Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Every Picture tells a Story: An Investigation of Data Models as Tools of Communication

A thesis presented in partial
Fulfillment of the requirements for the degree of
Master of Information Sciences
At Massey University, Palmerston North, New Zealand

SOMRUTAI MALAIPONG

2000

Abstract

Data models are important in information system (IS) development, particularly as tools for expressing and communicating information business requirements. The ability to understand the information content of data models is a fundamental skill required by anyone involved with them. The aim of data modelling is usually the creation of a database design but without everyone having a clear understanding of what the data model 'says', the quality of the design may suffer. When we build a model, we obviously want it to be understood but that understanding is dependent on the ability of the model to communicate its meaning. The better the model is as a vehicle of communication the clearer the understanding will be.

This research report explores this important aspect of data models and investigates the research that has been undertaken in this area including the NaLER (Natural Language for E-R) technique for reading data models. It describes an experiment conducted to explore this aspect. Subjects were tested on their ability to accurately and comprehensively interpret or 'read' a data model both before and after learning the NaLER technique. Measurement was done by using a questionnaire which consisted of two types of questions. The results show that when the subjects used NaLER, they improved their scores on the difficult questions but not the simple and medium questions. In addition, the results show that after learning NaLER, subjects' confidence in their ability to understand a model was increased, even though they actually scored less well overall.

Acknowledgements

I would like to express my deep gratitude to my supervisor, Ms. Clare Atkins, for her invaluable guidance, advice and encourage provided throughout the course of my study.

I would like to thank Mr. Barry Jackson for his support and advice.

I would like to thank to my friends for their support and commiseration.

I would also like to thank to staff of the Information System Department for their supports and advice.

A special thanks to my husband and my family for their support and encouragement they have given me throughout my study.

Finally, I would like to thank to the New Zealand Government for their financial support during my study and to the Thai Government for granting me study leave.

Table of Contents

ABSTRACT ..... **II**

ACKNOWLEDGEMENTS ..... **III**

TABLE OF CONTENTS ..... **IV**

LIST OF TABLES ..... **VII**

LIST OF FIGURES ..... **VIII**

CHAPTER ONE - INTRODUCTION **1**

CHAPTER TWO - CONTEXT **7**

2.1 TERMINOLOGY 7

2.1.1 Data Model *7*

2.1.2 Data Modelling *9*

2.1.3 Conceptual Data Model *10*

2.1.4 The Entity-Relationship Model *14*

2.1.5 The O-O model *16*

2.2 RELATED RESEARCH AREAS 17

2.2.1 Infological vs Datalogical *17*

2.2.2 Understanding graphical designs *19*

2.2.3 Communication *21*

2.3 DATA MODELLING RESEARCH 25

2.3.1 General studies *25*

2.3.2 Comparative Studies *29*

2.3.3 Data Models as Communication Tools *39*

CHAPTER THREE - READING E-R MODELS **43**

3.1 STANDARD TECHNIQUE 43

3.2 NALER TECHNIQUE 48

3.2.1 Uses of the NaLER technique *48*

3.2.2 The NaLER Construction *49*

CHAPTER FOUR - THE EXPERIMENT **55**

4.1 SECTION 1 – RESEARCH METHODOLOGY 55

4.1.1 Overview *55*

4.1.2 Reason why an experiment approach was considered the most suitable *56*

4.1.3 Research model *57*

4.1.4 Independent Variable *58*

4.1.5 Dependent variable *58*

4.1.6 Control variable *58*

4.2 SECTION 2 – RESEARCH PROCESS 59

4.2.1 The questionnaire *59*

4.2.2 Pilot study *60*

4.2.3 Select the subjects *61*

4.2.4 The procedure *61*

4.2.5 Measurement *63*

4.2.6 Limitation *63*

CHAPTER FIVE - RESULTS AND DISCUSSION **65**

5.1 RESULTS 65

5.1.1 Experiment 1 ..... *66*

5.1.2 Experiment 2 ..... *67*

5.1.3 Overall *69*

5.2 DISCUSSION 76

5.3 EVALUATION OF RESULTS AND HYPOTHESES 78

CHAPTER SIX – CONCLUSIONS AND FURTHER RESEARCH **79**

6.1 CONCLUSIONS 79

6.2 FURTHER RESEARCH 80

REFERENCES **82**

GLOSSARY **90**

APPENDICES **92**

APPENDIX 1 CASE DESCRIPTION (BATRA ET AL., 1990) 92

APPENDIX 2 ER/R MODEL 93

APPENDIX 3 QUESTIONNAIRE (COMMON SET) 94

APPENDIX 4 QUESTIONNAIRE (SET A) 96

APPENDIX 5 QUESTIONNAIRE (SET B) 98

APPENDIX 6 ETHICS COMMITTEE APPROVAL 100

APPENDIX 7 INFORMATION SHEET 109
APPENDIX 8 CONSENT FORM 111
APPENDIX 9 TRANSCRIPTS OF THE DESCRIPTIONS FROM EXPERIMENT 1 112
APPENDIX 10 TRANSCRIPTS OF THE NALER SENTENCES FROM EXPERIMENT 2 119
APPENDIX 11THE DATA RESULTS FROM EXPERIMENT 1 133
APPENDIX 12 THE DATA RESULTS FROM EXPERIMENT 2 135
APPENDIX 13 DATA ANALYSIS FROM MINITAB TRUE AND FALSE QUESTIONS 137
APPENDIX 14 DATA ANALYSIS FROM MINITAB CONFIDENCE QUESTIONS 141

List of Tables

Table 1 Data modelling studies 30
Table 2 The criteria of 3 types statements 60
Table 3 Technique 70
Table 4 Question 70
Table 5 Question*Technique 70
Table 6 Technique 71
Table 7 Question 71
Table 8 Question*Technique 71
Table 9 Technique 72
Table 10 Question 72
Table 11 Question*Technique 72
Table 12 Technique 73
Table 13 Question 73
Table 14 Question*Technique 73
Table 15 Confidence correct 74
Table 16 Confidence complete 74
Table 17 Confidence correct 75
Table 18 Confidence complete 75

List of Figures

Figure 1 ANSI/X3/SPARC Architecture 11
Figure 2 Data-driven approach to information systems design 12
Figure 3 An Analyst/User perception of the world 24
Figure 4 Three stage learning model of cognitive skills 34
Figure 5 A process model of information requirement determination 40
Figure 6 Alternative Entities Notations 43
Figure 7 Alternative Attributes Notations 44
Figure 8 Alternative Relationship Notations 45
Figure 9 Example ER diagram 51
Figure 10 NaLER sentences for E-R diagram in Figure 9 53
Figure 11 Research model 57
Figure 12 The procedure of two experiments 62
Figure 13 The correct scores (%) from experiment 1 66
Figure 14 Confidence level using the standard technique 67
Figure 15 The correct scores (%) from experiment 2 68
Figure 16 Confidence level using the NaLER technique 68
Figure 17 The correct scores (%) from experiment 1 & 2 69
Figure 18 Confidence level of two techniques and two sets of questions 69