a Die: MoaD ..................................................... 225
   P Machanick

Further Cache and TLB Investigation of the RAMpage Memory Hierarchy ................ 225
   P Machanick & Z Patel

The Influence of Facilitation in a Group Decision Support Systems Environment ........ 226
   T Nepal & D Petkov

Managing the operational implications of Information Systems .................................. 226
   B Potgieter

Finding Adjacencies in Non-Overlapping Polygons .................................................. 226
   J Adler, GD Christelis, JA Deneys, GD Konidaris, G Lewis, AG Lipson, RL Phillips,
   DK Scott-Dawkins, DA Shell, BV Strydom, WM Trakman & LD Van Gool
The South African Institute of Computer Scientists and Information Technologists (SAICSIT) was formed in 1982 and focuses on research and development in all fields of computing and information technology in South Africa. Now in the 20th year of its existence, SAICSIT has come of age, and through its flagship series of annual conferences provides a showcase of not only the best research from the Southern-African region, but also of international research, attracting contributions from far afield. SAICSIT does, however, not exist or operate in isolation.

More than 50 years have passed since the first electronic computer appeared in our society. In the intervening years technological development has been exponential. Over the last 20 years there has been a vast growth and pervasiveness of computing and information technology throughout the world. This has led into the expansion and consolidation of research into a diversity of new technologies and applications in diverse cultural environments. During this period huge strides have also been made in the development of computing devices. The processing speed of computers has increased thousand-fold and memory capacity from megabytes to gigabytes in the last decade alone. The Southern African region did not miss out on these developments.

It is hardly possible for such quantitative expansion not to bring a change in quality. Initially computers had been developed mainly for purposes such as automation for the improvement of processing, labour-reduction in production and automation control of machinery, with artificial intelligence, which made great strides in the 1980s, seen as the ultimate field to which computers could be applied. As we moved into the 1990s it was recognized that such an automation route was not the only direction in the improvement of computers. The expansion of processing power has enabled image data to be incorporated into computer systems, mainly for the purpose of improving human utilisation. For most computer technologies of the 1990s, including the Internet and virtual reality, automation was not the ultimate purpose. Humans were increasingly actively involved in the information-processing loop. This involvement has gradually increased as we move into the 21st century. Development of computer technology based not on automation, but on interaction, is now fully established.

The method of interaction has significantly changed as well. The expansion of computer ability means that the same function can be performed far more cheaply and on smaller computers than ever before. The advent of portable and mobile computers and pervasive computing devices is ample evidence of this. The need for users to be at the same location as a computer in order to reap the benefits of software installed on that computer is becoming an obsolete notion. Time and space are no longer constraints. One of the most discussed impacts of computing and information technology is communication and the easy accessibility of information. This changes the emphasis for research and development – issues such as cultural, political, and economic differences must, for example, be accommodated in ways that researchers have not previously considered. Our goal should be to enable users to benefit from technological advances, hence matching the skills, needs, and expectations of users of available technologies to their immense possibilities.
The conference theme for the SAICSIT 2001 Conference – *Hardware, Software and Peopleware: The Reality in the Real Millennium* – aims to reflect technological developments in all aspects related to computerised systems or computing devices, and especially reflect the fact that each influences the others.

Not only has SAICSIT come of age in the 21st century, but so has the research and development community in Southern Africa. The outstanding quality of papers submitted to SAICSIT 2001, of which only a small selection is published in this collection, illustrates both the exciting and developing nature of the field in our region. I hope that you will enjoy SAICSIT 2001 and that it will provide opportunities to cultivate and grow the seeds of discussion on innovative and new developments in computing and information technology.

Paula Kotzé  
SAICSIT President
Running this conference has been rewarding, exciting and exhausting. The response to the call for papers we sent out in March was overwhelming. We received 64 paper submissions for our main conference and twelve for the postgraduate symposium. We had a panel of internationally recognized reviewers, both local and international. The response from the reviewers was impressive – accepting a variety of papers and mostly returning the reviews long before the due date. We were struck, once again, by the sheer magnanimity of academia – as busy as we all are, we still manage to contribute fully to a conference such as SAICSIT.

After an exhaustive review process, where each paper was reviewed by at least three reviewers, the program committee accepted 26 full research papers and 14 electronic papers. Five papers were referred to the postgraduate symposium, since they represented work in progress – not yet ready for presentation to a full conference but which nevertheless represented sound and relevant research. The papers published in this volume therefore represent research of an internationally high standard and we are proud to publish it. Full electronic papers will be available on the conference web site (http://www.cs.unisa.ac.za/saicsit2001/).

Computer Science and Information Systems academics in South Africa labour under difficult circumstances. The popularity of IT courses stems from the fact that IT qualifications are in high demand in industry, which leads in turn to a shortage of IT academic staff to teach the courses, even when posts are available. The net result is that fewer people teach more courses to more students. IT departments thus rake in ever-increasing amounts of state subsidy for their universities. These profits, euphemistically labelled “contribution to overhead costs”, are deployed in various ways: cross-subsidization of non-profitable departments; maintenance of general facilities; salaries for administrative personnel, etc. Sweeteners of generous physical resources for the IT departments may be provided. We have yet to hear of a University in South Africa where significant concessions have been made in terms of industry-related remuneration. At best, small subventions are provided. As a result, shortages of quality staff remain acute in most IT departments – especially at senior teaching levels. What is even worse is that academics in these departments have to motivate the value of their conference contributions and other IT outputs to selection committees, often dominated by sceptical academic power-brokers from the more traditional departments whose continued survival is underwritten by IT’s contribution to overhead costs. 1

The papers published in this volume are conclusive evidence of the indefatigability and pertinacity of Computer Science and Information Systems academics and technologists in South Africa. We are proud to be part of such a prestigious and innovative group of people.

In conclusion, we would like to thank the conference chair, Prof Paula Kotzé, for her support. We also specially thank Prof Derrick Kourie for his substantial contribution. Finally, to all of you, contributors, presenters, reviewers and organisers – a big thank you – without you this conference could not be successful.

Enjoy the Conference!
Karen Renaud & Andries Barnard

1 This taken almost verbatim from Professor Derrick Kourie’s SACLA 2001 paper titled: “The Benefits of Bad Teaching”.
Conference Organisation

General Chair
Paula Kotzé

Programme Chairs
Karen Renaud
Andries Barnard

Organising Committee
Chairs
Lucas Venter, Alta van der Merwe

Art and Design
Tersia Parsons

Sponsor Liaison
Paula Kotzé, Chris Bornman

Secretarial & Finances
Christa Prinsloo, Elmarie Havenga

Marketing & Public Relations
Klarissa Engelbrecht, Elmarie van Solms, Adriaan Pottas, Mac van der Merwe

Audio Visual
Tobie van Dyk, Andre van der Poll, Mac van der Merwe

Program Committee

Bob Baber – McMaster University, Canada
Andries Barnard – University of South Africa
Judy Bishop – University of Pretoria
Andy Bytheway – University of the Western Cape
Andre Calitz – University of Port Elizabeth
Elsabe Cloete – University of South Africa
Carina de Villiers – University of Pretoria
Alan Dix – Lancaster University, United Kingdom
Jan Eloff – Rand Afrikaans University
Andries Engelbrecht – University of Pretoria
Chris Johnson – University of Glasgow, United Kingdom
Paul Licker – University of Cape Town
Paula Kotzé – University of South Africa
Derrick Kourie – University of Pretoria
Philip Machanick – University of the Witwatersrand
Gary Marsden – University of Cape Town
Don Petkov – University of Natal in Pietermaritzburg
Karen Renaud – University of South Africa
Ian Sanders – University of the Witwatersrand
Derrick Smith – University of Cape Town
Harold Thimbleby – Middlesex University, United Kingdom
Theda Thomas – Port Elizabeth Technikon
Herna Viktor – University of Pretoria, South Africa
Bruce Watson – Universities of Pretoria and Eindhoven
Janet Wesson – University of Port Elizabeth
Referees

Molla Alemayehu
Trish Alexander
Adi Attar
Bob Baber
Andries Barnard
John Barrow
Judy Bishop
Gordon Blair
Arina Britz
Andy Bytheway
André Calitz
Charmain Cilliers
Elsabe Cloete
Gordon Cooper
Richard Cooper
Annemieke Craig
Thad Crews
Quintin Cutts
Michael Dales
Carina de Villiers
Alan Dix
Dunlop Mark
Elize Ehlers
Jan Eloff
Andries Engelbrecht
Klarissa Engelbrecht
David Forsyth
John Galletly
Vashti Galpin
Wayne Goddard
Alexandre Hardy
Scott Hazelhurst
Johannes Heidema
Tersia Hörne
Chris Johnson
Bob Jolliffe
Paula Kotzé
Derrick Kourie
Les Labuschagne
Paul Licker
Philip Machanick
Anthony Maeder
David Manlove
Gary Marsden
Thomas Meyer
Elsa Naudé
Martin Olivier
Don Petkov
Pekka Pihlajasaaari
Nelisha Pillay
Laurette Pretorius
Karen Renaud
Ingrid Rewitzky
Sheila Rock
Markus Roggenbach
Ian Sanders
Justin Schoeman
Martie Schoeman
Elsje Scott
Derek Smith
Elmé Smith
Adrie Stander
Harold Thimbleby
Theda Thomas
Judy Van Biljon
Alta Van der Merwe
André van der Poll
Tobias Van Dyk
Lynette van Zijl
Lucas Venter
Herna Viktor
Bruce Watson
Janet Wesson

Conference
Sponsors

UNISA
ABSA Group
Microsoft
NRF

viii
Keynote Abstracts
Issues Affecting the Adoption of Data Mining in South Africa

M L Hart, K Davies, E-J Barker-Goldie, A Theron

Department of Information Systems, University of Cape Town, South Africa, *mhart@commerce.uct.ac.za

Abstract: This research describes a study into the adoption of data mining (DM) techniques and tools in industry. They are seen by most companies interviewed as strategically important, with key applications varying per sector, but have not been seriously adopted by many. A framework reflecting issues pertaining to the adoption of data mining was proposed, and tested in interviews. Issues were divided into four main categories: Business Issues; Data Mining Technology Issues; Resource Issues and Organisational Issues. Results imply widely differing levels of importance attached to these, suggesting a revised framework of key barriers and enablers for successful adoption.

Keywords: data mining, knowledge discovery, adoption, framework, business intelligence

Computing Review Categories: H.2.8, H.3.3, H.4.2, I.5.1

1. Introduction

With the pervasive use of point of sale scanners, automatic teller machines and automated order processing, data is being produced and collected at unprecedented rates [4], but many organisations lack the ability to capitalise on it. The competitive importance of information, an increasingly service-based economy and the advent of mass customisation is creating the need for companies to rapidly turn terabytes of data into significant insights to guide marketing, investment, and management strategies. They are attempting to make the most of their data assets through decision support (DSS) and business intelligence systems, including query and reporting, OLAP (Online Analytical Processing) or multi-dimensional analysis, modelling and data mining (DM) [3]. This has been facilitated partly over the past decade by the development of data warehouses and data marts [9, 11, 15, 17, 20].

2. Literature review

2.1 Data Mining (DM) defined

Many interpretations of the term “data mining” exist, exacerbated by marketing hype and vendors positioning non-DM technology as data mining [6, 24]. DM can be described as “an information extraction activity whose goal it is to discover hidden facts contained in databases, using a combination of machine learning, statistical analysis, modelling techniques and database technology, which infers rules that allow the prediction of future results” [6]. DM may alternatively be referred to as a knowledge extraction process. OLAP and DM are complementary and important parts of exploiting data [4] which should reinforce each other. OLAP and query systems rely on the analyst to pose the original hypothesis and interpret the results. Effectiveness of this analysis is however limited by the ability of the analyst to pose appropriate questions, think in non-traditional directions, and interpret the results [27]. DM allows companies more flexibility than OLAP technology. For example, they may want to know the profile of people who bought the most profitable product last year [21]. DM tools, using pattern recognition technologies, statistical and mathematical techniques, can answer this question by sifting through large amounts of data and discovering meaningful correlations, patterns and trends amongst buyers. They are used to generate a hypothesis rather than verify one. [12].

Internationally DM is used widely today [16] for customer relationship management (CRM), customer profiling, market basket analysis, and fraud detection inter alia. In CRM, DM finds patterns of product usage and consumer behaviour, helping managers understand causes of customer attrition and how to improve customer retention. Customer profiling helps identify customer groups that are more likely to respond to offers and provide better profits. [3, 18, 25]. Market-basket analysis helps retailers to understand what products or services customers tend to purchase at the same time, or in sequence. DM tools allow telecommunications firms, credit-card companies, insurance companies, stock exchanges, and government agencies to identify potentially fraudulent transactions and contain the damage [12]. Credit scoring DM applications can predict who is likely to be a good credit risk or which existing customers are likely to become delinquent in payments [7]. DM may in future be increasingly carried out on structured data over the Internet [34].

2.2 Issues pertaining to the adoption of DM

Because DM does not directly translate into improvements in operational efficiency costs, expenditure on it is difficult to justify [31]. It
can be argued that costs involved are justified by increasing revenues through more effective targeting of prospective customers, identifying cross-selling opportunities, finding growth possibilities, decreasing exposure to risk or fraud and allocating resources more efficiently [2]. While DM can provide managers with meaningful answers, it can not automatically make a company more profitable. Whether DM becomes a valuable strategic tool or an expense depends as much on the organisation using it, as on the technology itself [17].

With decreasing costs of data processing and storage even small and medium sized organisations may hold gigabytes of data. DM tools should perform acceptably on large volumes of data, regardless of the computing architecture used, and be scalable to allow for larger and more complex problems in future [29]. DM results generated may be affected by incomplete or faulty data or by subjective decisions taken on its structure. Contents of databases change constantly, raising problems in ensuring that rules are up-to-date and consistent with the most current information. The business value of DM is limited by the confidence of the user in the technology. If users don't trust it, they won't use it [1]. Data warehouses and DM tools potentially give companies weapons with which to invade individual privacy [28]. Information about an individual's purchases and sites visited can be sold without their knowledge or permission [34]. Thearling [33] expects an increased level of scrutiny of DM in terms of its impact on privacy.

DM tools are increasingly being used by business people lacking skills in analysis technologies, but Foley and Russell [16] state "In their current generation, DM tools are a little premature and require technical expertise to manipulate correctly". "You need the skills of a statistician, the skills of a computer scientist and the skills of a business analyst to frame the problem in a way that makes sense. Setting up the parameters to build a DM system requires expertise that the business user doesn't, and shouldn't be expected to have"[30]. Since DM encompasses numerous algorithms and techniques, proper deployment requires expertise in analysis, statistics, and data refinement. A DM "scientist" must be employed to obtain real value from DM tools, and these experts are in scarce supply, effectively limiting the number of organisations that can benefit [24].

As sophistication and complexity of the analysis techniques increase, the gap between the business professional and the technical analyst may continue to widen. This needs to be addressed if DM is to gain wide acceptance in the business environment, rather than being a skill offered mainly by specialists and consultants. With an increase in usability of tools, employees will be empowered to perform more extensive data analysis in their areas of business expertise. It is equally important to have an acute understanding of the business being analysed. DM is being touted as a business solution when it is simply the base technology upon which business solutions are built. DM tools merely discover patterns or co-incidences in data; only human expertise can decide on reasonable analysis and interpretation of the data and results generated. It remains the responsibility of the business to ensure rigorous testing of any model or prediction, and to decide whether to act on its "recommendations". [8, 9, 14].

3. Research Objectives

The subject of DM as an analysis technique has been covered fairly extensively in literature. Its use in South African organisations is less documented, and previous research by Hart et al [19] indicated adoption was fairly limited. It was therefore decided that an exploratory case study approach with semi-structured interviews be employed to gain a richer picture of the situation, rather than attempt a wider but possibly less informative survey.

3.1 Research problem and propositions

The research problem chosen was to identify a framework of issues pertaining to the adoption of DM in the Western Cape, looking specifically at the attitudes and perceptions of those interviewed about DM. In doing this, two propositions based on a literature study would be examined:

**Proposition 1:** DM is perceived to be a prevalent and effective technique in analysing data within Western Cape organisations

**Proposition 2:** Issues pertaining to the successful use of DM include: Business Issues, Data Mining Technology Issues, Resource Issues and Organisational Issues

A draft framework was constructed from two sources: a literature review highlighting DM issues prevalent internationally and Licker's [23] User Acceptance Model, demonstrating user characteristics influencing adoption of a
new technology: An analysis model was drawn up to outline issues pertaining to the adoption of DM, categorised into Business Issues, Technology Issues, Resource Issues and Organisational Issues. Due to space limitations the 47 issues are not listed here but are shown in later figures of results.

4. Research Methodology

4.1 Research Approach Adopted

Eisenhardt [13] defines a case study as “a research strategy which focuses on understanding the dynamics present within single settings”. The research undertaken uses a multiple case study approach; such designs are desirable when the intent of research is description, theory building or theory testing [5, 13]. The guidelines set by Lee [22] and Yin [in Licker, 23] were followed as closely as possible.

4.2 Selection of Participants in the Research Sample

Five industry sectors represented in the Western Cape (Insurance, Retail, Internet/Cellular, Consultancies and Entertainment) were chosen to cover a wide spectrum of the McFarlan grid [26], although in analysis no company fell into the turnaround quadrant. The target population of interviewees included people in Marketing, Finance, and Information Technology, consultants and executives. Interviewees were chosen at random from the companies who participated in the research. Of 24 companies contacted, 16 agreed to participate. Of these, one had no intention of doing DM, and the study concentrated on the five who were currently involved in a DM project, and the ten who were planning one.

4.3 Research Method Used

A semi-structured questionnaire was developed to guide the interviewer in discussions with respondents as there was no suitable validated research instrument available. The bulk of the questionnaire comprised open questions. It was based on the literature review and a questionnaire designed for a DM survey by the Two Crows Corporation [19]. It comprised five sections, the first used to set the interviewee at ease and obtain information regarding the company and its understanding of DM. The next section concerned organisational structure and culture, looking at the drivers of a DM project and current or planned end users of DM applications. Later questions examined resource skills available and required during a DM project. In the fourth section the framework issues pertaining to the adoption of DM were discussed, to decide whether each one should be regarded as either Very important, Important or Not Important to the organisation. Finally any other comments were recorded.

The questionnaire was piloted on academics and consultants, and some questions were altered. Slight modification was also carried out after feedback from the initial respondent. Instead of just asking “Are you using DM?” it proved preferable to ask “What is your understanding of DM?” Several companies thought they were using DM tools when in fact they were using query or OLAP tools, and it could not be easily established if they were using them in the way that had been intended. One or more people were interviewed in each company, with each participant being contacted telephonically beforehand about the type of questions to be asked. Interviews ranged from 30 minutes to two hours and all questions were answered as completely as possible (See Appendix).

4.4 Limitations of Research

The interviews were limited to 15 companies, and findings should be viewed as exploratory and strictly only applicable to the Western Cape and to those sectors included. However, the fact that people in various functional positions in a number of economic sectors were interviewed suggests that many aspects may be relevant to a broader audience. The competitive nature of the environment in which DM tools are used also meant that many companies were not willing to share information about their DM success and/or failure.

5. Findings

5.1 General Points

All companies sampled felt a need to manage and obtain value from the large amount of data accumulating in their organisations. Of the fifteen companies interviewed, fourteen believed the adoption of DM techniques to be either strategically very important or important. Five were currently involved in DM projects, and ten stated their intent to use data-mining tools in the future.

The majority of companies sampled feel that a DM project suits a flat organisational structure as this facilitates better communication between the diverse skills needed and allows more than one area of the business to benefit. Recent merging and restructuring has caused
several problems in some companies interviewed. Although data analysis is viewed as highly important, other priorities such as establishing a unified mission statement and a unified data warehouse have resulted in little present focus on DM projects. Generally companies felt that the skills required include an understanding of statistics, IT, and business knowledge, and most employed different people to fulfil each role. A common success factor was drawing these skills together in a team. The team structure facilitates “better communication as well as a tighter focus on the objectives the company wishes to attain from the project” [24]. Most are hesitant to outsource a data-mining project, feeling that consultants could never fully understand the data and company to the extent required, and it would cost more in the long run. If the DM project were to be outsourced the company would never come to grips with its data. DM was viewed by these companies as an “ongoing project that needs to be fully understood by the company”, with data having highly strategic value.

Companies have heard what international companies have done with DM, and view it as a technique to help them gain more knowledge about their customers, enabling them to become internationally competitive. They realise clients now have many choices due to globalisation and “we need to stop assuming what their needs are and gain more information about the customers we are dealing with”. Many companies feel the effective use of their voluminous data should be a management, not an IT decision. DM, according to most, is very low on the IT priority list. The IT department has been somewhat isolated from business issues in many companies, and is unable to properly evaluate the benefit of DM. A project sponsor with some goal or business problem in mind is important as the return on investment can be assessed more easily when the costs and the associated benefits from solving it can be measured.

Customer Relationship Management (CRM) in industry is currently a strong topic, in which DM can play a very important role [32]. Figure 1 demonstrates the relative importance to the companies of CRM over the other main uses of DM (Fraud Detection, Credit Scoring, and The Internet and Direct Marketing). In the insurance sector, fraud detection and credit scoring DM applications were more frequent than CRM.

Based on discussions with the respondents from the fifteen companies, the issues categorised into the four sets were evaluated as either not important, important or very important. A summary figure of each, and a brief overview of the issues deemed most important will now be given.

5.2 Business Issues

All companies sampled believe that DM should be business driven and that a strong driver of the project, coupled with management support, determines its success. Most agreed that cost was a factor, but fewer than half rated it very important. See Figure 2.

5.3 DM Technology Issues

Figure 3 reveals that the availability of a data warehouse, data integrity and age of accumulated data were key, while confusion on the real meaning of data mining made some companies uncertain as to whether they were using DM techniques or not.
5.4 Resource Issues

Teamwork was considered as an essential consideration by all companies interviewed. Ninety three percent of the companies were concerned about the availability of skills and many felt that the number of resources needed for such a project may deter companies from using DM tools. Most agreed that expenditure on training and usefulness of the tool were important. (See Figure 4)

5.5 Organisational Issues

Ninety three percent of companies recognised DM as providing a strategic advantage, and of increasing their international competitiveness. Integration across functional business units and the company infrastructure were seen as other important factors. (See Figure 5)
6. Proposition Evaluation and Framework Revision

6.1 Proposition 1: DM is perceived to be a prevalent and effective technique in analysing data within the Western Cape organisations.

From the above findings we have determined that although DM is perceived as an effective and strategic data analysis technique, it is not the most prevalent technique within the companies sampled. Therefore, we need to consider the issues pertaining to its adoption to discover the reasons for this.

6.2 Proposition 2: Issues pertaining to the successful use of DM include: Business Issues, Data Mining Technology Issues, Resource Issues and Organisational Issues.

Issues listed are either barriers that inhibit the use of DM in the Western Cape or enablers of this technology. It has been shown in the previous section that although a number are not viewed as very important or even important by many respondents, there are several from each of the four categories that clearly need to be taken strongly into account for a successful DM operation, thus
substantiating the proposition. Although the findings are based on Western Cape interviews only, all companies operate nationally, and it appears reasonable to accept that the issues strongly endorsed would carry a fair weight in a national sample of medium to large companies.

7. Conclusions & Recommendations

All companies interviewed agreed on the necessity of obtaining value from their data repositories. However, although those sampled were large enough to afford DM tools, most were still planning their full and effective use. A framework of issues drawn from published sources listed potential barriers to and enablers of DM success. The findings from the discussions with interviewees show that in the Western Cape certain issues play a far greater role than others, and some may effectively be ignored as an influencing factor. It is recommended that based on results obtained, the list of issues can in future be shortened to cover only those gaining reasonable support from this study. Ideally a further study should be undertaken over a wider geographic area and in additional economic sectors, extending too to smaller companies. Further insight is still being gleaned from deeper analysis of interview transcripts. Other questions of interest include the impact of the statistical sophistication of DM products on their use in a business environment, whether the tools are being correctly used, and characteristics of sponsors of successful DM projects. These may also be examined using Davis' Technology Acceptance Model framework.

Acknowledgement

The authors would like to thank all participants in this study for the time given to interviews and discussions, and the anonymous referees of this paper for their valuable comments.

References

[1] G.H. Anthes. Minding the storage; As the amount of digital data doubles each year, IT managers struggle to keep up. Computerworld, 22 March 1999, p.72


[16] J. Foley and D. Russel. Mining Your own Business - Vendors seek to ease deployment as more companies look to
data mining to turn data into profits. *Information Week*, 16 March 1998: 18


[34]N. Wreden. The Mother Lode - Data Mining Digs Deep For Business Intelligence. *InternetWeek*, 17 February 1997

**Appendix – Interview Promptsheet**

**Section One - Introductory Questions**
1. What is your understanding of the term Data Mining (DM)?
2. What is your perception of DM based on?
3. What do you see are the benefits (if any) of using DM over other data analysis techniques.
4. How would you classify your organisation in terms of: Small, Medium or Large? Small < 100 Medium 100 - 5000 Large 5000 - plus
5. What is your job title?
6. What were or are your expectations with regards to DM?
7. Do you believe that these expectations have been or will be met?

**Section Two - Organisational, Structural and Cultural Questions**
1. Do you have a central database repository across the functional areas of your organisation?
2. What is the communication structure within your organisation?
3. Do you perceive your organisation's technological infrastructure to be suitable for DM?
4. Who would be (or is) the driver of a DM project?
5. Who in the organisation would use (or is using) the results of a DM exercise?
6. Who are the current/planned end users of the DM applications?
7. What is the level of management commitment to DM and how important do you perceive this to be?

**Section Three - Resources**
1. What do you perceive to be the necessary skill to perform DM?
2. Are these skills currently available within your organisation or would your organisation have to acquire them?
3. Would your company consider outsourcing a DM project?
4. What percentage of the budget would be spent on a DM project?
   (I.e. how important is DM to your organisation?)

Section Four - Issues Surrounding DM
1. Having reviewed international literature regarding the adoption of DM, these were the issues that were found to be pertinent:
2. Do you feel that these issues are relevant to your organisation?
   (Each issue and its relative importance was explored)
3. Did you identify any additional issues regarding DM?

Section Five - Other Comments Made